RADIO ENGINEERING HANDBOOK

HENNEY

THIRD EDITION

McGRAW-HILL BOOK COMPANY

35

THE RADIO ENGINEERING HANDBOOK

THE
RADIO ENGINEERING
HANDBOOK

PREPARED BY A STAFF OF
TWENTY-THREE SPECIALISTS

KEITH HENNEY, EDITOR-IN-CHIEF
Member, The Institute of Radio Engineers; Author, "Principles of Radio," "Electron Tubes in Industry"; Editor, "Electronics"

THIRD EDITION
EIGHTH IMPRESSION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK AND LONDON
1941

THE RADIO ENGINEERING HANDBOOK

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

In preparing new material and in revising existing material for the Third Edition, the same principles were followed as in the First Edition. An endeavor has been made to prepare a comprehensive working manual of the radio science and to compile in a single book concise information on each of the branches of radio engineering. As in earlier editions, there is in this volume a considerable amount of what may be called fundamental background, but the emphasis is on practice rather than on theory.

Each of the sections has been brought up to date. Several have been completely rewritten, notably those on television, high-frequency technique, loud-speakers and acoustics, detection and modulation, facsimile, and aircraft radio. In each of these fields, much progress has been made since 1935. The authors of the individual sections have the requisite theoretical background as well as the very necessary practical experience in the field.

The engineer will find in this book many man-hours of effort compiled in the form of tables and curves and converted into concise English by the engineers, physicists, and teachers who have aided the editor in preparing this new edition.

KEITH HENNEY.

NEW YORK,
April, 1941.

CONTENTS

THE RADIO ENGINEERING HANDBOOK

SECTION 1

MATHEMATICAL AND ELECTRICAL TABLES

1. Greek Alphabet.

Name	Letters Cap.	Letters Small	Commonly used to designate
Alpha	Α	α	Angles. Coefficients. Area
Beta	Β	β	Angles. Coefficients
Gamma	Γ	γ	Angles. Specific gravity. Conductivity
Delta	Δ	δ	Decrements. Increments. Variation. Density
Epsilon	Ε	ε	E.m.f. Base of natural logarithms. Very small quantity
Zeta	Ζ	ζ	(Cap.) Impedance. Coordinates
Eta	Η	η	Hysteresis coefficient. Efficiency
Theta	Θ	θ, ϑ	Angular phase displacement. Time constant
Iota	Ι	ι	Current in amperes
Kappa	Κ	κ	Dielectric constant. Susceptibility. Visibility
Lambda	Λ	λ	(Small) Wave length
Mu	Μ	μ	Permeability. Amplification factor. Prefix micro-
Nu	Ν	ν	Reluctivity
Xi	Ξ	ξ	
Omicron	Ο	ο	
Pi	Π	π	Circumference divided by diameter 3.1416
Rho	Ρ	ρ	Resistivity
Sigma	Σ	σ, ς	(Cap.) Sign of summation
Tau	Τ	τ	Time constant. Time-phase displacement
Upsilon	Υ	υ	
Phi	Φ	φ, ϕ	Flux. Angle of lag or lead
Chi	Χ	χ	(Cap.) Reactance
Psi	Ψ	ψ	Angular velocity in time. Phase difference. Dielectric flux. Angles
Omega	Ω	ω	Resistance in ohms. Resistance in megohms. $2\pi F$. Angular velocity

2. Decimal Equivalents of Parts of One Inch.

1/64	0.015625	17/64	0.265625	33/64	0.515625	49/64	0.765625
1/32	0.031250	9/32	0.281250	17/32	0.531250	25/32	0.781250
3/64	0.046875	19/64	0.296875	35/64	0.546875	51/64	0.796875
1/16	0.062500	5/16	0.312500	9/16	0.562500	13/16	0.812500
5/64	0.078125	21/64	0.328125	37/64	0.578125	53/64	0.828125
3/32	0.093750	11/32	0.343750	19/32	0.593750	27/32	0.843750
7/64	0.109375	23/64	0.359375	39/64	0.609375	55/64	0.859375
1/8	0.125000	3/8	0.375000	5/8	0.625000	7/8	0.875000
9/64	0.140625	25/64	0.390625	41/64	0.640625	57/64	0.890625
5/32	0.156250	13/32	0.406250	21/32	0.656250	29/32	0.906250
11/64	0.171875	27/64	0.421875	43/64	0.671875	59/64	0.921875
3/16	0.187500	7/16	0.437500	11/16	0.687500	15/16	0.937500
13/64	0.203125	29/64	0.453125	45/64	0.703125	61/64	0.953125
7/32	0.218750	15/32	0.468750	23/32	0.718750	31/32	0.968750
15/64	0.234375	31/64	0.484375	47/64	0.734375	63/64	0.984375
1/4	0.250000	1/2	0.500000	3/4	0.750000	1	1

3. Trigonometric Functions.

° ′	sin	tan	cot	cos	
0 0	0.0000	0.0000	infinit.	1.0000	0 90
10	0.0029	0.0029	343.7737	1.0000	50
20	0.0058	0.0058	171.8854	1.0000	40
30	0.0087	0.0087	114.5887	1.0000	30
40	0.0116	0.0116	85.9398	0.9999	20
50	0.0145	0.0145	68.7501	0.9999	10
1 0	0.0175	0.0175	57.2900	0.9998	0 89
10	0.0204	0.0204	49.1039	0.9998	50
20	0.0233	0.0233	42.9641	0.9997	40
30	0.0262	0.0262	38.1885	0.9997	30
40	0.0291	0.0291	34.3678	0.9996	20
50	0.0320	0.0320	31.2416	0.9995	10
2 0	0.0349	0.0349	28.6363	0.9994	0 88
10	0.0378	0.0378	26.4316	0.9993	50
20	0.0407	0.0407	24.5418	0.9992	40
30	0.0436	0.0437	22.9038	0.9990	30
40	0.0465	0.0466	21.4704	0.9989	20
50	0.0494	0.0495	20.2056	0.9988	10
3 0	0.0523	0.0524	19.0811	0.9986	0 87
10	0.0552	0.0553	18.0750	0.9985	50
20	0.0581	0.0582	17.1693	0.9983	40
30	0.0610	0.0612	16.3499	0.9981	30
40	0.0640	0.0641	15.6048	0.9980	20
50	0.0669	0.0670	14.9244	0.9978	10
4 0	0.0698	0.0699	14.3007	0.9976	0 86
10	0.0727	0.0729	13.7267	0.9974	50
20	0.0756	0.0758	13.1969	0.9971	40
30	0.0785	0.0787	12.7062	0.9969	30
40	0.0814	0.0816	12.2505	0.9967	20
50	0.0843	0.0846	11.8262	0.9964	10
5 0	0.0872	0.0875	11.4301	0.9962	0 85
10	0.0901	0.0904	11.0594	0.9959	50
20	0.0929	0.0934	10.7119	0.9957	40
30	0.0958	0.0963	10.3854	0.9954	30
40	0.0987	0.0992	10.0780	0.9951	20
50	0.1016	0.1022	9.7882	0.9948	10
6 0	0.1045	0.1051	9.5144	0.9945	0 84
10	0.1074	0.1080	9.2553	0.9942	50
20	0.1103	0.1110	9.0098	0.9939	40
30	0.1132	0.1139	8.7769	0.9936	30
40	0.1161	0.1169	8.5555	0.9932	20
50	0.1190	0.1198	8.3450	0.9929	10
7 0	0.1219	0.1228	8.1443	0.9925	0 83
10	0.1248	0.1257	7.9530	0.9922	50
20	0.1276	0.1287	7.7704	0.9918	40
30	0.1305	0.1317	7.5958	0.9914	30
40	0.1334	0.1346	7.4287	0.9911	20
50	0.1383	0.1376	7.2687	0.9907	10
8 0	0.1392	0.1405	7.1154	0.9903	0 82
	cos	cot	tan	sin	′ °

° ′	sin	tan	cot	cos	
8 0	0.1392	0.1405	7.1154	0.9903	0 82
10	0.1421	0.1435	6.9682	0.9899	50
20	0.1449	0.1465	6.8269	0.9894	40
30	0.1478	0.1495	6.6912	0.9890	30
40	0.1507	0.1524	6.5606	0.9886	20
50	0.1536	0.1554	6.4348	0.9881	10
9 0	0.1564	0.1584	6.3138	0.9877	0 81
10	0.1593	0.1614	6.1970	0.9872	50
20	0.1622	0.1644	6.0844	0.9868	40
30	0.1650	0.1673	5.9758	0.9863	30
40	0.1679	0.1703	5.8708	0.9858	20
50	0.1708	0.1733	5.7694	0.9853	10
10 0	0.1736	0.1763	5.6713	0.9848	0 80
10	0.1765	0.1793	5.5764	0.9843	50
20	0.1794	0.1823	5.4845	0.9838	40
30	0.1822	0.1853	5.3955	0.9833	30
40	0.1851	0.1883	5.3093	0.9827	20
50	0.1880	0.1914	5.2257	0.9822	10
11 0	0.1908	0.1944	5.1446	0.9816	0 79
10	0.1937	0.1974	5.0658	0.9811	50
20	0.1965	0.2004	4.9894	0.9805	40
30	0.1994	0.2035	4.9152	0.9799	30
40	0.2022	0.2065	4.8430	0.9793	20
50	0.2051	0.2095	4.7729	0.9787	10
12 0	0.2079	0.2126	4.7046	0.9781	0 78
10	0.2108	0.2156	4.6382	0.9775	50
20	0.2136	0.2186	4.5736	0.9769	40
30	0.2164	0.2217	4.5107	0.9763	30
40	0.2193	0.2247	4.4494	0.9757	20
50	0.2221	0.2278	4.3897	0.9750	10
13 0	0.2250	0.2309	4.3315	0.9744	0 77
10	0.2278	0.2339	4.2747	0.9737	50
20	0.2306	0.2370	4.2193	0.9730	40
30	0.2334	0.2401	4.1653	0.9724	30
40	0.2363	0.2432	4.1126	0.9717	20
50	0.2391	0.2462	4.0611	0.9710	10
14 0	0.2419	0.2493	4.0108	0.9703	0 76
10	0.2447	0.2524	3.9617	0.9696	50
20	0.2476	0.2555	3.9136	0.9689	40
30	0.2501	0.2586	3.8667	0.9681	30
40	0.2532	0.2617	3.8208	0.9674	20
50	0.2560	0.2648	3.7760	0.9667	10
15 0	0.2588	0.2679	3.7321	0.9659	0 75
10	0.2616	0.2711	3.6891	0.9652	50
20	0.2644	0.2742	3.6470	0.9644	40
30	0.2672	0.2773	3.6059	0.9636	30
40	0.2700	0.2805	3.5656	0.9628	20
50	0.2728	0.2836	3.5261	0.9621	10
16 0	0.2756	0.2867	3.4874	0.9613	0 74
	cos	cot	tan	sin	′ °

° ′	sin	tan	cot	cos	
16 0	0.2756	0.2867	3.4874	0.9613	0 74
10	0.2784	0.2899	3.4495	0.9605	50
20	0.2812	0.2931	3.4124	0.9596	40
30	0.2840	0.2962	3.3759	0.9588	30
40	0.2868	0.2994	3.3402	0.9580	20
50	0.2896	0.3026	3.3052	0.9572	10
17 0	0.2924	0.3057	3.2709	0.9563	0 73
10	0.2952	0.3089	3.2371	0.9555	50
20	0.2979	0.3121	3.2041	0.9546	40
30	0.3007	0.3153	3.1716	0.9537	30
40	0.3035	0.3185	3.1397	0.9528	20
50	0.3062	0.3217	3.1084	0.9520	10
18 0	0.3090	0.3249	3.0777	0.9511	0 72
10	0.3118	0.3281	3.0475	0.9502	50
20	0.3145	0.3314	3.0178	0.9492	40
30	0.3173	0.3346	2.9887	0.9483	30
40	0.3201	0.3378	2.9600	0.9474	20
50	0.3228	0.3411	2.9319	0.9465	10
19 0	0.3256	0.3443	2.9042	0.9455	0 71
10	0.3283	0.3476	2.8770	0.9446	50
20	0.3311	0.3508	2.8502	0.9436	40
30	0.3338	0.3541	2.8239	0.9426	30
40	0.3365	0.3574	2.7980	0.9417	20
50	0.3393	0.3607	2.7725	0.9407	10
20 0	0.3420	0.3640	2.7475	0.9397	0 70
10	0.3448	0.3673	2.7228	0.9387	50
20	0.3475	0.3706	2.6985	0.9377	40
30	0.3502	0.3739	2.6746	0.9367	30
40	0.3529	0.3772	2.6511	0.9356	20
50	0.3557	0.3805	2.6279	0.9346	10
21 0	0.3584	0.3839	2.6051	0.9336	0 69
10	0.3611	0.3872	2.5826	0.9325	50
20	0.3638	0.3906	2.5605	0.9315	40
30	0.3665	0.3939	2.5386	0.9304	30
40	0.3692	0.3973	2.5172	0.9293	20
50	0.3719	0.4006	2.4960	0.9283	10
22 0	0.3746	0.4040	2.4751	0.9272	0 68
10	0.3773	0.4074	2.4545	0.9261	50
20	0.3800	0.4108	2.4342	0.9250	40
30	0.3827	0.4142	2.4142	0.9239	30
40	0.3854	0.4176	2.3945	0.9228	20
50	0.3881	0.4210	2.3750	0.9216	10
23 0	0.3907	0.4245	2.3559	0.9205	0 67
10	0.3934	0.4279	2.3369	0.9194	50
20	0.3961	0.4314	2.3183	0.9182	40
30	0.3987	0.4348	2.2998	0.9171	30
40	0.4014	0.4383	2.2817	0.9159	20
50	0.4041	0.4417	2.2637	0.9147	10
24 0	0.4067	0.4452	2.2460	0.9135	0 66
	cos	cot	tan	sin	′ °

° ′	sin	tan	cot	cos	
24 0	0.4067	0.4452	2.2460	0.9135	0 66
10	0.4094	0.4487	2.2286	0.9124	50
20	0.4120	0.4522	2.2113	0.9112	40
30	0.4147	0.4557	2.1943	0.9100	30
40	0.4173	0.4592	2.1775	0.9088	20
50	0.4200	0.4628	2.1609	0.9075	10
25 0	0.4226	0.4663	2.1445	0.9063	0 65
10	0.4253	0.4699	2.1283	0.9051	50
20	0.4279	0.4734	2.1123	0.9038	40
30	0.4305	0.4770	2.0965	0.9026	30
40	0.4331	0.4806	2.0809	0.9013	20
50	0.4358	0.4841	2.0655	0.9001	10
26 0	0.4384	0.4877	2.0503	0.8988	0 64
10	0.4410	0.4913	2.0353	0.8975	50
20	0.4436	0.4950	2.0204	0.8962	40
30	0.4462	0.4986	2.0057	0.8949	30
40	0.4488	0.5022	1.9912	0.8936	20
50	0.4514	0.5059	1.9768	0.8923	10
27 0	0.4540	0.5095	1.9626	0.8910	0 63
10	0.4566	0.5132	1.9486	0.8897	50
20	0.4592	0.5169	1.9347	0.8884	40
30	0.4617	0.5206	1.9210	0.8870	30
40	0.4643	0.5243	1.9074	0.8857	20
50	0.4669	0.5280	1.8940	0.8843	10
28 0	0.4695	0.5317	1.8807	0.8829	0 62
10	0.4720	0.5354	1.8676	0.8816	50
20	0.4746	0.5392	1.8546	0.8802	40
30	0.4772	0.5430	1.8418	0.8788	30
40	0.4797	0.5467	1.8291	0.8774	20
50	0.4823	0.5505	1.8165	0.8760	10
29 0	0.4848	0.5543	1.8040	0.8746	0 61
10	0.4874	0.5581	1.7917	0.8732	50
20	0.4899	0.5619	1.7796	0.8718	40
30	0.4924	0.5658	1.7675	0.8704	30
40	0.4950	0.5696	1.7556	0.8689	20
50	0.4975	0.5735	1.7437	0.8675	10
30 0	0.5000	0.5774	1.7321	0.8660	0 60
10	0.5025	0.5812	1.7205	0.8646	50
20	0.5050	0.5851	1.7090	0.8631	40
30	0.5075	0.5890	1.6977	0.8616	30
40	0.5100	0.5930	1.6864	0.8601	20
50	0.5125	0.5969	1.6753	0.8587	10
31 0	0.5150	0.6009	1.6643	0.8572	0 59
10	0.5175	0.6048	1.6534	0.8557	50
20	0.5200	0.6088	1.6426	0.8542	40
30	0.5225	0.6128	1.6319	0.8526	30
40	0.5250	0.6168	1.6212	0.8511	20
50	0.5275	0.6208	1.6107	0.8496	10
32 0	0.5299	0.6249	1.6003	0.8480	0 58
	cos	cot	tan	sin	′ °

° ′	sin	tan	cot	cos	′ °	° ′	sin	tan	cot	cos	′ °
32 0	0.5299	0.6249	1.6003	0.8480	0 58	39 0	0.6293	0.8098	1.2349	0.7771	0 51
10	0.5324	0.6289	1.5000	0.8465	50	10	0.6316	0.8146	1.2276	0.7753	50
20	0.5348	0.6330	1.5798	0.8450	40	20	0.6338	0.8195	1.2203	0.7735	40
30	0.5373	0.6371	1.5697	0.8434	30	30	0.6361	0.8243	1.2131	0.7716	30
40	0.5398	0.6412	1.5597	0.8418	20	40	0.6383	0.8292	1.2059	0.7698	20
50	0.5422	0.6453	1.5497	0.8403	10	50	0.6406	0.8342	1.1988	0.7679	10
33 0	0.5446	0.6494	1.5399	0.8387	0 57	40 0	0.6428	0.8391	1.1918	0.7660	0 50
10	0.5471	0.6536	1.5301	0.8371	50	10	0.6450	0.8441	1.1847	0.7642	50
20	0.5495	0.6577	1.5204	0.8355	40	20	0.6472	0.8491	1.1778	0.7623	40
30	0.5519	0.6619	1.5108	0.8339	30	30	0.6494	0.8541	1.1708	0.7604	30
40	0.5544	0.6661	1.5013	0.8323	20	40	0.6517	0.8591	1.1640	0.7585	20
50	0.5568	0.6703	1.4919	0.8307	10	50	0.6539	0.8642	1.1571	0.7566	10
34 0	0.5592	0.6745	1.4826	0.8290	0 56	41 0	0.6561	0.8693	1.1504	0.7547	0 49
10	0.5616	0.6787	1.4733	0.8274	50	10	0.6583	0.8744	1.1436	0.7528	50
20	0.5640	0.6830	1.4641	0.8258	40	20	0.6604	0.8796	1.1369	0.7509	40
30	0.5664	0.6873	1.4550	0.8241	30	30	0.6626	0.8847	1.1303	0.7490	30
40	0.5688	0.6916	1.4460	0.8225	20	40	0.6648	0.8899	1.1237	0.7470	20
50	0.5712	0.6959	1.4370	0.8208	10	50	0.6670	0.8952	1.1171	0.7451	10
35 0	0.5736	0.7002	1.4281	0.8192	0 55	42 0	0.6691	0.9004	1.1106	0.7431	0 48
10	0.5760	0.7046	1.4193	0.8175	50	10	0.6713	0.9057	1.1041	0.7412	50
20	0.5783	0.7089	1.4106	0.8158	40	20	0.6734	0.9110	1.0977	0.7392	40
30	0.5807	0.7133	1.4019	0.8141	30	30	0.6756	0.9163	1.0913	0.7373	30
40	0.5831	0.7177	1.3934	0.8124	20	40	0.6777	0.9217	1.0850	0.7353	20
50	0.5854	0.7221	1.3848	0.8107	10	50	0.6799	0.9271	1.0786	0.7333	10
36 0	0.5878	0.7265	1.3764	0.8090	0 54	43 0	0.6820	0.9325	1.0724	0.7314	0 47
10	0.5901	0.7310	1.3680	0.8073	50	10	0.6841	0.9380	1.0661	0.7294	50
20	0.5925	0.7355	1.3597	0.8056	40	20	0.6862	0.9435	1.0599	0.7274	40
30	0.5948	0.7400	1.3514	0.8039	30	30	0.6884	0.9490	1.0538	0.7254	30
40	0.5972	0.7445	1.3432	0.8021	20	40	0.6905	0.9545	1.0477	0.7234	20
50	0.5995	0.7490	1.3351	0.8004	10	50	0.6926	0.9601	1.0416	0.7214	10
37 0	0.6018	0.7536	1.3270	0.7986	0 53	44 0	0.6947	0.9657	1.0355	0.7193	0 46
10	0.6041	0.7581	1.3190	0.7069	50	10	0.6967	0.9713	1.0295	0.7173	50
20	0.6065	0.7627	1.3111	0.7051	40	20	0.6988	0.9770	1.0235	0.7153	40
30	0.6088	0.7673	1.3032	0.7034	30	30	0.7009	0.9827	1.0176	0.7133	30
40	0.6111	0.7720	1.2954	0.7016	20	40	0.7030	0.9884	1.0117	0.7112	20
50	0.6134	0.7766	1.2876	0.7898	10	50	0.7050	0.9942	1.0058	0.7092	10
38 0	0.6157	0.7813	1.2799	0.7880	0 52	45 0	0.7071	1.0000	1.0000	0.7071	0 45
10	0.6180	0.7860	1.2723	0.7862	50						
20	0.6202	0.7907	1.2647	0.7844	40						
30	0.6225	0.7954	1.2572	0.7826	30						
40	0.6248	0.8002	1.2497	0.7808	20						
50	0.6271	0.8050	1.2423	0.7790	10						
39 0	0.6293	0.8098	1.2349	0.7771	0 51						
	cos	cot	tan	sin	′ °		cos	cot	tan	sin	′ °

4. Functions of Angles in Various Quadrants.

Function	$-x$	$90° \pm x$	$180° \pm x$	$270° \pm x$	$360° \pm x$
Sin	$-\sin x$	$+\cos x$	$\mp\sin x$	$-\cos x$	$\pm\sin x$
Cos	$+\cos x$	$\mp\sin x$	$-\cos x$	$\pm\sin x$	$+\cos x$
Tan	$-\tan x$	$\mp\cot x$	$\pm\tan x$	$\mp\cot x$	$\pm\tan x$
Cot	$-\cot x$	$\mp\tan x$	$\pm\cot x$	$\mp\tan x$	$\pm\cot x$
Sec	$+\sec x$	$\mp\operatorname{cosec} x$	$-\sec x$	$\pm\operatorname{cosec} x$	$+\sec x$
Cosec	$-\operatorname{cosec} x$	$+\sec x$	$\mp\operatorname{cosec} x$	$-\sec x$	$\pm\operatorname{cosec} x$

5. Mathematical and Physical Constants.

$\pi = 3.14159$ $\log_{10} \pi = 0.49714$
$1/\pi = 0.31830$ $\log_\epsilon \pi = 1.14472$
$\pi^2 = 9.86960$ $\log_{10} 2 = 0.30102$
$\sqrt{\pi} = 1.77245$ $\log_{10} \epsilon = 0.43429$
$\epsilon = 2.71828$ $\log_\epsilon 10 = 2.30258$
$\log_\epsilon 2 = 0.69314$

$$\text{Velocity of light} = 2.99796 \times 10^{10} \text{ cm per second}$$

$$\text{Electron charge} = \begin{cases} 1.5911 \times 10^{-20} \text{ abs. e.m.u.} \\ 4.770 \times 10^{-10} \text{ abs. e.s.u.} \end{cases}$$

$$\text{Planck's constant} = h = 6.547 \times 10^{-27} \text{ erg-sec.}$$

6. Table of Circuit Constants. (Pages 6, 7, 8 and 9).

Values of ω, $1/\omega$, inductive and capacitive reactance, wave length, and LC products for frequencies from 10 cycles to 100 Mc for inductance in henrys and capacity in microfarads.

The following table, in conjunction with the multiplying factors given below, gives the values of circuit constants, for any frequency between 10 cycles and 100 mc:

MULTIPLYING FACTORS

For frequencies between	Mult. ω by	Mult. $1/\omega$ by	Mult. λ (wave length) by	Mult. LC by
10.5 cycles and 100 cycles	1.0	10^{-4}	10^5	10
105 cycles and 1,000 cycles	10.0	10^{-5}	10^4	10^{-2}
1,050 cycles and 10,000 cycles	10^2	10^{-6}	10^3	10^{-4}
10.5 kc and 100 kc	10^3	10^{-7}	10^2	10^{-6}
105 kc and 1,000 kc	10^4	10^{-8}	10^1	10^{-8}
1,050 kc and 10,000 kc	10^5	10^{-9}	1.0	10^{-10}
10.5 mc and 100 mc	10^6	10^{-10}	0.1	10^{-12}

Inductive Reactance. To obtain the inductive reactance of an inductance of L henrys at any frequency:

a. Apply the proper multiplying factor to column 2.
b. Multiply by L, the number of henrys.

Capacitive Reactance. To obtain the capacitive reactance of a condenser of C μf at any frequency:

a. Apply the proper multiplying factor to column 3.
b. Divide the result by C, the number of microfarads.
c. Multiply by 10^6.

If C is in micromicrofarads instead of microfarads, multiply by 10^{12} instead of 10^6.

Example. Thus an inductance of 250 mh at 2,500 cycles has a reactance of $250 \times 10^{-3} \times 157.08 \times 10^2 = 3{,}940$ ohms. A capacity of 250 μμf at 2,500 kc has a reactance of $10^{-9} \times 63.665 \times 10^{12} \div 250 = 254$ ohms.

Frequency	$\omega = 2\pi f$	$1/\omega = 1/2\pi f$	λ Wave length	LC
105	65.974	151.57	285.71	229.75
110	69.115	144.79	272.73	209.34
115	72.257	138.49	260.87	191.52
120	75.398	132.63	250.00	175.90
125	78.540	127.33	240.00	162.18
130	81.682	122.43	230.77	149.88
135	84.823	117.89	222.22	138.99
140	87.965	113.68	214.28	129.23
145	91.106	109.76	206.90	120.48
150	94.248	106.10	200.00	112.58
155	97.389	102.60	193.55	105.44
160	100.53	99.472	187.50	98.945
165	103.67	96.459	181.82	93.040
170	106.81	93.624	176.47	87.646
175	109.96	90.983	171.43	82.708
180	113.10	88.418	166.67	78.179
185	116.24	86.030	162.16	74.011
190	119.38	83.766	157.90	70.167
195	122.52	81.618	153.85	66.615
200	125.66	79.562	150.00	63.325
205	128.81	77.633	146.35	60.274
210	131.95	75.785	142.85	57.637
215	135.09	74.024	139.54	54.796
220	138.23	72.395	136.36	52.335
225	141.37	70.736	133.33	50.035
230	144.51	69.245	130.43	47.880
235	147.65	67.727	127.66	45.866
240	150.80	66.315	125.00	43.975
245	153.94	64.959	122.45	42.198
250	157.08	63.665	120.00	40.545
255	160.22	62.415	117.65	38.954
260	163.36	61.215	115.38	37.470
265	166.50	60.060	113.20	36.068
270	169.65	58.995	111.11	34.747
275	172.89	57.841	109.09	33.494
280	175.93	56.840	107.14	32.307
285	179.07	55.844	105.26	31.185
290	182.21	54.880	103.45	30.120
295	185.35	53.952	101.70	29.107
300	188.47	53.050	100.00	28.145
305	191.64	52.181	98.36	27.229
310	194.78	51.300	96.77	26.360
315	197.92	50.525	95.238	25.528
320	201.06	49.736	93.700	24.736
325	204.20	48.977	92.308	23.981
330	207.35	48.229	90.910	23.260
335	210.49	47.508	89.559	22.571
340	213.63	46.812	88.245	21.911
345	216.77	46.132	86.956	21.281
350	219.91	45.491	85.715	20.677
355	223.05	44.833	84.390	20.099
360	226.20	44.209	83.335	19.565
365	229.34	43.602	82.192	19.013
370	232.48	43.015	81.080	18.503
375	235.62	42.440	80.000	18.013

See multiplying factors on page 5.

Frequency	$\omega = 2\pi f$	$1/\omega = 1/2\pi f$	λ Wave length	LC
380	238.76	41.883	78.950	17.542
385	241.90	41.339	77.922	17.089
390	245.04	40.809	76.975	16.654
395	248.19	40.293	75.948	16.234
400	251.33	39.781	75.000	15.831
405	254.47	39.298	74.073	15.442
410	257.61	38.816	73.175	15.068
415	260.75	38.355	72.288	14.707
420	263.89	37.892	71.425	14.409
425	267.04	37.448	70.588	14.023
430	270.18	37.012	69.770	13.699
435	273.32	36.587	68.965	13.386
440	276.46	36.197	68.180	13.084
445	279.60	35.764	67.416	12.788
450	282.74	35.368	66.666	12.509
455	285.89	34.980	65.934	12.238
460	289.03	34.622	65.215	11.970
465	292.17	34.227	64.516	11.715
470	295.31	33.863	63.830	11.466
475	298.45	33.505	63.161	11.227
480	301.59	33.157	62.500	10.994
485	304.74	32.815	61.856	10.768
490	307.88	32.479	61.225	10.549
495	311.02	32.152	60.604	10.337
500	314.16	31.832	60.000	10.136
505	317.30	31.516	59.406	9.9322
510	320.44	31.207	58.825	9.7380
515	323.59	30.903	58.251	9.5524
520	326.73	30.607	57.690	9.3675
525	329.87	30.317	57.142	9.1898
530	333.01	30.030	56.600	9.0170
535	336.15	29.748	56.075	8.8498
540	339.29	29.497	55.555	8.6867
545	342.43	29.203	55.045	8.5276
550	345.58	28.920	54.545	8.3735
555	348.72	28.676	54.054	8.2234
560	350.86	28.420	53.570	8.0767
565	355.00	28.169	53.097	7.9348
570	358.14	27.922	52.630	7.7962
575	361.28	27.679	52.174	7.6610
580	364.43	27.440	51.725	7.5296
585	367.57	27.207	51.280	7.4013
590	370.71	26.976	50.850	7.2767
595	373.85	26.749	50.420	7.1547
600	376.99	26.525	50.000	7.0362
605	380.13	26.308	49.586	6.9200
610	383.28	26.090	49.180	6.8072
615	386.42	25.878	48.780	6.6968
620	389.56	25.650	48.385	6.5900
625	392.70	25.468	48.000	6.4844
630	395.84	25.262	47.619	6.3820
635	398.98	25.063	47.244	6.2819
640	402.12	24.868	46.850	6.1840
645	405.27	24.674	46.511	6.0885
650	408.41	24.488	46.154	5.9952

See multiplying factors on page 5.

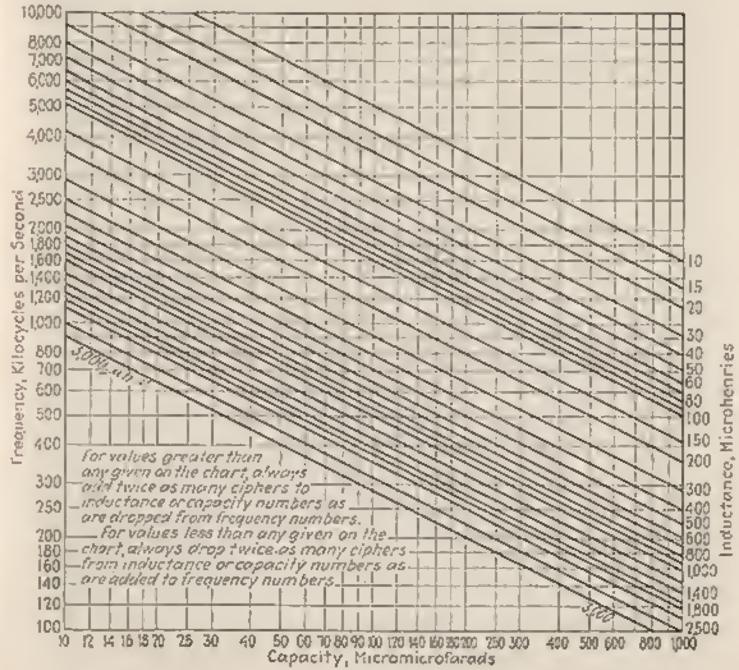
Frequency	$\omega = 2\pi f$	$1/\omega = 1/2\pi f$	λ Wave length	LC
655	411.55	24.298	45.801	5.9040
660	413.69	24.114	45.455	5.8150
665	417.83	23.933	45.113	5.7279
670	420.97	23.754	44.779	5.6425
675	424.12	23.578	44.445	5.5466
680	427.26	23.406	44.122	5.4777
685	430.39	23.238	43.796	5.3982
690	433.54	23.068	43.478	5.3202
695	436.68	22.900	43.166	5.2441
700	439.82	22.745	42.857	5.1492
705	442.97	22.575	42.553	5.0962
710	446.11	22.416	42.195	5.0247
715	449.25	22.259	41.957	4.9546
720	452.39	22.104	41.667	4.8912
725	455.53	21.953	41.379	4.8189
730	458.67	21.801	41.096	4.7532
735	461.82	21.655	40.817	4.6887
740	464.96	21.507	40.540	4.6257
745	468.10	21.363	40.268	4.5636
750	471.24	21.220	40.000	4.5032
755	474.38	21.080	39.735	4.4436
760	476.52	20.941	39.475	4.3855
765	480.67	20.804	39.215	4.3282
770	483.81	20.669	38.961	4.2722
775	486.95	20.536	38.710	4.2173
780	490.09	20.404	38.487	4.1635
785	493.23	20.275	38.216	4.1105
790	496.37	20.146	37.974	4.0585
795	499.51	20.019	37.735	4.0076
800	502.66	19.891	37.500	3.9577
805	505.80	19.770	37.267	3.9087
810	508.94	19.649	37.036	3.8605
815	512.08	19.528	36.810	3.8134
820	515.22	19.408	36.587	3.7670
825	518.36	19.292	36.364	3.7216
830	521.51	19.177	36.144	3.6767
835	524.65	19.060	35.927	3.6337
840	527.79	18.946	35.712	3.6022
845	530.93	18.835	35.502	3.5474
850	534.07	18.724	35.294	3.5062
855	537.21	18.614	35.087	3.4657
860	539.36	18.506	34.885	3.4242
865	543.50	18.399	34.682	3.3852
870	546.64	18.293	34.487	3.3465
875	549.78	18.189	34.285	3.3082
880	552.92	18.098	34.090	3.2710
885	556.06	17.988	33.898	3.2341
890	558.92	17.882	33.708	3.1970
895	562.35	17.783	33.520	3.1622
900	565.49	17.689	33.333	3.1272
905	568.63	17.586	33.150	3.0928
910	571.77	17.490	32.967	3.0595
915	574.91	17.378	32.787	3.0254
920	578.05	17.311	32.607	2.9925
925	581.20	17.206	32.432	2.9604

See multiplying factors on page 5.

Frequency	$\omega = 2\pi f$	$1/\omega = 1/2\pi f$	λ Wave length	LC
930	584.34	17.113	32.258	2.9287
935	587.48	17.022	32.086	2.8974
940	590.62	16.931	31.915	2.8665
945	593.76	16.842	31.746	2.8364
950	596.90	16.752	31.580	2.8067
955	600.05	16.665	31.414	2.7774
960	602.19	16.578	31.250	2.7485
965	606.33	16.492	31.088	2.7200
970	609.47	16.407	30.928	2.6920
975	612.61	16.324	30.770	2.6646
980	615.75	16.239	30.617	2.6372
985	618.90	16.158	30.456	2.6108
990	622.04	16.071	30.302	2.5842
995	625.18	15.995	30.150	2.5586
1000	628.32	15.916	30.000	2.5330

See multiplying factors on page 5.

7. L, C, λ Chart.



8. Dimensions, Weights, and Resistances of Pure, Solid, Bare Copper Wire.

(Copper-wire Tables, *Circ.* 31, Bur. Standards.)

B. & S. or American wire gage	Diam. in mils at 20°C. (68°F.)	Cross-sectional area at 20°C. (68°F.): Circular mils (d^2), 1 m = 0.001 in.	Cross-sectional area: Square inches	Carrying capacities: Rubber insul. amps. A	Carrying capacities: Varn. cloth insul. amps. B	Carrying capacities: Other insul. amps. C	Weight: Pounds per 1,000 ft.	Weight: Pounds per mile
0000	460.0	211,600.0	0.166,2	225	270	325	640.5	3,381.840
000	409.6	167,800.0	0.131,8	175	210	275	507.9	2,681.712
00	364.8	133,100.0	0.104,5	150	180	225	402.8	2,126.784
0	324.9	105,500.0	0.082,89	125	150	200	319.5	1,686.960
1	289.3	83,690.0	0.065,73	100	120	150	253.3	1,337.424
2	257.6	66,370.0	0.052,13	90	110	125	200.9	1,060.752
3	229.4	52,640.0	0.041,34	80	95	100	159.3	841.104
4	204.3	41,740.0	0.032,78	70	85	90	126.4	667.392
5	181.9	33,100.0	0.026,00	55	65	80	100.2	529.056
6	162.0	26,250.0	0.020,62	50	60	70	79.46	419.548.8
7	144.3	20,820.0	0.016,35	38	...	54	63.02	332.745,6
8	128.5	16,510.0	0.012,97	35	40	50	49.98	263.894,4
9	114.4	13,090.0	0.010,28	28	...	38	39.63	209.246,1
10	101.9	10,380.0	0.008,155	25	30	30	31.43	165.950,4
11	90.74	8,234.0	0.006,467	20	...	27	24.02	131.577,6
12	80.81	6,530.0	0.005,129	20	25	25	19.77	104.385,6
13	71.96	5,178.0	0.004,067	17	...	...	15.08	82.790,4
14	64.08	4,107.0	0.003,225	15	18	20	12.43	65.630,4
15	57.07	3,257.0	0.002,558	...	...	...	9.858	52.050,24
16	50.82	2,583.0	0.002,028	6	...	10	7.818	41,279,04
17	45.26	2,048.0	0.001,609	...	...	...	6.200	32,736,00
18	40.30	1,624.0	0.001,276	3	...	6	4.917	25.961,76
19	35.89	1,288.0	0.001,012	...	...	...	3.899	20.586,72
20	31.96	1,022.0	0.000,802,3	The above values are those specified in the 1931 National Electrical Code. In lighting work, no wire smaller than No. 14 is used, except in fixtures			3.092	16.325,70
21	28.46	810.1	0.000,636,3				2.452	12.946,56
22	25.35	642.4	0.000,504,6				1.945	10.269,60
23	22.57	509.5	0.000,400,2				1.542	8.141,76
24	20.10	404.0	0.000,317,3				1.223	6.457,44
25	17.90	320.4	0.000,251,7				0.969,9	5.121,072
26	15.94	254.1	0.000,199,6				0.769,2	4.061,376
27	14.20	201.5	0.000,158,3				0.610,0	3.220,800
28	12.64	159.8	0.000,125,5				0.483,7	2.553,936
29	11.26	126.7	0.000,099,53				0.383,6	2.025,408
30	10.03	100.5	0.000,078,94	....			0.304,2	1.606,176
31	8.928	79.70	0.000,062,60	....			0.241,3	1.274,060
32	7.950	63.20	0.000,049,64	....			0.191,3	1.010,064
33	7.080	50.13	0.000,039,37	....			0.151,7	0.800,976
34	6.305	39.75	0.000,031,22	....			0.120,3	0.635,184
35	5.615	31.52	0.000,024,76	....			0.095,42	0.513,717,6
36	5.000	25.00	0.000,019,64	....			0.075,68	0.399,590,4
37	4.453	19.83	0.000,015,57	....			0.060,01	0.316,852,8
38	3.965	15.72	0.000,012,35	....			0.047,59	0.251,275,2
39	3.531	12.47	0.000,009,793	....			0.037,74	0.199,267,2
40	3.145	9.888	0.000,007,766	....			0.029,93	0.158,030,4

Length, 25°C. (77°F.): Feet per pound	Length: Feet per ohm	Resistance at 25°C. (77°F.): R ohms per 1,000 ft.	Resistance: Ohms per mile	Resistance: Ohms per pound	B. & S. or American wire gage
1.561	20,010.0	0.049,98	0.263,894,4	0.000,078,03	0000
1.968	15,870.0	0.063,02	0.332,745,6	0.000,124,1	000
2.482	12,580.0	0.079,47	0.419,501,6	0.000,197,3	00
3.130	9,980.0	0.100,2	0.529,056	0.000,313,7	0
3.947	7,914.0	0.126,4	0.667,392	0.000,498,8	1
4.977	6,270.0	0.159,3	0.841,104	0.000,793,1	2
6.276	4,977.0	0.200,9	1.060,752	0.001,261	3
7.914	3,947.0	0.253,3	1.337,424	0.002,005	4
9.980	3,130.0	0.319,5	1.686,960	0.003,188	5
12.58	2,482.0	0.402,8	2.126,784	0.005,069	6
15.87	1,969.0	0.508,0	2.682,240	0.008,061	7
20.01	1,561.0	0.640,5	3.381,840	0.012,82	8
25.23	1,238.0	0.807,7	4.264,656	0.020,38	9
31.82	981.8	1.018	5.375,04	0.032,41	10
40.12	778.7	1.284	6.779,52	0.051,53	11
50.59	617.5	1.619	8.548,32	0.081,93	12
63.80	489.7	2.042	10.781,76	0.130,3	13
80.44	388.3	2.575	13.596,00	0.207,1	14
101.4	308.0	3.247	17.144,16	0.329,4	15
127.9	244.2	4.094	21.616,32	0.523,7	16
161.3	193.7	5.163	27.260,64	0.832,8	17
203.4	153.6	6.510	34.372,80	1.324	18
256.5	121.8	8.210	43.348,80	2.105	19
323.4	96.60	10.35	54.648.0	3.348	20
407.8	76.61	13.05	68.904,0	5.323	21
514.2	60.75	16.46	86.908,8	8.464	22
648.4	48.18	20.76	109.612,8	13.46	23
817.7	38.21	26.17	138.177,6	21.40	24
1,031.0	30.30	33.00	174.240.0	34.03	25
1,300.0	24.03	41.62	219.753,6	54.11	26
1,639.0	19.06	52.48	277.094,4	86.03	27
2,067.0	15.11	66.17	349.377,6	136.8	28
2,607.0	11.98	83.44	440.563,2	217.5	29
3,287.0	9.504	105.2	555.456	345.9	30
4,145.0	7.537	132.7	700.656	549.9	31
5,227.0	5.977	167.3	883.344	874.4	32
6,591.0	4.740	211.0	1,114.080	1,390.0	33
8,310.0	3.759	266.0	1,404.480	2,211.0	34
10,480.0	2.981	335.5	1,771.440	3,515.0	35
13,210.0	2.364	423.0	2,233,440	5,590.0	36
16,600.0	1.875	533.4	2,816.352	8,888.0	37
21,010.0	1.487	672.6	3,551.328	14,130.0	38
26,500.0	1.179	848.1	4,477.968	22,470.0	39
33,410.0	0.935	1,069.0	5,644.32	35,730.0	40

9. Tensile Strength of Pure Copper Wire in Pounds.

Size, B. & S. gage	Hard drawn		Annealed		Size, B. & S. gage	Hard drawn		Annealed	
	Actual	Average per square inch	Actual	Average per square inch		Actual	Average per square inch	Actual	Average per square inch
0000	8,260	49,700	5,320	32,000	7	1050.0	64,200	556.0	34,000
000	6,550	49,700	4,220	32,000	8	843.0	65,000	441.0	34,000
00	5,440	52,000	3,340	32,000	9	678.0	66,000	350.0	34,000
0	4,530	54,600	2,650	32,000	10	546.0	67,000	277.0	34,000
1	3,680	56,000	2,100	32,000	12	343.0	67,000	174.0	34,000
2	2,970	57,000	1,670	32,000	14	219.0	68,000	110.0	34,000
3	2,380	57,600	1,323	32,000	16	138.0	68,000	68.9	34,000
4	1,900	58,000	1,050	32,000	18	86.7	68,000	43.4	34,000
5	1,580	60,800	884	34,000	19	68.8	68,000	34.4	34,000
6	1,300	63,000	700	34,000	20	54.7	68,000	27.3	34,000

10. Insulated Copper Wire.

Size, B. & S. gage	Enamel wire			Single-silk covered			Double-silk covered		
	Outside diameter, mils	Turns per linear inch	Pounds per 1,000 ft.	Outside diameter, mils	Turns per linear inch	Pounds per 1,000 ft.	Outside diameter, mils	Turns per linear inch	Pounds per 1,000 ft.
8	130.0	7.7	50.6						
9	116.5	8.6	40.2						
10	104.0	9.6	31.8						
11	92.7	10.8	25.3						
12	82.8	12.1	20.1						
13	74.0	13.5	15.90						
14	66.1	15.1	12.60						
15	59.1	16.9	10.00						
16	52.8	18.9	7.930	52.8	18.9	7.89	54.6	18.3	8.00
17	47.0	21.3	6.275	47.3	21.1	6.26	49.1	20.4	6.32
18	42.1	23.8	4.980	42.4	23.6	4.97	44.1	22.7	5.02
19	37.7	26.5	3.955	37.9	26.4	3.94	39.7	25.2	3.99
20	33.7	29.7	3.135	34.0	29.4	3.13	35.8	28.0	3.17
22	26.9	37.2	1.970	27.3	36.6	1.98	29.1	34.4	2.01
24	21.5	46.5	1.245	22.1	45.3	1.25	23.9	41.8	1.27
26	17.1	58.5	0.785	17.9	55.9	0.791	19.7	50.8	0.810
28	13.6	73.5	0.494	14.6	68.5	0.498	16.4	61.0	0.514
30	10.9	91.7	0.311	12.0	83.3	0.316	13.8	72.5	0.333
32	8.7	115	0.196	9.9	101	0.210	11.8	84.8	0.217
34	6.9	145	0.123	8.3	121	0.129	10.1	99.0	0.141
36	5.5	180	0.078	7.0	143	0.082	8.8	114	0.092
38	4.4	227	0.049	6.0	167	0.053	7.8	128	0.062
40	3.5	286	0.031	5.1	196	0.035	6.9	145	0.043

11. Insulated Copper Wire.

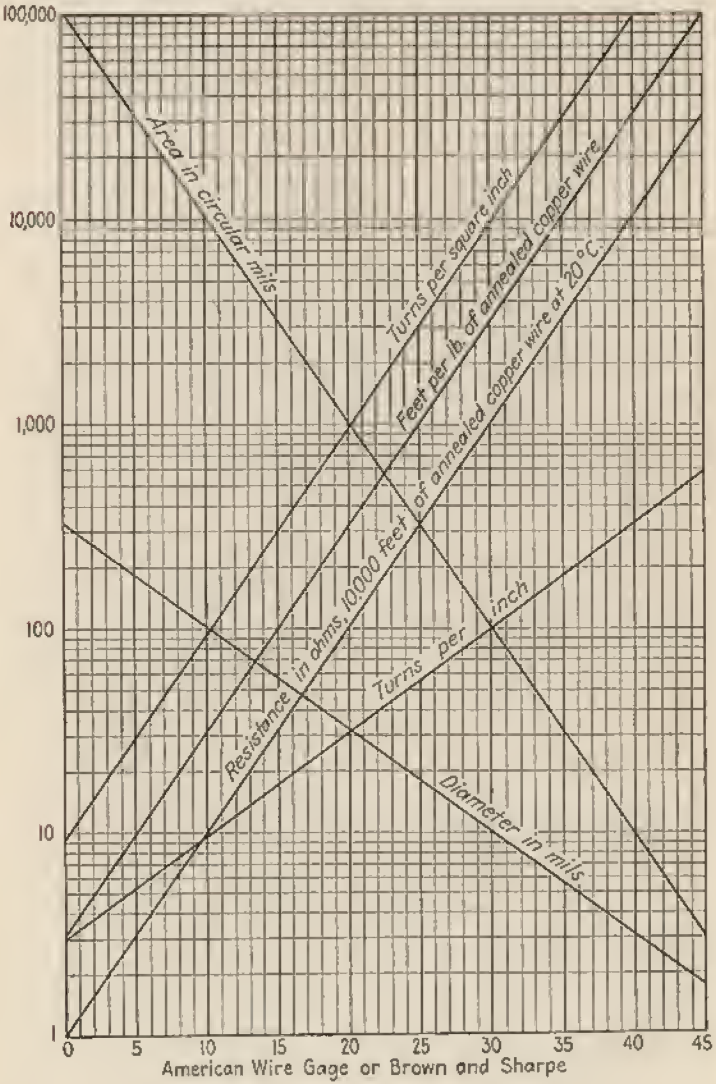
Size, B. & S. gage	Ohms per 1,000 ft.	Single-cotton covered			Double-cotton covered		
		Outside diameter, mils	Turns per linear inch	Pounds per 1,000 ft.	Outside diameter, mils	Turns per linear inch	Pounds per 1,000 ft.
0000	0.0500	467	2.14	......	477	2.10	
000	0.0630	418	2.39	......	428	2.34	
00	0.0795	373	2.68	......	382	2.62	
0	0.100	334	3.00	......	343	3.00	
1	0.126	300	3.33	......	308	3.25	
2	0.159	267	3.75	......	275	3.64	
3	0.201	239	4.18	......	248	4.03	
4	0.253	214	4.67	......	222	4.51	
5	0.319	192	5.21	......	200	5.00	
6	0.403	170	5.88	......	175	5.62	
7	0.508	153	6.54	......	160	6.25	
8	0.641	136	7.35	50.6	142	7.05	51.2
9	0.808	121	8.26	40.2	127	7.87	40.6
10	1.02	108	9.25	31.9	113	8.85	32.2
11	1.28	97	10.3	25.3	102	9.80	25.6
12	1.62	87	11.5	20.1	92	10.9	20.4
13	2.04	78	12.8	16.0	82	12.2	16.2
14	2.58	70	14.3	12.7	74	13.5	12.9
16	4.1	56	17.9	8.03	60	16.7	8.21
18	6.5	45	22.2	5.08	49	20.4	5.24
20	10.4	37	27	3.22	41	24.4	3.37
22	16.6	29.5	33.9	2.05	33.3	30.0	2.17
24	26.2	24.1	41.5	1.3	28.1	35.6	1.4
26	41.6	19.9	50.2	0.834	23.9	41.8	0.914
28	66.2	16.6	60.2	0.533	20.6	48.6	0.608
30	105	14	71.4	0.340	18.0	55.6	0.400
32	167	12	83.4	0.223	16.0	62.9	0.270
34	266	10.3	97.1	0.148	14.3	70.0	0.193
36	423	9.0	111	0.099	13.0	77.0	0.136
38	673	8.0	125	0.070	12.0	83.3	0.105
40	1,070	7.1	141	0.052	11.1	90.9	0.084

12. Properties of Commercial Insulating Oils.[1]

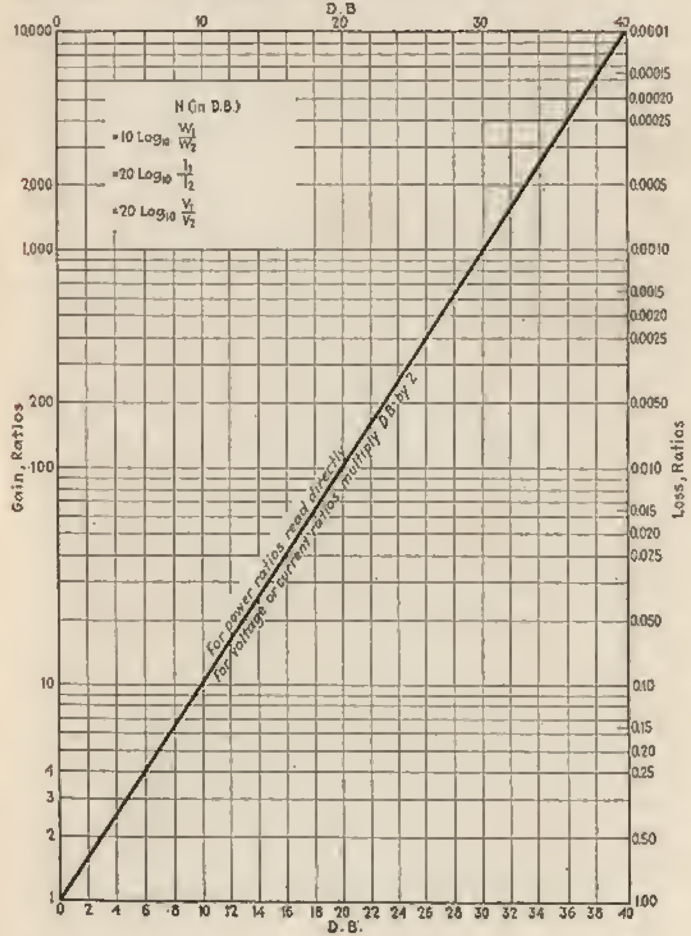
Oil	Dielectric constant	Resistivity at 500 volts d.c. 100°C., ohm-cm	Power factor 100°C.	Dielectric strength, 25°C.; 0.1-in. gap, kv
Mineral oil	2.23	21.0×10^{12}	0.0004	30 to 40
Whale oil	3.05	0.032×10^{12}	0.0015	30 to 40
Linseed oil	3.3	0.61×10^{12}	0.0027	30 to 40
Castor oil	4.7	0.066×10^{12}	0.0070	30 to 40
Cottonseed oil	3.2	0.01×10^{12}	0.0005	30 to 40
China wood oil	3.2	0.08×10^{12}	0.0090	30 to 40

[1] CLARK, F. M., Liquids as Insulators, *Gen. Elec. Rev.*, April, 1928.

13. Wire Table Chart.



14. Chart for Converting Loss or Gain into Decibels.



15. Logarithms of Numbers.

N	0	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996

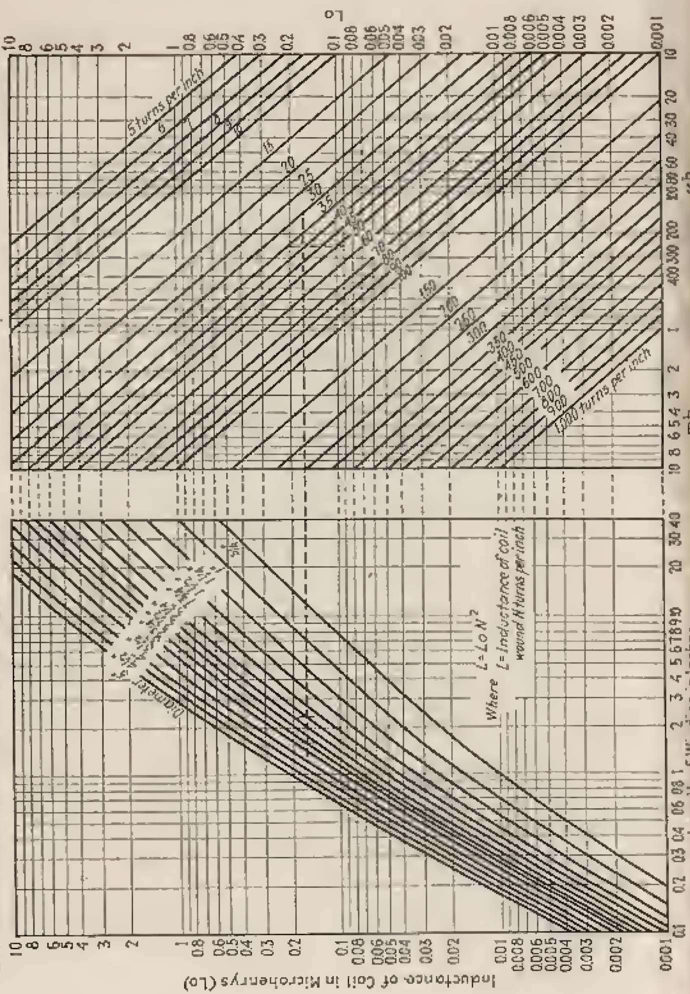
16. Exponential and Hyperbolic Functions.

$e = 2.71828$; $\frac{1}{e} = 0.36787$; $\log_{10} e = 0.43429$; $\sinh x = \frac{e^x - e^{-x}}{2}$; $\cosh x = \frac{e^x + e^{-x}}{2}$

x	Natural values					$\log_{10}$			
	e^x	e^{-x}	sinh x	cosh x	tanh x	e^x	sinh x	cosh x	tanh x
0.00	1.000	1.0000	0.000	1.000	0.0000	0.0000	∞	0.0000	∞
0.10	1.1052	0.9048	0.100	1.005	0.0997	0.0434	$\bar{1}$.0007	0.0022	$\bar{2}$.9986
0.20	1.2214	0.8187	0.201	1.020	0.1974	0.0869	$\bar{1}$.3039	0.0086	$\bar{1}$.2953
0.30	1.3499	0.7408	0.304	1.045	0.2913	0.1303	$\bar{1}$.4836	0.0193	$\bar{1}$.4644
0.40	1.4918	0.6703	0.411	1.081	0.3796	0.1737	$\bar{1}$.6136	0.0336	$\bar{1}$.5797
0.50	1.6487	0.6065	0.521	1.128	0.4621	0.2172	$\bar{1}$.7169	0.0522	$\bar{1}$.6647
0.60	1.8221	0.5488	0.637	1.186	0.5371	0.2606	$\bar{1}$.8040	0.0739	$\bar{1}$.7300
0.70	2.0138	0.4966	0.759	1.255	0.6044	0.3040	$\bar{1}$.8800	0.0987	$\bar{1}$.7813
0.80	2.2255	0.4493	0.888	1.337	0.6640	0.3474	$\bar{1}$.9485	0.1263	$\bar{1}$.8222
0.90	2.4596	0.4066	1.026	1.433	0.7163	0.3909	0.0114	0.1563	$\bar{1}$.8551
1.00	2.7183	0.3679	1.175	1.543	0.7616	0.4343	0.0701	0.1884	$\bar{1}$.8817
1.10	3.0042	0.3329	1.335	1.669	0.8005	0.4777	0.1257	0.2223	$\bar{1}$.9034
1.20	3.3201	0.3012	1.509	1.811	0.8337	0.5212	0.1788	0.2578	$\bar{1}$.9210
1.30	3.6693	0.2725	1.698	1.971	0.8617	0.5646	0.2300	0.2947	$\bar{1}$.9354
1.40	4.0552	0.2466	1.904	2.151	0.8854	0.6080	0.2797	0.3326	$\bar{1}$.9471
1.50	4.4817	0.2231	2.129	2.352	0.9052	0.6514	0.3282	0.3715	$\bar{1}$.9567
1.60	4.9530	0.2019	2.376	2.578	0.9217	0.6949	0.3758	0.4112	$\bar{1}$.9646
1.70	5.4739	0.1827	2.646	2.828	0.9354	0.7383	0.4225	0.4515	$\bar{1}$.9710
1.80	6.0496	0.1653	2.942	3.108	0.9468	0.7817	0.4687	0.4924	$\bar{1}$.9763
1.90	6.6859	0.1496	3.268	3.412	0.9562	0.8252	0.5143	0.5337	$\bar{1}$.9806
2.00	7.3891	0.1353	3.627	3.762	0.9640	0.8686	0.5595	0.5754	$\bar{1}$.9841
2.10	8.1662	0.1225	4.022	4.144	0.9705	0.9120	0.6044	0.6175	$\bar{1}$.9870
2.20	9.0250	0.1108	4.457	4.568	0.9757	0.9554	0.6491	0.6597	$\bar{1}$.9893
2.30	9.9742	0.1003	4.937	5.037	0.9801	0.9989	0.6935	0.7022	$\bar{1}$.9913
2.40	11.023	0.0907	5.466	5.557	0.9837	1.0423	0.7377	0.7448	$\bar{1}$.9928
2.50	12.182	0.0821	6.050	6.132	0.9866	1.0857	0.7818	0.7876	$\bar{1}$.9942
2.60	13.464	0.0743	6.695	6.770	0.9890	1.1292	0.8257	0.8305	$\bar{1}$.9952
2.70	14.880	0.0672	7.406	7.473	0.9910	1.1726	0.8696	0.8735	$\bar{1}$.9961
2.80	16.445	0.0608	8.192	8.253	0.9926	1.2160	0.9134	0.9166	$\bar{1}$.9968
2.90	18.174	0.0550	9.050	9.115	0.9940	1.2595	0.9571	0.9597	$\bar{1}$.9974
3.00	20.086	0.0498	10.018	10.068	0.9951	1.3029	1.0008	1.0029	$\bar{1}$.9979
3.10	22.198	0.0451	11.077	11.122	0.9960	1.3463	1.0444	1.0462	$\bar{1}$.9982
3.20	24.533	0.0408	12.246	12.287	0.9967	1.3897	1.0880	1.0894	$\bar{1}$.9986
3.30	27.113	0.0369	13.538	13.575	0.9973	1.4332	1.1316	1.1327	$\bar{1}$.9988
3.40	29.964	0.0334	14.965	14.999	0.9978	1.4766	1.1751	1.1761	$\bar{1}$.9990

x	Natural values					$\log_{10}$			
	e^x	e^{-x}	sinh x	cosh x	tanh x	e^x	sinh x	cosh x	tanh x
3.50	33.115	0.0302	16.543	16.573	0.9982	1.5200	1.2186	1.2194	$\bar{1}$.9992
3.60	36.598	0.0273	18.285	18.313	0.9985	1.5635	1.2621	1.2628	$\bar{1}$.9994
3.70	40.447	0.0247	20.211	20.236	0.9988	1.6069	1.3056	1.3061	$\bar{1}$.9995
3.80	44.701	0.0224	22.339	22.362	0.9990	1.6503	1.3491	1.3495	$\bar{1}$.9996
3.90	49.402	0.0202	24.691	24.711	0.9992	1.6938	1.3925	1.3929	$\bar{1}$.99964
4.00	54.598	0.0183	27.290	27.308	0.9993	1.7372	1.4360	1.4363	$\bar{1}$.99971
4.10	60.340	0.0166	30.162	30.178	0.99945	1.7806	1.4795	1.4797	$\bar{1}$.99976
4.20	66.686	0.0150	33.336	33.351	0.99955	1.8240	1.5229	1.5231	$\bar{1}$.99980
4.30	73.700	0.0136	36.843	36.857	0.99963	1.8675	1.5664	1.5665	$\bar{1}$.99984
4.40	81.451	0.0123	40.719	40.732	0.99970	1.9109	1.6098	1.6099	$\bar{1}$.99987
4.50	90.017	0.0111	45.003	45.014	0.99975	1.9543	1.6532	1.6534	$\bar{1}$.99989
4.60	99.484	0.0101	49.737	49.747	0.99980	1.9976	1.6967	1.6968	$\bar{1}$.99991
4.70	109.95	0.0091	54.969	54.978	0.99983	2.0412	1.7401	1.7402	$\bar{1}$.99993
4.80	121.51	0.0082	60.751	60.759	0.99986	2.0846	1.7836	1.7836	$\bar{1}$.99994
4.90	134.29	0.0075	67.141	67.149	0.99989	2.1280	1.8270	1.8270	$\bar{1}$.99995
5.00	148.41	0.0067	74.203	74.210	0.99991	2.1715	1.8704	1.8704	$\bar{1}$.99996
5.10	164.02	0.0061	82.008	82.014	0.99993	2.2149	1.9137	1.9139	$\bar{1}$.99997
5.20	181.27	0.0055	90.633	90.639	0.99994	2.2583	1.9573	1.9573	$\bar{1}$.99997
5.30	200.34	0.0050	100.17	100.17	0.99995	2.3018	2.0007	2.0007	$\bar{1}$.99998
5.40	221.41	0.0045	110.70	110.71	0.99996	2.3452	2.0442	2.0442	$\bar{1}$.99998
5.50	244.69	0.0041	122.34	122.35	0.99997	2.3886	2.0876	2.0876	$\bar{1}$.99999
5.60	270.43	0.0037	135.21	135.22	0.99997	2.4321	2.1310	2.1310	$\bar{1}$.99999
5.70	298.87	0.0034	149.43	149.44	0.99998	2.4755	2.1744	2.1744	$\bar{1}$.99999
5.80	330.30	0.0030	165.15	165.15	0.99998	2.5189	2.2179	2.2179	$\bar{1}$.99999
5.90	365.04	0.0027	182.52	182.52	0.99998	2.5623	2.2613	2.2613	$\bar{1}$.99999
6.00	403.43	0.0025	201.71	201.72	0.99999	2.6058	2.3047	2.3047	$\bar{1}$.99999

17. Inductance of Various Windings.



18. Systems of Electrical Units.

Practical	E.m.u.	E.s.u.	R.f.	A.f.
Volt(v)	10^{-8} v	300 v	v	v
Ampere(a)	10 a	3.33×10^{-10} a	ma	ma
Second	sec.	sec.	μsec.	msec.
Cycle	cycle	cycle	Mc	kc
Ohm	10^{-9} ohm	0.9×10^{12} ohm	k-ohm	k-ohm
Mho	10^{9} mho	1.11×10^{-12} mho	m-mho	m-mho
Henry(h)	10^{-9} h (cm)	0.9×10^{12} h	mh	h
Farad(f)	10^{9} f	1.1 μμf (cm)	mμf	μf
Watt(w)	10^{-7} w	10^{-7} w	mw	mw
Joule(j)	10^{-7} j (erg)	10^{-7} j	mμj	μj
Coulomb(c)	10 c	3.33×10^{-10} c	mμc	μc

$\mu = 10^{-6}$; $m = 10^{-3}$; $k = 10^{3}$; $M = 10^{6}$; $m\mu = 10^{-9}$.

19. Computing the Harmonic Content of Any Given Periodic Complex Wave Form. When an oscillogram (or other graphical representation) of a periodic complex wave is available, it is possible to compute the percentage of each harmonic up to and including the sixth, by means of the following scheme:[1]

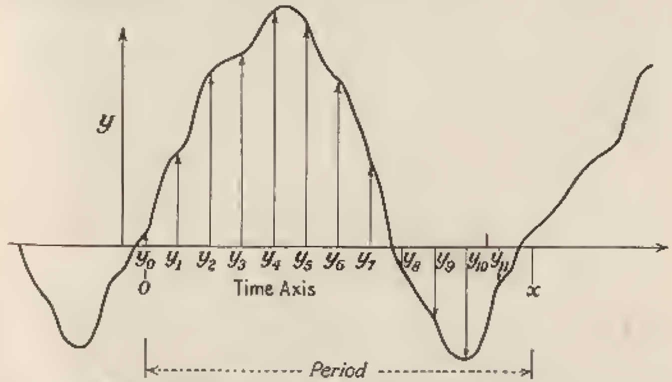


FIG. 1.—Example of complex wave for analysis.

The oscillogram must contain at least one complete period of the wave, *i.e.*, from any given point on the wave to the corresponding point at the left or right at which the form of the wave begins to repeat itself. In Fig. 1 the complete period is given by the distance OX, a distance of 360 electrical degrees. With a compass or dividers, divide this complete period into 12 equal parts, and erect the 12 equally spaced ordinates $y_0, y_1, y_2, \ldots, y_{11}$. Each of these vertical lines is drawn from the

[1] This method is known as the *twelve ordinate scheme*, and is a convenient form for solving the equations of the Fourier analysis. The form given here has been adapted from "Graphical and Mechanical Computation," Part II, Experimental Data, by Joseph Lipka, published by John Wiley & Sons, Inc., New York, pp. 181–185. See also Terebesi, "Rechenschablonen für harmonische Analyse und Synthese," Julius Springer, Berlin, 1930.

horizontal time axis to the curve. With a rule (preferably one divided into tenths of inches or a millimeter rule, so that the lengths can be expressed in decimal form), measure the length of each of these ordinates. It makes no difference whether inches, millimeters, or any other arbitrary unit is used, so long as all ordinates are measured with the same unit. Record the length of each ordinate in the spaces given in the table below:

Ordinate number	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}	y_{11}
Length of ordinate	5.5	37.0	68.6	76.4	93.2	89.6	66.7	34.3	−8.8	−28.4	−44.1	−15.0

The lengths given are the lengths taken from Fig. 1.

The computation consists in substituting these lengths in the following schedule of additions, subtractions, and multiplications, and in performing the indicated operations. First set down the values of the ordinates in the following arrangement, adding and subtracting as indicated:

$$
\begin{array}{lccccccc}
 & y_0 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \\
 & & y_{11} & y_{10} & y_9 & y_8 & y_7 & \\
\hline
\text{Sum:} & s_0 & s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\
\text{Difference:} & & d_1 & d_2 & d_3 & d_4 & d_5 &
\end{array}
$$

Then take the sum terms in the following arrangement:

$$
\begin{array}{lcccc}
 & s_0 & s_1 & s_2 & s_3 \\
 & s_6 & s_5 & s_4 & \\
\hline
\text{Sum:} & S_0 & S_1 & S_2 & S_3 \\
\text{Difference:} & D_0 & D_1 & D_2 &
\end{array}
$$

Take the difference terms in this arrangement:

$$
\begin{array}{lccc}
 & d_1 & d_2 & d_3 \\
 & d_5 & d_4 & \\
\hline
 & S_4 & S_5 & S_6 \\
 & D_3 & D_4 &
\end{array}
$$

Finally:

$$
\begin{array}{lcc}
 & S_0 & S_1 \\
 & S_2 & S_3 \\
\hline
\text{Sum:} & S_7 & S_8
\end{array}
\quad \text{and} \quad
\begin{array}{lcc}
 & S_4 & D_0 \\
 & S_6 & D_2 \\
\hline
\text{Difference:} & D_5 & D_6
\end{array}
$$

We are now in a position to find the coefficients in the equation of the complex wave. This equation is written:

$$
y = A_0 + A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + A_4 \cos 4\omega t + A_5 \cos 5\omega t + A_6 \cos 6\omega t + B_1 \sin \omega t + B_2 \sin 2\omega t + B_3 \sin 3\omega t + B_4 \sin 4\omega t + B_5 \sin 5\omega t
$$

where A and B are the coefficients of the cosine and sine terms, respectively.

The formulas for the A's and B's are as follows:

$$
A_0 = \frac{S_7 + S_8}{12};\ A_1 = \frac{D_0 + 0.866D_1 + 0.5D_2}{6};\ A_2 = \frac{S_0 + 0.5S_1 - 0.5S_2 - S_3}{6}
$$

$$
A_3 = \frac{D_6}{6};\ A_4 = \frac{S_0 - 0.5S_1 - 0.5S_2 + S_3}{6};\ A_5 = \frac{D_0 - 0.866D_1 + 0.5D_2}{6}
$$

$$
A_6 = \frac{S_7 - S_8}{12};\ B_1 = \frac{0.5S_4 + 0.866S_5 + S_6}{6};\ B_2 = \frac{0.866(D_3 + D_4)}{6}
$$

$$
B_3 = \frac{D_5}{6};\ B_4 = \frac{0.866(D_3 - D_4)}{6};\ B_5 = \frac{0.5S_4 - 0.866S_5 + S_6}{6}
$$

There are several checks which may be made on the arithmetic of the above computations:

$$
y_0 = A_0 + A_1 + A_2 + A_3 + A_4 + A_5 + A_6
$$

$$
y_1 - y_{11} = (B_1 + B_5) + \sqrt{3}(B_2 + B_4) + 2B_3
$$

For computing the percentage harmonic content of the wave, it is convenient to express the equation of the wave in somewhat simpler form, reducing the cosine terms to sine terms in the following manner:

$$
\begin{aligned}
y = A_0 &+ \sqrt{A_1^2 + B_1^2}\sin(\omega t + \alpha_1) + \sqrt{A_2^2 + B_2^2}\sin(2\omega t + \alpha_2) + \\
&\sqrt{A_3^2 + B_3^2}\sin(3\omega t + \alpha_3) + \sqrt{A_4^2 + B_4^2}\sin(4\omega t + \alpha_4) + \\
&\sqrt{A_5^2 + B_5^2}\sin(5\omega t + \alpha_5) + A_6 \sin(6\omega t + \alpha_6)
\end{aligned}
$$

The coefficient of each sine term in the above equation is proportional to the magnitude of the harmonic, that is, $\sqrt{A_1^2 + B_1^2}$ is the amplitude of the fundamental, $\sqrt{A_2^2 + B_2^2}$ the amplitude of the second harmonic (double frequency), $\sqrt{A_3^2 + B_3^2}$ the amplitude of the third harmonic (triple frequency), and so on. A_0 is the d-c component of the wave, ω is equal to $2\pi f$, where f is the fundamental frequency. The angles α_1, α_2, α_3, etc., are equal to $\tan^{-1}\frac{A_1}{B_1}$, $\tan^{-1}\frac{A_2}{B_2}$, etc. These angles do not enter into the computation, unless the phase displacements between the various harmonics are desired.

To find the percentages of the various harmonics, in terms of the magnitude of the fundamental, use the following expressions:

Per cent second harmonic:

$$
\text{Per cent} = \frac{\sqrt{A_2^2 + B_2^2}}{\sqrt{A_1^2 + B_1^2}} \times 100 \text{ per cent}
$$

For the third harmonic:

$$
\text{Per cent} = \frac{\sqrt{A_3^2 + B_3^2}}{\sqrt{A_1^2 + B_1^2}} \times 100 \text{ per cent}
$$

and so on. For all harmonics up to the sixth taken together, the total harmonic content expressed as a percentage is:

$$
\text{Per cent} = \frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + A_5^2 + A_6^2 + B_2^2 + B_3^2 + B_4^2 + B_5^2}}{\sqrt{A_1^2 + B_1^2}} \times 100 \text{ per cent}
$$

It is sometimes useful to compare the r-m-s value of the fundamental with the d-c component, expressed as a percentage. To obtain this percentage from above figures, substitute in the following expression:

D-c component, expressed as a per cent of r-m-s fundamental,

$$
= \frac{A_0}{0.707\sqrt{A_1^2 + B_1^2}} \times 100 \text{ per cent}
$$

Example (see Fig. 1 and values in table above):

	5.5	37.0	68.6	76.4	93.2	89.6	66.7
		−15.0	−44.1	−28.4	−8.8	34.3	
Sum:	5.5	22.0	24.5	48.0	84.4	123.9	66.7
	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Difference:		52.0	112.7	104.8	102.0	55.3	
		d_1	d_2	d_3	d_4	d_5	

	5.5	22.0	24.5	48.0
	66.7	123.9	84.4	
	72.2	145.9	108.9	48.0
	S_0	S_1	S_2	S_3
	−61.2	−101.9	−59.9	
	D_0	D_1	D_2	

	52.0	112.7	104.8
	55.3	102.0	
Sum	107.3	214.7	104.8
	S_4	S_5	S_6
Difference:	−3.3	10.7	
	D_3	D_4	

	72.2	145.9
	108.9	48.0
Sum:	181.1	193.9
	S_7	S_8

	107.3	−61.2
	104.8	−59.9
Difference:	2.5	−1.3
	D_5	D_6

$$A_0 = \frac{181.1 + 193.9}{12} = +31.3$$

$$A_1 = \frac{-61.2 + 0.866(-101.9) + 0.5(-59.9)}{6} = -29.6$$

$$A_2 = \frac{72.2 + 0.5(145.9) - 0.5(108.9) - 48.0}{6} = +7.1$$

$$A_3 = \frac{-1.3}{6} = -0.2$$

$$A_4 = \frac{72.2 - 0.5(145.9) - 0.5(108.9) + 48.0}{6} = -1.2$$

$$A_5 = \frac{-61.2 - 0.866(-101.9) + 0.5(-59.9)}{6} = -0.4$$

$$A_6 = \frac{181.1 - 193.9}{12} = -1.1$$

$$B_1 = \frac{0.5(107.3) + 0.866(214.7) + 104.8}{6} = +57.3$$

$$B_2 = \frac{0.866(-3.3 + 10.7)}{6} = +1.1$$

$$B_3 = \frac{2.5}{6} = +0.4$$

$$B_4 = \frac{0.866(-3.3 - 10.7)}{6} = -2.0$$

$$B_5 = \frac{0.5(107.3) - 0.866(214.7) + 104.8}{6} = -4.5$$

Result:

$$\begin{aligned} y = 31.3 &- 29.6 \cos \omega t + 7.1 \cos 2\omega t - 0.2 \cos 3\omega t \\ &- 1.2 \cos 4\omega t - 0.4 \cos 5\omega t - 1.1 \cos 6\omega t \\ &+ 57.3 \sin \omega t + 1.1 \sin 2\omega t + 0.4 \sin 3\omega t \\ &- 2.0 \sin 4\omega t - 4.5 \sin 5\omega t \end{aligned}$$

Percentage of various harmonics:

$$\text{Second: Per cent} = \frac{\sqrt{(7.1)^2 + (1.1)^2}}{\sqrt{(29.6)^2 + (57.3)^2}} \times 100 \text{ per cent} = 11.1 \text{ per cent}$$

$$\text{Third: Per cent} = \frac{\sqrt{(0.2)^2 + (0.4)^2}}{64.5} \times 100 \text{ per cent} = 0.7 \text{ per cent}$$

$$\text{Fourth: Per cent} = \frac{\sqrt{(1.2)^2 + (2.0)^2}}{64.5} \times 100 \text{ per cent} = 3.6 \text{ per cent}$$

$$\text{Fifth: Per cent} = \frac{\sqrt{(0.4)^2 + (4.5)^2}}{64.5} \times 100\ \% = 7.0 \text{ per cent}$$

$$\text{Sixth: Per cent} = \frac{1.1}{64.5} \times 100\ \% = 1.7 \text{ per cent}$$

Total harmonic content:

$$\text{Per cent} = \frac{\sqrt{(7.1)^2 + (0.2)^2 + (1.2)^2 + (0.4)^2 + (1.1)^2 + (1.1)^2 + (0.4)^2 + (2.0)^2 + (4.5)^2}}{64.5}$$

$$= 13.8 \text{ per cent}$$

Percentage d-c component:

$$\text{Per cent} = \frac{31.3}{0.707(64.5)} = 68.9 \text{ per cent}$$

20. Evaluation of Square Root of the Sum of the Squares of Two Numbers. In the calculation of impedance as the square root of the sum of the squares of a reactance and a resistance, a useful and convenient method of solution consists in rewriting the equation as follows.

$$\sqrt{a^2 + b^2} = b\sqrt{1 + \frac{a^2}{b^2}}$$

where a is the large number.

The operations can now be carried out fairly simply with the slide rule. If the right-hand side of this equation be multiplied and divided by a/b the solution becomes simply one of multiplying the larger number a by a factor which is a function of the ratio of a/b.

A table may be worked out for this function. W. J. Seeley of Duke University, Durham, N. C., has copyrighted such a table in which the factor has been worked out to five decimal places for various values of a/b from 0.001 to 30. Curves may be drawn from calculations of this nature which will be useful in graphically determining the value of the function a/b.

21. Shunt and Multiplier Data for Meters. It is often useful to convert a low-reading current meter to a voltmeter or a current meter of higher maximum current reading. The following table will cover the usual situations arising in the average laboratory. The values of shunt are calculated from the equation for meter shunts,

$$\frac{Rm \times Im}{I - Im}$$

where Rm = meter resistance in ohms
Im = full-scale current of meter
I = current desired to be read.

SHUNT AND MULTIPLIER VALUES
27-ohm (0-1) Milliammeter

Scale	Use as	Resistance in ohms of multiplier or shunt		Multiply old scale by
0-10	Voltmeter	10,000	M	10
0-50	Voltmeter	50,000	M	50
0-100	Voltmeter	100,000	M	100
0-250	Voltmeter	250,000	M	250
0-500	Voltmeter	500,000	M	500
0-1000	Voltmeter	1,000,000	M	1000
0-10	Milliammeter	3	S	10
0-50	Milliammeter	0.551	S	50
0-100	Milliammeter	0.272	S	100
0-500	Milliammeter	0.0541	S	500
35-ohm (0-1.5) Milliammeter				
0-15	Voltmeter	10,000	M	10
0-150	Voltmeter	100,000	M	100
0-750	Voltmeter	500,000	M	500
0-15	Milliammeter	3.89	S	10
0-75	Milliammeter	0.714	S	50
0-150	Milliammeter	0.351	S	100
0-750	Milliammeter	0.0701	S	500

SECTION 2

ELECTRIC AND MAGNETIC CIRCUITS

BY E. A. UEHLING[1]

FUNDAMENTALS OF ELECTRIC CIRCUITS

1. Nature of Electric Charge. According to modern views all natural phenomena may be explained on the basis of fundamental postulates regarding the nature of electric charge. In the neighborhood of an electric charge is postulated the existence of an electric field to explain such phenomena as repulsion and attraction. The force which acts between electric charges by virtue of the electric fields surrounding them is expressed by Coulomb's law which states that

$$F = \frac{q_1 q_2}{r^2}$$

The value of the unit charge in the electrostatic system is based on this law and is defined, therefore, as that value of electric charge which when placed at 1 cm distance from an equal charge repels it with a force of 1 dyne.

2. Electrons and Protons. There are two types of electricity: positive and negative. The electron is representative of the latter and the proton of the former. All matter is made up simply of electrons and protons. Exhaustive experiment has proved that all electrons, no matter how derived, are identical in nature. They are easily isolated and as a consequence have been thoroughly studied. Among the most important results of this study are the following facts:[2]

Charge of the electron	4.770×10^{-10} e.s.u.
Mass	9.04×10^{-28} g
Radius	2×10^{-13} cm, approx.

The proton has not been so thoroughly studied. It is not so easily isolated, and the effects of electric and magnetic fields on its motion are considerably smaller than similar effects obtained when electrons are studied. The proton apparently has a mass of about 1,838 times that of the electron and a considerably smaller radius.

The mass of electrons and protons is purely inertial in character. In other words these fundamental units of electric charge consist simply of pure electricity. For the sake of completeness it should be added that this mass is not independent of velocity and that the values given for both the electron and proton assume velocities which are small in comparison with that of light.

[1] Department of Physics, University of Washington.
[2] MILLIKAN, R. A., "The Electron."

3. Atomic Structure. The atoms of matter consist of a central positive nucleus surrounded by such a number of electrons as will neutralize the nuclear charge. The central positive nucleus consists of both electrons and protons with an excess of the latter. This excess determines the chemical characteristics of the atom by determining the number of electrons outside the nucleus, while the total number of protons determines the atomic weight of the element. According to one view the electrons outside the nucleus move in planetary elliptic orbits about it. The radius of the different orbits varies within a single atom, and as a consequence the strength of the bond existing between the nucleus and the different electrons varies.

4. Ionization. The outer electrons are in general loosely bound to the nucleus and under favorable conditions may be completely dissociated from the remainder of the atom. This process of the removal of an electron is known as *ionization.* It is the process by which electrons are removed from a heated filament in a vacuum tube, from an alkali metal surface in the photoelectric cell, and from the plate and grid of vacuum tubes when bombarded by the filament electrons giving rise to the secondary emission so commonly experienced.

5. The Nature of Current. The modern view of electricity regards a current as a flow of negative charge in one direction plus a flow of positive charge in the opposite direction. In electrolytic conduction the unit of negative charge is an atom with one or more additional electrons called a *negative ion,* and the unit of positive charge is an atom with one or more electrons less than its normal number known as the *positive ion.*

In conduction through gases, as, for example, through the electric arc, the negative ion is usually a single electron, whereas the positive ion is as before an atom with one or more electrons removed.

In conduction through solids, however, the current is strictly electronic and is not made up of two parts as in the previous cases. The electrons constituting the current are the outer orbital electrons of the atoms. Since these electrons are less tightly bound to the atom than the other electrons they are comparatively free and are often spoken of as *free electrons.* These electrons move through the solid under the influence of an electric field colliding with the atoms as they move and continuously losing energy gained from the field. As a consequence the motion of the electrons in the direction of the field is of a comparatively small velocity[1] (of the order of 1 cm per second), whereas the velocity of thermal agitation of the free electrons is high (about 10^7 cm per second). According to this view of the electric current in solids, conductors and insulators differ only in the relative number of free electrons possessed by the substance.

Since current consists of a motion of electric charges, it may be defined as a given amount of charge passing a point in a conductor per unit time. In the electrostatic system the unit of current is defined to be a current such that an electrostatic unit of electricity crosses any selected cross section of a conductor in unit time. In the practical system the unit of current is the *ampere* which is approximately equal to 3×10^9 electrostatic units of current and is defined on the basis of material constants as that current which will deposit 0.00111800 g of silver from a solution of silver nitrate in 1 sec.

[1] JEANS, J. H., "Electricity and Magnetism," p. 306.

6. The Nature of Potential. An electric charge that is resident in an electric field experiences a force of repulsion or attraction depending on the nature of the charge. Its position in the field may be considered as representing a certain quantity of potential energy which may be taken as the amount of work which is capable of being done when the electric charge moves from the point in question to an infinite distance. If the convention of considering a unit positive charge as the test charge is adopted, the potential energy at a point may be taken as characteristic of the field and consequently will be regarded simply as the potential.

In a similar manner the difference of potential of two points may be described as the amount of work required to move a unit positive test charge from one point to another. More specifically a difference of potential in a conductor may be spoken of as equal to the energy dissipated when an electron moves through the conductor from the point of low potential to the point of high potential. This energy is dissipated in the form of heat caused by the bombardment of the molecules of the conductor by the electrons as they proceed from one point to another.

7. Concept of E.M.F. The idea of potential leads directly to a conception of an electromotive force. If a difference of potential between two points of a conductor is maintained by some means or other, electrons will continue to flow, giving rise to a continuous current. A difference in potential maintained in this way while the current is flowing is known as an *electromotive force.* Only two important methods of maintaining a constant e.m.f. exist: the battery and the generator. Other methods, as, for example, the thermocouple, are not primarily intended for the purpose of maintaining a current.

The unit of e.m.f. in the practical system is the *volt.* It is defined as 10^8 e.s.u. of potential or as 1.0000/1.0183 of the voltage generated by a standard Weston cell.

8. Ohm's Law and Resistance. The free electrons which contribute to the electric current have a low drift velocity in the negative direction of the field within the conductor. In moving through the metal in a common general direction they enter into frequent collisions with the molecules of the metal, and as a consequence they are continually retarded in their forward motion and are not able to attain a velocity greater than a certain terminal velocity u, which depends on the value of the field and the nature of the substance. The collisions which tend to reduce the drift velocity of the electrons act as a retarding force. When a current is flowing, this retarding force must be exactly equal to the accelerating force of the field. The retarding force is proportional to N, the number of free electrons per unit length of conductor, and to u, their drift velocity. It may be designated as kNu. The accelerating force is proportional to the field E per unit length of conductor, to the number N of electrons per unit length, and to the electronic charge e and may be represented as NEe. Then $NEe = kNu$. Since the current i has been given as

$$i = Neu$$

$$NEe = k\frac{i}{e}$$

$$E = \frac{k}{Ne^2}i = Ri$$

where

$$R = \frac{k}{Ne^2}$$

The statement $E = Ri$ is known as Ohm's law. R is here defined as the *resistance* per unit length. The unit of resistance is the *ohm*. It may be obtained from Ohm's law when the e.m.f. is expressed in volts and the current in amperes.

9. Inductance. Circuits possess inductance by virtue of the electromagnetic field which surrounds a conductor carrying a current. The coefficient of self-inductance is defined as the total number of lines of force passing through a circuit and due entirely to one c.g.s. unit of current traversing the circuit. If N is the number of lines of force linked with any circuit of inductance L and conveying C c.g.s. units of current, $N = LC$.

The practical unit of inductance is the *henry*. It is equal to 10^9 c.g.s. units of inductance. If the number of lines of force N through a circuit is changed, an e.m.f. due to this change of flux is induced in the circuit. This e.m.f. is given by the equation

$$e = -\frac{dN}{dt} = -L\frac{dC}{dt}$$

The inductance of a circuit is equal to 1 henry if an opposing e.m.f. of 1 volt is set up when the current in the circuit varies at the rate of 1 amp. per second.

10. Mutual Inductance. The coefficient of mutual inductance is defined in the same way as that of self-inductance and is given in c.g.s. units as the total magnetic flux which passes through one circuit when the other is traversed by one c.g.s. unit of current, or

$$N = MC$$

$$e = -\frac{dN}{dt} = -M\frac{dC}{dt}$$

The practical unit is the henry as in self-inductance.

11. Energy in Magnetic Field. Energy is stored in the electromagnetic field surrounding a circuit representing the energy accumulated during the time when the free electrons were initially set in motion and the current established. This energy is given by the equation, $W = \frac{1}{2}LI^2$, where, if L is in henrys and I in amperes, the energy is in *joules*.

12. Capacitance. The ratio of the *quantity* of charge on a conductor to the *potential* of the conductor represents its *capacity*. If one conductor is at zero potential and another at the potential V, the capacity is given as the ratio of the charge stored to the potential difference of the conductors

$$C = \frac{Q}{V}$$

If Q is in coulombs (the quantity of charge carried by 1 amp. flowing for 1 sec.) and V is in volts, C is known as the *farad*.

The energy stored in a condenser is given by the equation, $W = \frac{1}{2}CV^2$, where, if V is in volts and C is in farads, W is in joules.

The force acting per unit area on the conductors of the condenser tending to draw them together is

$$F = \frac{E^2}{8\pi} = \frac{V^2}{8\pi d^2}$$

where d is the distance separating the condenser plates, and V is the potential difference.

Other expressions relating charge or current to capacity and potential difference are

$$V = \frac{\int i dt}{C}$$

and

$$i = C\frac{dV}{dt}$$

13. Units. The practical units that have been described are related to the electrostatic units as shown by the following table. A third set of units, known as the electromagnetic, is also related to the practical units, the ratios of which are given in this table.

Quantity	Name of unit	Measure in electromagnetic units	Measure in electrostatic units
Charge of electricity	Coulomb	10^{-1}	3×10^9
Potential	Volt	10^8	$\frac{1}{300}$
Capacity	Farad	10^{-9}	9×10^{11}
Current	Ampere	10^{-1}	3×10^9
Resistance	Ohm	10^9	$\frac{1}{9} \times 10^{-11}$
Inductance	Henry	10^9	

14. Continuous and Alternating Currents. If the free electrons of a conductor move with a constant drift velocity under the impelling force of an invariant electric field, the electric current in the conductor is spoken of as being *continuous*, or *direct*. If, however, the impressed electric field is varying in both direction and magnitude, the drift velocity of the electrons will vary in both direction and magnitude, since electrons always flow in a direction opposite to that of the electric field. A current of this kind which varies periodically with the time is known as an *alternating* current.

15. Wave Form. The current or the e.m.f. may be represented graphically as a function of the time by assigning to successive values of the latter variable the value of the former. There is an infinite variety of functional relationships between current and time, but of all the laws by which these two variables may be connected there is one that can be differentiated from all others. This law is that of the *sine* or *cosine* function. All other relationships can be resolved into a linear combination of functions of this simple type.

The form of the sine function is shown in Fig. 1*a*. It is represented analytically by the following type of equations

$$i = I_0 \sin \omega t$$

$$e = E_0 \sin \omega t$$

where i and e are the instantaneous values of the current and voltage, I_0 and E_0 are the maximum values, and ω is 2π times the frequency with which the current or voltage alternates. The sine wave is the ideal toward which practical types approach more or less closely. Since it cannot be resolved into other types, it is the pure wave form.

16. Harmonics. Current and voltage waves, in practice, are not pure and may therefore be resolved into a series of sine or cosine functions. One of the functions into which the original wave is resolved will have a frequency term equal to that of the original wave. All of the other functions will have frequency terms of higher value, which will in general be designated as *harmonics* of the lowest or fundamental frequency. A few types of complex waves which may be resolved into two or more pure sine waves are shown in Fig. 1*b* and *c*. The resolution of a complex wave into its component parts may be accomplished physically as well as mathematically. This may be demonstrated by means of high- and low-pass filters in the output circuit of an ordinary vacuum-tube oscillator.

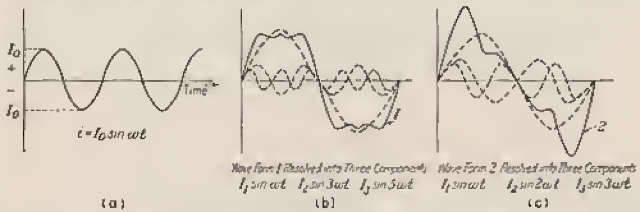


FIG. 1.—Sine wave and complex waves.

17. Effective and Average Values. The *effective value* of an a-c wave is the value of continuous current which gives the same power dissipation as the a. c. in a resistance. For a sine wave this value of continuous current is equal to the maximum value divided by $\sqrt{2}$. The average value of an alternating current is equal to the integral of the current over the time for one-half period divided by the elapsed time. For a sine wave the average value is equal to the maximum value of the current divided by $\pi/2$. The ratio of the effective value of the current to the average value is often taken as the *form factor* of the wave. Thus all types of waves may be simply characterized by means of this ratio.

Direct-current meters read average values of currents over a complete period. Such meters therefore read zero in an a-c circuit. Thermocouple and hot-wire-type meters read effective values. Such meters are therefore used for making a-c measurements at radio- as well as at audio-frequencies.

18. Phase. The current in a circuit may have its maximum and zero values at the same time as those of the e.m.f. wave, or these values may occur earlier or later than those of the latter. These three cases are illustrated in Fig. 2. When the corresponding values of the current and e.m.f. occur at the same time they are said to be in phase. If the current values occur before the corresponding values of the voltage wave, the current is said to be in *leading* phase, and if these values occur after the corresponding values of the voltage wave, it is said to be in *lagging* phase.

19. Power. The power consumed in a continuous-current circuit is $W = EI = I^2R$, where R is the *effective resistance* of the circuit. The power consumed in an a-c circuit having negligible inductance and

capacitance is given by the same equation with the necessary restrictions on I so that it represents the effective value of the current and not the average value. The power consumed in an inductive or capacitative circuit is $W = EI \cos \varphi$, where φ is the *phase angle*, that is, the angle of lag or lead of current. The term "$\cos \varphi$" is commonly referred to as the *power factor* of the circuit.

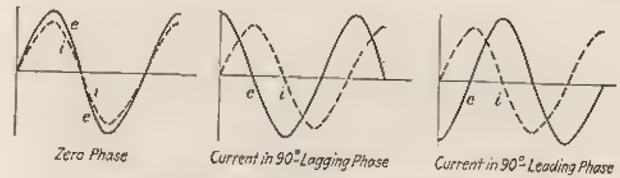


FIG. 2.—Phase in a-c circuits.

DIRECT-CURRENT CIRCUITS

20. Direction of Current Flow. An electric current is a flow of electric charges. Electric charges will move through a medium of finite resistance if a difference of electric potential exists between two points of that medium. In metallic conductors there is but one type of charge which is free to move, the negative charge or the free electrons of the conductor. The current in a metallic conductor then consists solely of an electron current. The convention arose historically of speaking of an electric current as flowing from the high potential (positive) to the low potential (negative) point, while, as a matter of fact, the electrons of the conductor actually move in the opposite direction. It is necessary to distinguish, therefore, between the direction of current flow in the historical sense and the direction of flow of electrons.

21. Constant Positive Resistance, Negative Resistance, and Infinite Resistance. In a d-c circuit the relationship between voltage and current is governed solely by the resistance of the circuit and all equivalent resistances such as counter e.m.fs. Some knowledge regarding the nature of this resistance is needed. Three cases present themselves. In the first case are those circuits in which

$$\frac{de}{di} = R$$

where R is positive and is constant in value over a rather large range. Conduction in solids and electrolytes is of this type. In the second class are those circuits in which de/di has a value which is negative and is usually not constant. Conduction in arcs and glow discharges is generally of this type. In the third class are those circuits in which

$$\frac{de}{di} = \infty$$

Conduction in the plate circuit of a vacuum tube under saturation conditions is of this type.

Circuits of the first class, in which the differential coefficient de/di has a positive value, may be subdivided into two other classes. If the value of de/di is constant over the entire range of voltage and current from zero to the maximum value, and if this value is designated by the quantity R, then Ohm's law may be used and $e = iR$. In this case, R is both the d-c and a-c resistance. If, however, R is not constant over this range of values, the value of R given at a particular value of e and i given by the equation

$$R = \frac{de}{di}$$

is only the a-c resistance of the circuit at the particular value of e and i chosen. The a-c resistance given by this equation may be quite different from the d-c value as given by the equation

$$R = \frac{e}{i}$$

In a vacuum-tube plate circuit the d-c value of the resistance is frequently about twice as high as the a-c value.

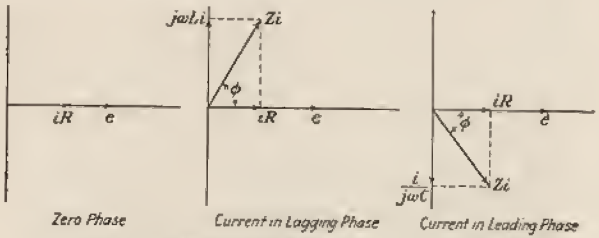


FIG. 3.—Vector representation of a-c circuits.

ALTERNATING-CURRENT CIRCUITS

22. Impedance. The resistance to the flow of an electric current having the value $i = I_0 \sin \omega t$ depends on the circuit element through which the current is passing. In a pure resistance the potential fall would be $E_1 = I_0 R \sin \omega t$, which is seen to be in phase with the current passing through it. In an inductance the potential fall would be

$$E_2 = L\frac{di}{dt} = \omega L I_0 \cos \omega t = j\omega L I_0 \sin \omega t = j\omega L i$$

and therefore leads the current by a phase angle of 90 deg. In a capacitance the potential fall would be

$$E_3 = \frac{1}{C}\int i\,dt = -\frac{I_0}{\omega C}\cos \omega t = -\frac{jI_0}{\omega C}\sin \omega t$$
$$= -\frac{ji}{\omega C}$$
$$= \frac{i}{j\omega C}$$

and is therefore led by the current by a phase angle of 90 deg. The potential fall through all three elements taken together is equal to

$$E = \left(R + j\omega L + \frac{1}{j\omega C}\right)i$$

The coefficient of i is termed the *impedance* of the circuit. It is written, in general, as

$$z = R + j\omega L + \frac{1}{j\omega C} = R + j\left(\omega L - \frac{1}{\omega C}\right)$$

where R is the total series resistance of the circuit, L is the total series inductance, and C is the effective series capacitance. The term involving j is of special importance, for it is this term which gives to the current its leading or lagging characteristics depending on whether ωL is smaller or larger than $1/\omega C$. This quantity is known as the circuit *reactance*

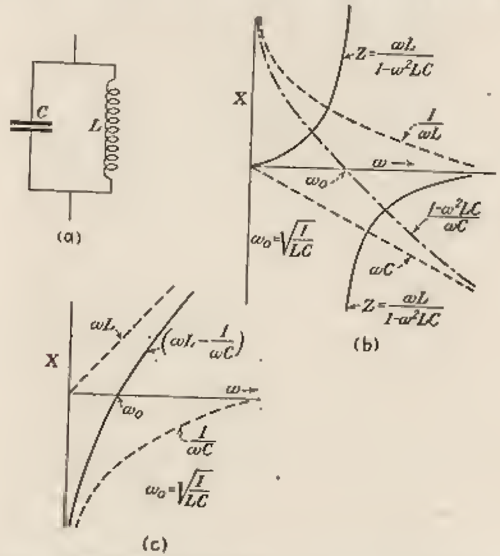


FIG. 4.—Reactance and impedance of parallel circuit.

and is designated by the letter X. The impedance may be written, therefore,

$$z = R + jX$$

Occasionally the absolute value of the circuit impedance is required. It is then written in the following form

$$z = Ze^{j\phi}$$

where

$$Z = \sqrt{R^2 + X^2}$$

$$\phi = \arctan \frac{X}{R}$$

In this expression Z represents the absolute value of the impedance, z the complex value, and ϕ the phase angle.

The impedance of a single circuit will be given to illustrate the method of obtaining this quantity for any circuit. For a parallel combination of circuit elements, such as illustrated in Fig. 4*a*, it would be obtained as follows:

$$z = \frac{1}{\dfrac{1}{1/j\omega C} + \dfrac{1}{j\omega L}} = \frac{j\omega L}{1 - \omega^2 LC}$$

This equation shows that when $\omega^2 = 1/LC$ the impedance is infinite. It may be represented graphically as a function of ω as shown in Fig. 4*b*. The figure and the equation illustrate the case of parallel resonance. The case of series resonance is illustrated in Fig. 4*c*, and the equation is $z = j\left(\omega L - \dfrac{1}{\omega C}\right)$, which holds for a circuit having only an inductance L and capacitance C in series with the e.m.f. In the series case, the impedance is zero at resonance; that is, when $\omega^2 = 1/LC$ and in the parallel case the impedance is infinite at resonance.

23. Circuit Parameters. Every electric circuit, no matter how complicated, is made up of a particular combination of inductances, capacitances, and resistances. These parameters and the manner in which they are combined with one another completely govern the performance of a circuit and determine the value of the current at any point of the circuit at any time for any given value of the impressed e.m.f. or combination of e.m.fs.

Inductances, capacitances, and resistances may be lumped or distributed in nature. They are regarded as of the former type if their values are more or less concentrated at one or a finite number of points in a circuit. For example, the inductance of a circuit would be considered as lumped if a definite number of places in the circuit is found where inductance exists, and at all other points a comparative non-existence of inductance. On the other hand the inductance of a uniform telephone line is considered as distributed since it exists along the entire line and may, at no point in the line, be neglected.

24. Circuit Equations. Every circuit may be completely expressed by a system of simultaneous equations. Having expressed a particular circuit in this manner, a solution may be obtained frequently without difficulty. Since the equations are of primary importance, methods of obtaining them will be given.

There are two distinct cases. When a sinusoidal voltage or combination of sinusoidal voltages is impressed on a circuit, a.c. flows in every branch of the circuit as a consequence of the impressed e.m.f. This current may be divided into two parts. One part is known as the *transient* current, and the other as the current of the *steady state*. The transient current disappears very shortly after the voltage has been impressed. The steady state continues as long as the e.m.f. continues in its initial state of voltage, frequency, and wave form. Often only the steady state is of interest. Examples of this are to be found in studies of r-f transformer performance and in studies of electric filters of the low-pass, high-pass, or band-pass types and in the studies of the various characteristics of different antenna-coupling methods. At other

times the transient condition may be of primary interest; as, for example, in the study of the fidelity of reproduction with regard to wave form of an electromagnetic or electrodynamic loud-speaker motor.

If interest centers only in the steady state the following method is to be used: Apply Kirchhoff's second law which states that the sum of all the e.m.fs. around any circuit is zero, writing one equation for each branch of the circuit, and using as the potential falls the values $j\omega LI$ for each inductance, $I/j\omega C$ for each capacitance, and IR for each resistance. If inductances, capacitances, and resistances occur that are common to two or more branches, they will be used once for each of the common branches paying due regard to the sign of the term.

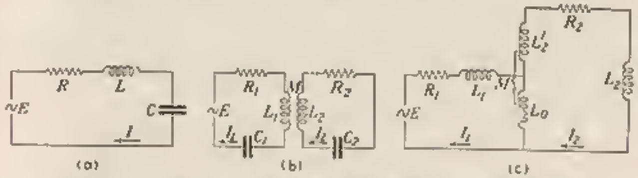


FIG. 5.—Circuits illustrating use of Kirchhoff's laws.

This method may be illustrated by the examples of Fig. 5 and the following equations:

For circuit *a*:

$$\begin{aligned} E &= IR + j\omega LI + \frac{I}{j\omega C} = I\left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right] \\ &= I(R + jX) \\ I &= \frac{E}{R + jX} \end{aligned}$$

For circuit *b*:

$$E = I_1R_1 + j\omega L_1I_1 + \frac{I_1}{j\omega C_1} - j\omega MI_2 = I_1z_1 - j\omega MI_2$$

$$0 = I_2R_2 + j\omega L_2I_2 + \frac{I_2}{j\omega L_2} - j\omega MI_1 = I_2z_2 - j\omega MI_1$$

where z_1 is the total complex impedance of circuit 1, and z_2 is the total complex impedance of circuit 2.

For circuit *c*:

$$\begin{aligned} E &= I_1R_1 + j\omega L_1I_1 + j\omega L_0I_1 - j\omega MI_2 - j\omega L_0I_2 \\ &= I_1z_1 - j\omega I_2(M + L_0) \\ 0 &= I_2R_2 + j\omega L'_2I_2 + j\omega L_0I_2 + j\omega L_2I_2 - j\omega MI_1 - j\omega L_0I_1 \\ &= I_2z_2 - j\omega I_1(M + L_0) \end{aligned}$$

In these equations I is the maximum value of the sinusoidal current, and E is the maximum value of the sinusoidal e.m.f. These equations may be solved for any of the currents by the method of simultaneous equations.

In the transient values of the various currents, Kirchhoff's second law may be used as before, but instead of using the values of potential fall as given in the preceding equations, use the instantaneous values. The equation for circuit *a* of Fig. 5 is then written

$$e = iR + L\frac{di}{dt} + \frac{1}{C}\int i\,dt$$

or

$$\frac{de}{dt} = L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{i}{C}$$

where e and i are the instantaneous values of the impressed e.m.f. and current respectively. For circuit b,

$$e = i_1R_1 + L_1\frac{di_1}{dt} + \frac{1}{C_1}\int i_1dt - M\frac{di_2}{dt}$$

$$0 = i_2R_2 + L_2\frac{di_2}{dt} + \frac{1}{C_2}\int i_2dt - M\frac{di_1}{dt}$$

To obtain the transient solution, e and de/dt are replaced by zero and the equation solved by the methods used for linear, homogeneous equations of the first degree.

25. General Characteristics of A-c Circuits. The general equations applied to a number of the more important radio circuits yield the following results.

Current Flow in an Inductive Circuit:

$$i = \frac{E}{R}\left(1 - \epsilon^{-\frac{Rt}{L}}\right)$$

where E is the constant impressed e.m.f.

Time Constant of an Inductive Circuit: The time required for a current to rise to $\left(1 - \frac{1}{\epsilon}\right)$ or to about 63 per cent of its final value. This time is equal to L/R.

Current Flow in a Capacitive Circuit:

$$i = \frac{E}{R}\epsilon^{-\frac{t}{RC}}$$

where E is the constant impressed e.m.f.

Time Constant of a Capacitive Circuit. The time required for the current to fall from its initial value to $1/\epsilon$ or about 0.37 of this value. This time is equal to RC.

Current Flow in an Inductive-capacitive Circuit:

$$i = \frac{E}{\omega L}\epsilon^{-\frac{Rt}{2L}} \sin \omega t, \text{ if } R^2 < \frac{4L}{C}$$

$$i = \frac{E}{\omega L}\epsilon^{-\frac{Rt}{2L}} \qquad , \text{ if } R^2 = \frac{4L}{C}$$

where ω is 2π times the natural frequency of the circuit which is given by the equation

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}}$$

Logarithmic Decrement. Ratio of successive maxima of the current in an oscillatory discharge is equal to

$$\epsilon^{\frac{RT}{2L}} = \epsilon^{\frac{R}{2Lf}}$$

where $R/2Lf$ is called the log. dec. of the circuit, T is the natural period, and f the natural frequency of the circuit.

Currents in Two Circuits Coupled by a Mutual Impedance, M, when a Sinusoidal E.M.F., E, Exists in Circuit 1:

$$I_1 = \frac{E}{z_1 + \frac{\omega^2M^2}{z_2}}$$

$$I_2 = \frac{j\omega MI_1}{z_2} = \frac{j\omega ME}{z_1z_2 + \omega^2M^2}$$

where z_1 and z_2 are the complex impedances of circuits 1 and 2 respectively.

Effective Reactance of One Circuit Coupled to a Second Circuit:

$$X' = X_1 - \frac{\omega^2M^2}{Z_2^2}X_2$$

where X_1 and X_2 are the actual reactances of circuits 1 and 2 respectively and Z_2 is the absolute value of the complex impedance of circuit 2.

Effective Resistance of One Circuit Coupled to a Second Circuit:

$$R' = R_1 + \frac{\omega^2M^2}{Z_2^2}R_2$$

where R_1 and R_2 are the actual resistances of circuits 1 and 2 respectively.

Effective Total Impedance of One Circuit Coupled to a Second Circuit:

$$z' = z_1 + \frac{\omega^2M^2}{z_2} = R_1 + jX_1 + \frac{\omega^2M^2}{R_2 + jX_2}$$

$$= R_1 + \frac{\omega^2M^2}{Z_2^2}R_2 + j\left\{X_1 - \frac{\omega^2M^2}{Z_2^2}X_2\right\}$$

Partial Resonance Relation Obtained When Only the Reactance of Circuit 1 Is Variable:[1]

$$X_1 = \frac{\omega^2M^2}{Z_2^2}X_2$$

Partial Resonance Relation Obtained when only the Reactance of Circuit 2 Is Variable:[1]

$$X_2 = \frac{\omega^2M^2}{Z_1^2}X_1$$

Total Optimum Resonance Relation when the Reactance of Both Circuits 1 and 2 Are Variable:[1]

Case I: If $\omega^2M^2 < R_1R_2$

Resonance relation $X_1 = 0$ and $X_2 = 0$

Case II: If $\omega^2M^2 > R_1R_2$

$$\text{Resonance relation } \frac{R_2}{R_1} = \frac{\omega^2M^2}{Z_1^2} = \frac{X_2}{X_1}$$

Case III: If $\omega^2M^2 = R_1R_2$

Resonance relation $X_1 = 0,\ X_2 = 0$

$$\frac{R_2}{R_1} = \frac{\omega^2M^2}{Z_1^2}$$

Total Secondary Current at Total Optimum Resonance Relation, the E.M.F., E, Being Impressed in Circuit 1.

Case I: If $\omega^2M^2 < R_1R_2$

$$I_2 = \frac{\omega ME}{R_1R_2 + \omega^2M^2}$$

[1] PIERCE, G. W., "Electric Oscillations and Electric Waves," Chap. XI.

Cases II and III: If $\omega^2M^2 \geq R_1R_2$

$$I_2 = \frac{E}{2\sqrt{R_1R_2}}$$

I_2 for cases *II* and *III* is seen to be greater than for case *I* and is independent of ωM.

MAGNETIC CIRCUITS

26. The Fundamental Quantities of Magnetic Circuits. The first fundamental quantity is the *magnetic flux* or *induction*. The unit of flux is known as the *maxwell* and is defined by the statement that from a unit magnetic pole, 4π maxwells, or lines of force, radiate.

The second fundamental quantity is the *reluctance*. It is analogous to the resistance of electric circuits, as the flux is analogous to the current. The unit of reluctance is the *oersted* and is defined as the reluctance offered by 1 cm cube of air.

The third fundamental quantity is the *magnetomotive force* (m.m.f.). It is analogous to the e.m.f. of electrical circuits. The unit of m.m.f. is the *gilbert* and is defined as the m.m.f. required to force a flux of 1 maxwell through a reluctance of 1 oersted. Thus the fundamental equation in which these three quantities are related to one another is:

$$M = \phi R$$

Other important quantities of magnetic currents may be defined as follows: the *magnetic field strength* is represented by the quantity H and is equal to the number of maxwells per unit of area when the medium through which the flux is passing is air. This unit is known as the *gauss* if the unit of area is the square centimeter.

In any medium other than air the lines of force are known as *lines of induction* and the symbol B is used instead of H to represent them. In air the induction B and the field strength H are equal to one another, but in other mediums this is not true.

The *permeability* μ is the ratio between the magnetic induction B and the field strength H. In air this ratio is unity. In *paramagnetic* materials the permeability is greater than unity, in *ferromagnetic* materials it may have a value of several thousand, and in *diamagnetic* materials it has a value of less than unity.

The intensity of magnetization I is the *magnetic moment* per unit volume or the *pole strength* per unit area. The unit of magnetic pole strength is a magnetic pole of such a value that when placed 1 cm from a like pole, a force of repulsion of 1 dyne will exist between them. The magnetic pole strength per unit area of any pole is measured in terms of this unit. The magnetic moment of a magnet is the product of the pole strength and the distance between the poles.

The *susceptibility* K of a material is equal to the ratio of the magnetization I produced in the material to the field strength H producing it. All of these quantities are connected by the following equations

$$B = \mu H$$
$$I = KH$$
$$B = 4\pi I + H$$
$$\mu = 4\pi K + 1$$

Magnetization curves are of great importance in the design of magnetic structures and should be immediately available for all materials with which one intends to work. These curves may give either the values of

B as a function of H for the material, or the values of I as a function of H. A typical B-H curve is shown in Fig. 6. The ratio of the coordinates of a B-H curve gives the value of μ for the material at the particular value of H chosen. The ratio of the coordinates in an I-H curve similarly gives the value of the susceptibility K.

Magnetic *saturation* is a phenomenon occurring at large values of H when the induction B increases at a much lower rate with increase of H than is the case for small values of H.

The *retentivity* of a substance is the value of B in the material when the field H is reduced to zero after having first been raised to above its saturation value. It is given by the point A of the B-H curve of Fig. 6.

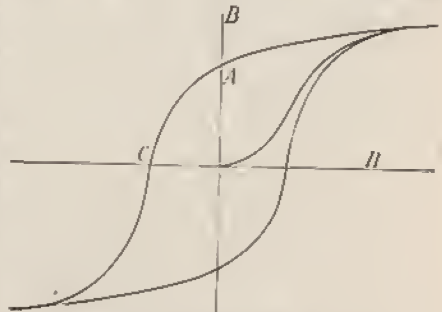


Fig. 6.—Typical B-H curve.

The *coercivity* of a material is the minimum negative value of H required to just reduce the induction to zero after the field strength H has first been raised to a positive value sufficiently large to saturate the material. It is given by the point C of the B-H curve of Fig. 6.

27. Magnetic Properties of Iron and Steel.

Material	Coercivity	Retentivity	Maximum permeability	$4\pi I$ at saturation
Electrolytic iron	2.83	11,400	1,850	21,620
Annealed	0.36	10,800	14,400	21,630
Annealed electrical iron in sheets	1.30	9,400	3,270	20,500
Cast steel	1.51	10,600	3,550	21,420
Annealed	0.37	11,000	14,800	21,420
Steel hardened	52.4	7,500	110	18,000
Cast iron	11.4	5,100	240	16,400
Annealed	4.6	5,350	600	16,800
Tungsten magnet steel	64.0	9,600	105	13,600
Chrome magnet steel	64.0	9,000	94	12,000
Cobalt steel (15 per cent)	192.0	8,000		

28. Electromagnetic Structures. In this type of structure the magnetic material is usually very soft; its coercivity is very low; and as a consequence the m.m.f. must be supplied by a continuous electric current. The m.m.f., M, due to an electric current, is given by the equation $M = 0.4\pi NI$, where I is the current in amperes, and N is the number of turns on the electromagnet.

By our most fundamental relation for magnetic circuits

$$M = \phi R$$
$$0.4\pi NI = R\phi$$
$$NI = \frac{R\phi}{0.4\pi}$$

The design of a magnetic structure is usually begun by a consideration of the flux requirements in a particular air gap. The size and shape of the air gap are generally given, and the flux density desired in the air gap is known. From these data one can compute R and ϕ. For the quantity ϕ, $\phi = BA$, where A is the area of the air gap and B is the flux density desired. This equation assumes no leakage flux, and since this is a condition never realized in practice and from which there may be a far from negligible departure, one must add to the value of ϕ given by this equation a correction the value of which is dictated by experience. For the quantity R, $R = L/A$, where L is the length of the air gap and A is the area. This equation neglects the reluctance of the magnet itself and of all other iron parts of the magnetic circuit. Since all reluctances but that residing in the air gap are very small in comparison, this procedure is usually justified, although there are cases in which additional reluctance must be taken into account. In such cases the reluctance of the other parts of the circuit is computed in the same manner as that of the air gap, except that an estimate of the permeability of the part in the circuit in question must be made and its equivalent air-gap reluctance computed by dividing by this permeability. Finally,

$$NI = \frac{R\phi}{0.4\pi} = \frac{LBA}{0.4\pi A} = \frac{LB}{0.4\pi}$$

This equation then completely determines the value of the ampere-turns NI from the original data. This is the important quantity in the design of the electromagnet. The separate values of N and I are undetermined by this equation, other considerations such as the nature of the current supply, the size of the coil, the heat dissipation that can be permitted and the cost being of paramount importance.

29. Core Materials for Receiver Construction (The Editor). Since such materials operate under widely different conditions each material must be properly selected for its particular task. For example, materials used in economical audio transformers are too expensive to be used in power transformers.

Power Transformers. Material for cores of transformers supplying energy for plate and filament circuits is selected as for any power transformer upon a watt-loss basis. This information is reliably supplied by manufacturers of such material, and measurements of this factor are not generally made by the user of the material. Loss tests are made on complete transformers to determine the suitability of the material under consideration.

The mechanical properties of the sheets submitted by various suppliers are important. By causing injury to or premature loss of a die, poor mechanical properties may tie up a production schedule. Wavy irregular sheets necessitate scrapping wide strips from both sides of each sheet and introduce an unexpected cost.

Permeability of the core material is of importance where limited space or weight requirements make necessary the use of flux densities of 14,000 gausses or higher. Here a high permeability is indicated to avoid high exciting copper losses and poor voltage regulation.

Audio Transformers: Filter Reactors. Here the permeability is of importance. The factor to be used is the *working permeability* or *apparent a-c permeability* instead of the theoretical value obtained from B-H or μ-B curves. This useful value must be obtained from the working inductance of some definite design of choke or transformer. Such values will take care of the fact that in audio transformers and chokes the core material is polarized by a relatively high unidirectional magnetizing force (plate current or load current through the filter).

The apparent a-c permeability may be determined from the following expression taken from the Allegheny Steel Company's book, "Magnetic Core Materials for Radio."

$$\mu a = \frac{La \times l \times 10^8}{1.256 \times AK_1N^2} = \frac{l \times 10^8}{1.256AN^2} \times \frac{La}{K_1}$$

where La = apparent inductance in henrys
A = cross-sectional area of core in square centimeters
K_1 = core stacking factor
N = number of turns in the winding
l = length of magnetic path in centimeters.

The quantity ($l \times 10^8 \div AN^2$) is a constant determined by the physical dimensions of the core and the number of turns in the coil. The quantity (La/K_1) indicates the way in which the stacking factor as affected by the punching characteristics enters into the determination of the permeability. Material which acts badly mechanically results in burrs in punching and gives a reduced number of pieces in a given design. This gives lower inductance but does not affect the permeability determination.

The value of the stacking factor for any design is given by dividing the product of the core volume (cubic centimeters) and the specific gravity of the core material into the actual measured weight of the core material in grams. Thus,

$$K_1 = \frac{W}{Vg}$$

where W = weight of core in grams; V = volume in cubic centimeters; g = specific gravity of the core material.

The value of g may vary as follows:

Silicon steel with silicon content 2 per cent or less........	7.7
Silicon steel with silicon content more than 2 per cent....	7.5
Allegheny electric metal.............................	8.3

Manufacturers of transformer iron supply curves from which a designer may learn the incremental or apparent a-c permeability of the iron he proposes to use. From these curves the inductance of a core winding may be determined by using the above formula.

To determine the inductance of a winding on a core with an air gap use the following schedule:

Total m.m.f. $= 1.256 \times I \times N = H_1l_1 + H_2l_2 = H_1l_1 + B_0l_2$

where I = current (d-c)
N = number of turns in the winding
l_1 and l_2 = the iron and air paths
H_1 and H_2 = magnetic potential gradients along those paths
$H_2 = B_0$ in air

This equation is that of a straight line intersecting the vertical axis of a B-H curve at a point corresponding to $H_1 = 0$ and $B_0 = \text{m.m.f.}/l_2$ and intersecting the horizontal axis at a point corresponding to $B_0 = 0$ and $H_1 = \text{m.m.f.}/l_1$. Thus the d-c flux density in the core and the magnetic potential gradient in the iron part of the circuit and the a-c permeability ($\mu_{a\text{-}c} = B/H$) may be determined. The a-c reluctivity is the reciprocal of the a-c permeability. The *apparent reluctivity* is equal (in cases where the air gap is 1 per cent or less of the iron path) to the a-c reluctivity plus the ratio of the air gap to the length of the mean iron path. The reciprocal of this value of apparent reluctivity is the apparent permeability which, substituted in the formula above, determines the inductance.

RADIATION

30. Nature of Radiation. Electromagnetic energy may arise from continuously varying electronic currents in a conductor, displacement currents, or oscillating dipoles. In order that this energy may be appreciable it is necessary that the system of conductors be of such a form that the electromagnetic field will not be confined in any way and that the frequency of oscillation of the current or charges be high. The various forms of antennas and the employment of radio frequencies satisfy these requirements.

The nature of radiation may be understood only after a complete examination of Maxwell's equations and the various transformations of the wave equation. Any attempt to give a simple yet accurate picture of the phenomenon of radiation must be fruitless, though such pictures may aid in an understanding of the subject. Such descriptions may be found in any text on radio. An exact analysis of Maxwell's equations shows that whenever an electric wave moves through space an associated magnetic wave having its vectors at right angles to that of the electric wave must accompany it. Both vectors, furthermore, are at right angles to the direction of propagation. This analysis also shows that an electromagnetic field due to an oscillating dipole or to an oscillating current in a conductor has two components. One of these varies inversely as the first power of the distance from the source and is, furthermore, directly proportional to the frequency, and the other varies inversely as the second power of the distance. The former is known as the *radiation field* and the latter as the *induction field*. Though indistinguishable physically, the induction and radiation fields have a separate mathematical existence accounting completely for the phenomenon of energy radiation. The energy of the induction field returns to the conductor with the completion of each cycle. Its existence is confined, as one might expect, to the neighborhood of the conductor, whereas the radiation field may be thought of as a detached field traveling outward into space with the velocity of light and varying much more slowly in intensity with distance from the conductor than the other.

31. Vertical Antenna. The most simple form of antenna is the vertical wire. The electromagnetic radiation field depends on the strength of the current in the wire, and as a consequence its intensity is increased if the current throughout the vertical wire is uniform. It is for this reason that a counterpoise is usually attached to the lower end of the antenna and a horizontal aerial to the upper end. The capacity of the counterpoise and aerial may be made so high that the current throughout the vertical portion of the wire is practically uniform.

Under these conditions the magnetic field at any distant point is given by the equation

$$H = -\frac{\omega h I_0}{10cl} \cos \omega\left(t - \frac{l}{c}\right) \text{ gauss}$$

where $\omega = 2\pi f$
f = frequency of oscillation
I_0 = maximum value of the current in the antenna
c = velocity of light in centimeters per second in vacuum
l = distance from the source in centimeters
h = height of antenna or length of vertical wire in centimeters

and

$$E = -\frac{300\omega h I_0}{10cl} \cos \omega\left(t - \frac{l}{c}\right) \text{ volts}$$

These equations[1] are derived by considering the antenna as an oscillating Hertzian doublet of separation h. The effective values of the magnetic and electric fields are

$$H_e = -\frac{\omega h I_e}{10cl} = -\frac{2\pi h I_e}{10\lambda l}$$

$$E_e = -\frac{300\omega h I_e}{10cl} = -\frac{600\pi h I_e}{10\lambda l}$$

where I_e is the effective value of the antenna current, and λ is the wave length of the electromagnetic wave.

32. Loop Antenna. The field due to a loop antenna is given by the equations

$$H_e = \frac{4\pi h I_e}{10\lambda l} \sin \frac{\pi s}{\lambda}$$

$$E_e = \frac{1{,}200\pi h I_e}{10\lambda l} \sin \frac{\pi s}{\lambda}$$

where s is the distance of separation of the vertical portions of the loop in centimeters.

33. Coil Antenna. For a coil of N turns having negligible capacity between turns at the frequency considered so that the current in all turns is substantially the same, the field is given by the equations

$$H_e = \frac{4\pi N h I_e}{10\lambda l} \sin \frac{\pi s}{\lambda}$$

$$E_e = \frac{1{,}200\pi N h I_e}{10\lambda l} \sin \frac{\pi s}{\lambda}$$

34. The fundamental and harmonic frequencies of oscillation in an antenna may be calculated in many cases. If the inductance and capacity of the vertical wire of the antenna are neglected, the low frequency capacity and inductance are given by the equations[2]

$$C = lC_l$$

$$L = \frac{l}{3}L_l$$

where C_l and L_l are the capacity and inductance per unit length of conductor, and l is the length of conductor. These equations may be calculated by means of accurate formulas which are available.[3]

Then the low-frequency reactance of the antenna is

$$X_l = \frac{\omega l L_l}{3} - \frac{1}{\omega l C_l}$$

[1] BERG, "Electrical Engineering," Advanced Course, pp. 278 *ff.*; MORECROFT, "Principles of Radio Communication," p. 706.
[2] *Bur. Standards Circ.* 74, pp. 72 *ff.*
[3] *Bur. Standards Circ.* 74, pp. 237–243.

The high-frequency reactance of the antenna is given by the equation

$$X_a = -\sqrt{\frac{L_i}{C_i}} \cot \omega l\sqrt{L_iC_i}$$

The reactance of the antenna becomes zero when

$$\omega l\sqrt{C_iL_i} = n\frac{\pi}{2}(n = 1, 3, 5 \cdots)$$

that is, when

$$f = \frac{\omega}{2\pi} = \frac{n}{4l\sqrt{C_iL_i}}$$

The reactance becomes infinite when

$$\omega l\sqrt{C_iL_i} = m\frac{\pi}{2}(m = 0, 2, 4 \cdots)$$

that is, when

$$f = \frac{\omega}{2\pi} = \frac{m}{4l\sqrt{C_iL_i}}$$

If the inductance of the vertical wire is to be considered, or if a series inductance is used with the antenna

$$X = \omega L_s - \sqrt{\frac{L_i}{C_i}} \cot \omega l\sqrt{C_iL_i}$$

where L_s is the total inductance of the vertical wire and any coils in series with the antenna.

The harmonic frequencies of the antenna at which the reactance is zero do not differ by multiples of π as before. The natural frequency of oscillation is given, however, quite generally by the equation

$$\omega L_s - \sqrt{\frac{L_i}{C_i}} \cot \omega l\sqrt{C_iL_i} = 0$$

$$\frac{\cot \omega l\sqrt{C_iL_i}}{\omega\sqrt{C_iL_i}} = \frac{L_s}{L_i}$$

35. Antenna Resistance. The resistance of an antenna may be divided into three parts in which the power dissipation is of the following kinds:

1. Radiation.
2. Joule heat.
3. Dielectric absorption.

The power radiated depends on the form of the antenna. It is proportional to the square of the frequency of oscillation and to the square of the current flowing in the antenna. Due to the latter consideration one may write $P = AI^2$, where A is a constant factor depending on the form of the antenna and the frequency. It may be called the radiation resistance. For a given antenna the radiation resistance varies inversely as the square of the wave length. The ohmic resistance to which the joule heat is due is approximately constant, the skin effect and other factors being comparatively small. The resistance due to dielectric absorption is directly proportional to the wave length. When these three components of resistance are added to obtain the total resistance, one finds that for every antenna there is a wave length for which the total resistance is a minimum.

36. Energy in the Field. The energy of an electromagnetic field at any point is given by the equation[1]

$$U = \frac{1}{8\pi}(\epsilon E^2 + \mu H^2)$$

where E is in electrostatic units instead of volts as in the previous equations, ϵ is the dielectric constant, and μ the permeability of the medium. In free space

$$U = \frac{1}{8\pi}(E^2 + H^2)$$

But, in general,

$$H = \sqrt{\frac{\epsilon}{\mu}}E$$

$$U = \frac{\epsilon}{4\pi}E^2 = \frac{\mu}{4\pi}H^2$$

$$= \frac{E^2}{4\pi} = \frac{H^2}{4\pi} \text{ in free space.}$$

The energy flux through 1 sq cm of surface, perpendicular to the direction of propagation, is given by the equation

$$S = vU = \frac{c}{\sqrt{\epsilon\mu}}U = \frac{c}{4\pi}\sqrt{\frac{\epsilon}{\mu}}E_e^2 = \frac{c}{4\pi}\sqrt{\frac{\mu}{\epsilon}}H_e^2$$

$$\left.\begin{aligned} &= \frac{c}{4\pi}E_e^2 = \frac{c}{4\pi}H_e^2 \\ &= \frac{c}{8\pi}E_m^2 = \frac{c}{8\pi}H_m^2 \end{aligned}\right\} \text{in free space.}$$

where E_e and H_e represent effective values, and E_m and H_m the maximum values of the electric and magnetic fields respectively. Therefore, for the effective values of the electric and magnetic fields due to a vertical wire antenna,

$$E_e = -\frac{2\pi hI_e}{10\lambda l} \text{ e.s.u.}$$

$$H_e = -\frac{2\pi hI_e}{10\lambda l}$$

$$S = \frac{c}{4\pi}\left(\frac{2\pi hI_e}{10\lambda l}\right)^2 = \frac{c\pi h^2I_e^2}{10^2\lambda^2l^2}$$

Then the total radiation from a vertical antenna, assuming that H has its maximum value in the equatorial plane of the antenna and that its variation in a vertical plane at a distance l from the antenna follows a sine law, is given by the expression

$$2\pi l^2\left(\frac{c\pi h^2I_e^2}{10^2\lambda^2l^2}\right) \text{ ergs per second}$$

or

$$\frac{60\pi^2h^2I_e^2}{\lambda^2} \text{ watts}$$

[1] JEANS, J. H., "Mathematical Theory of Electricity and Magnetism," p. 518

SECTION 3

RESISTANCE

By Jesse Marsten, B.S.[1]

1. General Concepts. In any electrical conductor or system in which there is a flow of current there is a certain amount of energy continually being lost or converted into forms not readily available for use. As far as is known at present this dissipation of energy may take one of two forms: there may be an evolution of heat, and there may be radiation of energy into space. Such energy dissipation is attributed to a property of electric conductors or systems termed *resistance.*

When dealing with continuous currents, the resistance of a conductor or network, R, is adequately defined by *Ohm's law,*

$$E = iR \tag{1}$$

where E is the voltage drop across the conductor or network and i is the current through it. This assumes no back e.m.f. due to polarization or other causes. In this case the dissipation of energy takes place entirely in the form of heat generation, and the rate at which electrical energy is thus converted into heat is given by *Joule's law,*

$$P = i^2R \tag{2}$$

where P is the power or rate at which electrical energy is being dissipated in the form of heat, i is the continuous current in the circuit, and R the resistance of the circuit.

Ohm's law is insufficient to define resistance in a-c circuits. It is found experimentally that the rate at which heat is evolved in a circuit exceeds that which would be necessitated by the resistance of the circuit as determined by Ohm's law. This is due to the fact that the electromagnetic and electrostatic fields around the circuit vary with time and introduce effects which increase the losses in the circuit. Among these effects may be enumerated the following major ones:

1. *Eddy-current* losses in conductors and other masses of metals in and near the circuit.
2. *Hysteresis* losses in magnetic materials.
3. *Dielectric* losses in the insulating mediums.
4. *Absorption* of energy by neighboring conductors or circuits by induction.
5. *Radiation* of electromagnetic energy into space.
6. *Skin Effect.* Increase of conductor resistance due to non-uniform current density.

[1] Member, Institute of Radio Engineers; associate member, American Institute of Electrical Engineers, chief engineer, International Resistance Company.

All these effects result in an increase in energy loss in the circuit over and above that given by Ohm's law. It therefore becomes necessary to introduce the concept of *a-c resistance* or *effective resistance*, which is defined by the more general joulean relationship,

$$P = i^2R \text{ effective} \tag{3}$$

where P is the power loss in the circuit due to all causes and i is the effective current in the circuit. Ohm's law for continuous currents follows directly from this more general definition.

2. Units of Resistance. The practical unit of resistance is the *ohm* and is defined by Ohm's law when the voltage and current are unity in the practical system. It has, however, been arbitrarily defined as the resistance at 0°C. of a column of mercury having a uniform cross section, a height of 106.3 cm, and weighing 14.4521 g. Owing to the increasing use of resistors having resistances of the order of millions of ohms, the *megohm* unit is also employed. The megohm is equal to 10^6 ohms.

3. Specific Resistance. It is found experimentally that the resistance of an electric conductor is directly proportional to its length and inversely to its cross section:

$$R = \rho\frac{l}{A} \tag{4}$$

The proportionality factor ρ is called the *specific resistance* of the conductor and is a function of the material of the conductor.

From this definition of specific resistance it is apparent that any number of units may be derived for specific resistance, depending upon the units chosen for l and A. The unit generally employed in practical engineering is the *ohms per circular mil foot*, and is the resistance of a 1 ft. length of the conductor having a section of 1 cir. mil (diameter 1 mil for a circular conductor).

4. Volume Resistivity. If, in the above definition, l and A are both unity, in the same system of units, then ρ is the resistance of a unit cube of the material and may be defined as the *volume resistivity* of the material. It should be noted that volume resistivity is not the resistance of any unit volume of the material but is specifically the resistance of unit volume measured across faces whose areas are each unity.

With a knowledge of the dimensions of a conductor and its specific resistance the resistance of the conductor to d.c. may be computed from Eq. (4). Consistent units must be employed. The resistance thus computed will be correct at the temperature for which the specific resistance applies. To obtain the resistance of the conductor at any other temperature a correction will have to be applied.

5. Temperature Coefficient. The resistance of a conductor is a function not only of the material and dimensions of the conductor but also of its temperature. Within the temperature limits generally encountered in practice the change in resistance due to temperature variation is directly proportional to the change in temperature:

$$R_{t_2} = R_{t_1}[1 + \alpha(t_2 - t_1)] \tag{5}$$

R_{t_1} and R_{t_2} are the conductor resistances at temperature t_1 and t_2 respectively.

The proportionality factor α is defined as the *temperature coefficient of resistance* of the material and is the change in resistance of any material per ohm per degree rise in temperature.

All conductors do not react alike to changes in temperature. Metals, for example, have a positive temperature coefficient. Some alloys, such as manganin and constantan, have practically zero temperature coefficient and are therefore used primarily for resistance standards.

A knowledge of the temperature coefficient of conductor materials enables one at times to make more accurate determinations of temperature change than is possible by thermometer measurements, especially in cases where parts to be measured are not readily accessible. Resistance determinations of the conductor are made at both temperatures and the temperature change computed from Eq. (5).

6. Properties of Materials as Conductors.

Material	Specific resistance at 0°C., ohms per cir. mil ft.	Temperature coefficient per °C. between 20° to 100°C., ohms per °C.
Silver	9.75	0.004
Copper	10.55	0.004
Aluminum	17.3	0.0039
Nickel (pure)	58.0	0.0048
Iron (pure)	61.1	0.0062
Phosphor bronze	70.0	0.004
Lead	114.7	0.0041
Nickel silver, 18 per cent (German silver)	190	0.00019
Manganin (copper, 82 per cent; manganese, 14 per cent; nickel, 4 per cent)	290	0.00002
Constantan (Advance, Cupron, Ideal, Ia-Ia) (copper, 55 per cent; nickel, 45 per cent)	294	0.00002
Nichrome (nickel, 60 per cent; chromium, 15 per cent; iron, balance)	650 to 675	0.00017

7. Resistors in Series and Parallel. Simple and complex networks of resistors may be represented by an equivalent resistor which may be expressed in terms of the individual resistances making up the network.

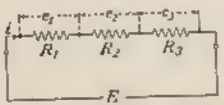


Fig. 1.—Simple series circuit.

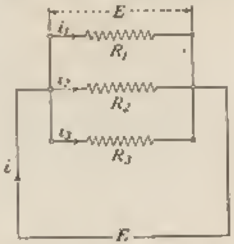


Fig. 2.—Parallel circuit.

The equivalent resistance of a number of resistors connected in series is equal to the sum of the individual resistances. Referring to Fig. 1:

$$E = iR_{\text{equiv.}} = e_1 + e_2 + \cdots + e_n = R_1 i + R_2 i + \cdots + R_n i = i(R_1 + R_2 + \cdots + R_n)$$

$$\frac{E}{i} = R_{\text{equiv.}} = (R_1 + R_2 + \cdots + R_n)$$

$$R_{\text{equiv.}} = \sum_1^n R$$

The reciprocal of the equivalent resistance of a number of resistors connected in parallel is equal to the sum of the reciprocals of the individual resistances. Referring to Fig. 2:

$$i = i_1 + i_2 + \cdots + i_n = \frac{E}{R_1} + \frac{E}{R_2} + \cdots + \frac{E}{R_n}$$

$$\frac{i}{E} = \frac{1}{R_{\text{equiv.}}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}$$

$$\frac{1}{R_{\text{equiv.}}} = \sum_1^n \frac{1}{R}$$

RESISTANCE AS FUNCTION OF FREQUENCY

8. Skin Effect. It may be shown that the resistance of a conductor is a minimum when the current density is uniformly distributed over the cross section of the conductor. This condition obtains for d.c. The resistance increases for non-uniform distribution of current density over the cross section of the conductor. This latter condition obtains in conductors carrying a.c. This is a result of the distribution of magnetic-flux lines, outside and inside the conductor. If the conductor is assumed to be made up of a number of conducting elements in parallel, then the interior elements, being surrounded by more flux lines than the exterior, will have greater reactance and, therefore, the current in the interior elements will be less than that in the exterior elements. As a result the current crowds toward the surface of the conductor, giving a non-uniform current density. This imperfect penetration of current in a conductor, resulting in an increase in resistance, is termed *skin effect*.

Skin effect in a conductor is a function of the following factors:

$$t\sqrt{\frac{\mu f}{\rho}} \tag{6}$$

where t = thickness of the conductor
f = frequency of current
μ = permeability of the conductor
ρ = specific resistance of the conductor in microhm-centimeters.

It is possible to compute accurately the h-f resistance of simple round cylindrical conductors from involved functions of the above factor. To facilitate these computations tables have been prepared from which the ratio of h-f resistance R_f to d-c resistance R_0 may be quickly determined. From this factor and the easily measured d-c resistance the h-f resistance may be computed.

The table below gives the values of R_f/R_0 for different values of the factor

$$x = \pi d \sqrt{\frac{2\mu f}{\rho}} \sqrt{\frac{1}{1000}}$$

where d is the diameter of the wire in centimeters, ρ is the volume resistivity in microhm-centimeters (1.724 at 10°C. for copper), x may be computed for any particular case, and R_0 may be measured at d.c. or computed.

9. Ratio of H-f Resistance to the D-c Resistance for Different Values of $x = \pi d\sqrt{2\mu f/\rho} \times \sqrt{1/1000}$.

x	R_f/R_0	x	R_f/R_0	x	R_f/R_0
0	1.0000	5.2	2.114	14.0	5.209
0.5	1.0003	5.4	2.184	14.5	5.386
0.6	1.0007	5.6	2.254	15.0	5.582
0.7	1.0012	5.8	2.324	16.0	5.915
0.8	1.0021	6.0	2.394	17.0	6.268
0.9	1.0034	6.2	2.463	18.0	6.621
1.0	1.005	6.4	2.533	19.0	6.974
1.1	1.008	6.6	2.603	20.0	7.328
1.2	1.011	6.8	2.673	21.0	7.681
1.3	1.015	7.0	2.743	22.0	8.034
1.4	1.020	7.2	2.813	23.0	8.387
1.5	1.026	7.4	2.884	24.0	8.741
1.6	1.033	7.6	2.954	25.0	9.094
1.7	1.042	7.8	3.024	26.0	9.447
1.8	1.052	8.0	3.094	28.0	10.15
1.9	1.064	8.2	3.165	30.0	10.86
2.0	1.078	8.4	3.235	32.0	11.57
2.2	1.111	8.6	3.306	34.0	12.27
2.4	1.152	8.8	3.376	36.0	12.98
2.6	1.201	9.0	3.446	38.0	13.69
2.8	1.256	9.2	3.517	40.0	14.40
3.0	1.318	9.4	3.587	42.0	15.10
3.2	1.385	9.6	3.658	44.0	15.81
3.4	1.456	9.8	3.728	46.0	16.52
3.6	1.529	10.0	3.799	48.0	17.22
3.8	1.603	10.5	3.975	50.0	17.93
4.0	1.678	11.0	4.151	60.0	21.47
4.2	1.752	11.5	4.327	70.0	25.00
4.4	1.826	12.0	4.504	80.0	28.54
4.6	1.899	12.5	4.680	90.0	32.07
4.8	1.971	13.0	4.856	100.0	35.61
5.0	2.043	13.5	5.033		

It is frequently useful to know the largest diameter of wire of differ materials which will give a ratio of R_f/R_0 of 1.01 for different frequenci. For a ratio of R_f/R_0 equal to 1.001, the diameters given below shou be multiplied by 0.55; and for R_f/R_0 equal to 1.1, the diameters should be multiplied by 1.78.

11. Reduction of Skin Effect. In view of the tendency of the current to crowd to the surface of the conductor at high frequencies, the remedies which have been found practical in effecting an improvement in the resistance ratio R_f/R_0 have been those in which the conductor has been designed so that it presents a skin to the current flow. These are:

1. *Use of Flat Copper Strip.* While skin effect is present, for the same cross-sectional area a flat strip gives a lower resistance ratio than do round conductors.
2. *Use of Tubular Conductors.* Here the external magnetic field is much greater than the internal field, and therefore all parts of the conductor are affected alike by the field, thus reducing the skin effect.
3. *Use of Litzendraht.* According to Eq. (6) the smaller the diameter of the wire the less the skin effect. Litzendraht is a braided cable made up of a large number of fine strands of wire. When certain precautions are taken this braid shows a very much lower resistance ratio than does a solid copper wire of equal section. These precautions are:
   a. Each strand must be thoroughly insulated from every other strand to avoid contact resistance.
   b. Braiding must be such that each strand passes from the center to the outside of the conductor at regular intervals—a sort of transposition. This ensures that all strands are affected alike by the magnetic flux.
   c. Each strand must be continuous.

12. Types of Resistors. Resistors generally used in radio and allied applications may be broadly classified as:

1. Fixed resistors.
2. Variable resistors.

10. Maximum Diameter of Wires for H-f Resistance Ratio of 1.01.

Frequency, kilocycles...... Wave length, meters......	100 3,000	400 750	1,000 300	1,600 187.5	2,000 150	3,000 100	6 Mc 50	10 Mc 30	20 Mc 15	60 Mc 5	300 Mc 1
Material	Diameter, centimeters										
Copper	0.0356	0.0177	0.0112	0.0089	0.0079	0.0065	0.00457	0.00355	0.00251	0.00145	0.00065
Silver	0.0345	0.0172	0.0109	0.0086	0.0077	0.0063	0.00445	0.00346	0.00267	0.00141	0.00063
Gold	0.0420	0.0210	0.0133	0.0105	0.0094	0.0077	0.00543	0.00422	0.00298	0.00172	0.00077
Platinum	0.1120	0.0560	0.0354	0.0280	0.0250	0.0205	0.01445	0.0112	0.00783	0.00456	0.00205
Mercury	0.264	0.132	0.0836	0.0661	0.0591	0.0483	0.03416	0.0265	0.0187	0.0108	0.00485
Manganin	0.1784	0.0892	0.0564	0.0446	0.0399	0.0325	0.02300	0.0179	0.0126	0.0073	0.00326
Constantan	0.1892	0.0946	0.0598	0.0473	0.0423	0.0345	0.02440	0.019	0.0134	0.00775	0.00346
German silver	0.1942	0.0970	0.0614	0.0485	0.0434	0.0354	0.02500	0.0195	0.0138	0.00784	0.00354
Graphite	0.765	0.383	0.242	0.191	0.171	0.140	0.0988	0.0767	0.0542	0.0312	0.0140
Carbon	1.60	0.801	0.506	0.400	0.358	0.292	0.2065	0.16	0.1135	0.0655	0.0292
Iron											
μ = 1,000	0.00263	0.00131	0.00083	0.00066	0.00059	0.00048	0.000339	0.000263	0.000186	0.000107	0.000048
μ = 500	0.00373	0.00187	0.00118	0.00093	0.00084	0.00068	0.00049	0.000374	0.000264	0.000152	0.000068
μ = 100	0.00838	0.00418	0.00264	0.00209	0.00186	0.00152	0.001075	0.000836	0.000591	0.000341	0.000153

Each of these groups may be further classified on the basis of the nature of the conducting material of the resistor, as

1. Wire wound.
2. Composition (employing carbon).

13. Fixed Wire-wound Resistors. As commonly made, these are wound on (1) ceramic forms, (2) strips of fiber or bakelite, and (3) cores of textile cord or glass fiber. These windings are then embedded in a covering or coating for protective purposes. The nature of the covering depends upon the core and power rating of the resistor. The characteristics of the wire-wound resistor are those of the particular wire employed and generally show a negligible or slight temperature coefficient and no voltage coefficient, *i.e.*, the resistance is independent of the applied voltage. Wire-wound resistors are used in radios at powers ranging from less than ½ watt to 200 watts or more. To cover this wide band, different designs and structures are used, which for convenience may be classified as low-, medium-, and high-power resistors, which correspond to the core structures (1), (2), and (3) above.

14. Protective Coatings for Wire-wound Resistors. Coatings on wire are employed to protect the windings from mechanical injury, to prevent electrolytic effects and consequent corrosion due to penetration of moisture, and to provide an insulating covering for the winding. Coatings most widely used in practice are as follows:

A. Vitreous enamel coatings.
B. Cement coatings employing inorganic binders.
C. Cement coatings employing organic binders.
D. Molded bakelite.

Coatings in the first two classifications *A* and *B*, are capable of withstanding temperatures in excess of 250°C. without deterioration. They afford a high measure of protection against humidity. Exception to the latter statement are coatings employing sodium silicate (water glass) binders which are highly hygroscopic and, therefore, unsuitable where resistance to humidity is an important factor.

Coatings in classification *C* are capable of withstanding temperatures up to about 175°C., this varying with the nature of the binder. Resinous binders stand lower temperatures than asphaltic binders. They are, however, superior to the higher temperature coatings in their moisture-resistant properties.

Coverings of the last classification, *D*, are capable of withstanding temperatures from 100°C. to 160°C., depending upon the nature of the bakelite used. The ordinary general-purpose molding materials with wood-flower base are good for the lower temperatures, whereas the asbestos- or mica-filled bakelite is good for the higher temperatures.

15. Rating Wire-wound Resistors. In view of the low temperature coefficient of the resistance wires generally employed in radio wire-wound resistors, the resistance change with loads normally encountered is small. The rating is, therefore, primarily determined by the power the resistor can dissipate continuously for an unlimited time without excessive temperature rise or deterioration of the resistor. Some manufacturers rate resistors on the basis of the power that will produce a temperature rise of 300°C. in an ambient temperature of 40°C., when the resistor is mounted in free air. Such perfect ventilation conditions are seldom encountered. As a result, it is generally recommended that such resistors be used at one-fourth to one-half the nominal rating, which results in a temperature rise of 100°C. to 150°C. In practice even these temperature rises may be excessive owing to such factors as poor ventilation, proximity of resistors to parts which may not be subjected to elevated temperatures, and Fire Underwriter's approval. The specific application, therefore, limits the practical use of a resistor rather than any nominal rating.

16. Factors Influencing Rating of Wire-wound Resistors.

1. Heat-resistant properties of protective coating.
2. Heat-resistant properties of winding core. (Ceramic cores are most widely used, which withstand very high temperatures.)
3. Use of intermediate taps. Taps reduce effective winding space, resulting in less active cooling surface, reducing the nominal rating. The extent of reduction depends upon length of the resistor, being smaller for long units than for short ones. On short units 2 in. long the rating may be reduced by as much as 15 to 20 per cent, whereas on long units 6 in. long the reductions may be 3 to 5 per cent.

17. Types of Resistors. 1. *Low-power Resistors.* These units dissipate ½ to 1 watt per square inch of surface. There are two general types (*a*) flexible resistors and (*b*) bakelite-molded resistors. Both have either a core of cord or glass fiber. The former has a textile or glass-fiber covering, the latter is molded in bakelite. The latter are made in sizes having ratings of ½, 1, and 2 watts, corresponding to dimensions of the order of 3⁄16 in. diameter by ⅝ in. long, ¼ in. diameter by 1¼ in. long, and 5⁄16 in. diameter by 1¾ in. long. They are equipped with wire leads making them very convenient for so-called *point to point* wiring in circuits, eliminating the necessity for special fittings for mounting. The flexible resistors with glass-fiber cores and coverings are capable of much higher ratings. Low-power resistors are used largely as biasing resistors, isolation resistors, and voltage-dropping resistors.

2. *Medium-power Resistors (Flat Wire-wound Type).* These units dissipate between 2 and 4 watts per square inch. They consist of wire wound on strips of fiber or laminated bakelite to which lug terminals are attached at appropriate points. The strip is covered with bakelite, either by molding or other means. This assembly is then tightly enclosed in a sheet-metal punching with mounting holes, or a metal mounting strip is attached in intimate engagement with one side of the resistor, enabling the other side to be mounted flat against a metal chassis.

This design has many advantages. It is easy to mount. The metal enclosure, or mounting strip, and the chassis act as heat distributors, preventing excessive differences in temperature along the length of the unit. Use is made of the metal chassis and metal mounting to conduct heat away from the resistor, which enables higher power ratings for a given temperature rise.

They are used in the power range from 2 to 20 watts. Lengths vary from 2 to 6 in. Widths vary from ½ to ¾ in.

3. *High-power Resistors.* These are wound on cylindrical ceramic cores and have cement or vitreous enamel coatings. When inorganic cement or vitreous coatings are used, they are made to handle powers from 5 to 200 watts, depending upon the size of the unit, at dissipations of 5 to 10 watts per square inch of surface. These ratings are based on 250°C. temperature rise.

When organic cement coatings are used, they are made to handle powers from 4 to 80 watts, depending upon the size of the unit, at dissipations of 2 to 4 watts per square inch of surface. These ratings are based on a temperature rise of 125°C. This coating is used primarily when maximum protection is desired against humidity.

18. Temperature Rise of Wire-wound Resistors. Figure 3 shows the temperature rise to be expected at various loadings of wire-wound resistors wound on ceramic forms, with vitreous-enamel and cement coverings. The 100 per cent rating is based on manufacturers' rating of 250°C. rise in open air for class *A* and *B* coatings (Art. 14) and 125°C. rise in open air for class *C* coating. Temperature is measured at the center of the outer surface of the resistor.

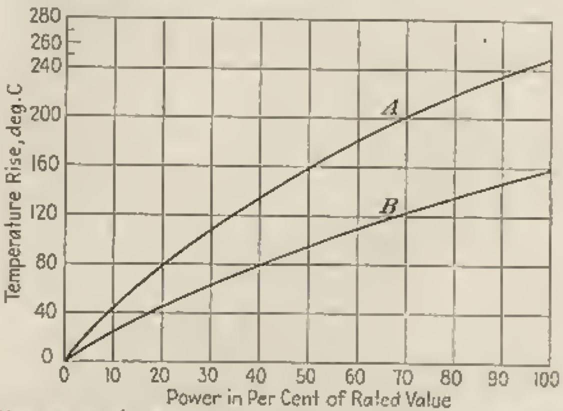


FIG. 3.—Temperature rise of wire-wound resistors. *A*, vitreous enamel or inorganic cement; *B*, organic cement covering.

19. Variable Wire-wound Resistors. These are usually of the continuously variable type, made by winding resistance wire on a flat strip of fiber, bakelite, or other insulating material. This strip may be formed into an arc and placed in a protecting container. A metallic sliding arm is arranged to travel over the winding, thus making contact with each turn as it is rotated. The choice of wire and size is determined by the resistance and space requirements.

In general, wire-wound continuously variable resistors are wound so that the resistance changes uniformly with the motion of the sliding contact. For certain uses, *e.g.*, antenna-type volume controls, it is desirable that the resistance change be non-uniform. In this case the form on which the wire is wound is sometimes tapered so that the resistance per degree rotation is not constant. Other methods of tapering employed are winding with variable pitch, winding sections of the control with different sizes of wire, and copper plating start and finish of the winding. Some of the factors to be considered in design are as follows:

1. Contact between slider and resistor element should be positive.
2. Winding should not become loose on the form.

3. Sliding contact should not wear away resistance wire.
4. Resistance change per turn should be as small as possible.
5. Slider material should be such that it will not oxidize.

20. Composition-type (Radio) Resistors. The term *composition-type resistor* is employed to cover that group of resistors in which a conductor is mixed with binder in definite proportions and suitably treated to produce a resistor material. This type of resistor has attained a wide popularity because of the following advantages: (1) Flexibility in range—it may be made in any value up to several megohms; (2) compactness—its physical dimensions are small for any range; they may be made in sizes as low as ⅛ in. diameter by ⅜ in. long.

Numerous types of these resistors have been produced, but they take two general forms:

1. *Solid-body Resistor.* In this type the resistor material is extruded, pressed, or molded into its final physical form, which generally is a solid rod, after which it may be subjected to some form of heat treatment. The so-called *carbon* resistors are examples of this type.

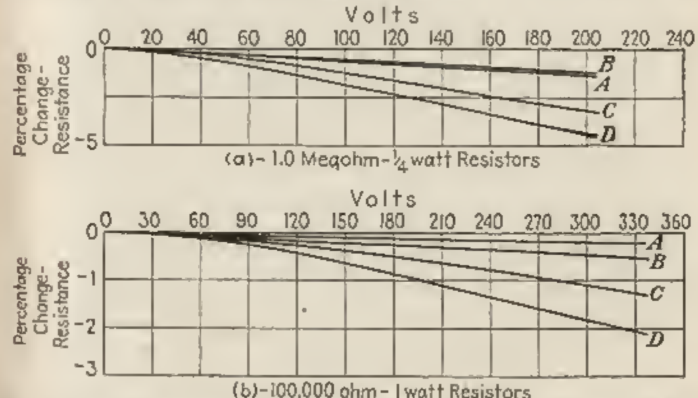


FIG. 4.—Voltage characteristic of various resistors. Curves *A* are metallized-filament type; others are carbon type.

2. *Filament-coated Resistors.* In this type a conducting coat or film is baked on the surface of a continuous glass filament or other form. In the case of the glass filament this is completely enclosed in an insulating tube. The so-called *metallized-filament* resistors are examples of this type.

21. Characteristics of Composition-type Resistors. Composition-type (commercially known as *radio*) resistors possess properties differing very markedly from those of metallic resistors. The most important ones are as follows and are possessed by all these types in varying degree:

1. *Voltage Characteristics.* The resistance is not independent of the applied voltage and generally falls with increasing voltage. Typical curves showing the manner in which the resistance varies with voltage (heating effect due to load not present or corrected for) are shown in Fig. 4.

The percentage change of resistance at a given voltage measurement referred to its resistance at some low voltage such as 1½ volts has arbitrarily been called the *voltage coefficient*. This coefficient increases as the physical size of the resistor decreases and increases with the resistance value. It is also a function of the ingredients or mix employed in the resistor. Figure 5 shows for a given type of carbon resistor the relationship between voltage coefficient and size and value of the resistor. The test voltage at which each measurement was made is indicated for each value of resistance.

2. *Radio-frequency Characteristics.* Unlike wire-wound resistors, composition-type resistors decrease in value with increasing frequency. This effect is very marked in the high-valued resistors such as 1 megohm but is absent, or very small, in the low values such as 100,000 ohms and under. The effect decreases with the diameter of the active resistor element. Skin effect is not the factor which determines this characteristic. Two factors play a prominent part here as follows: (1) the shunting effect of the indi-

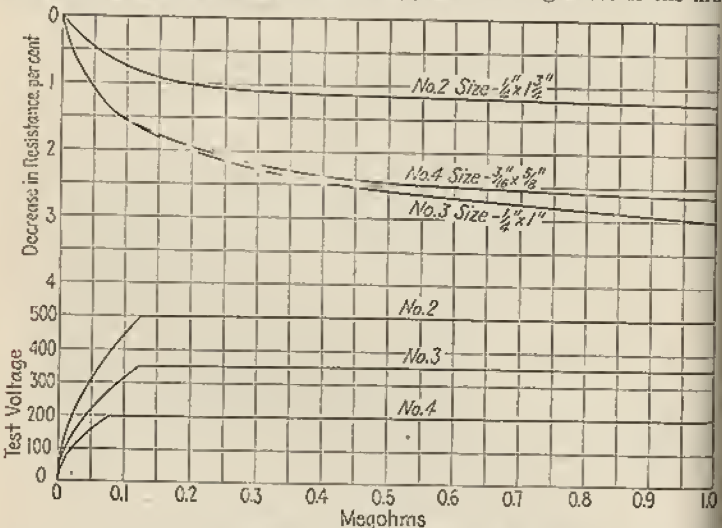


FIG. 5.—Voltage coefficient of carbon resistors.

vidual capacities between conducting masses in the resistor element tends to reduce the effective resistance; (2) the dielectric in binder and fillers of these resistors and their housings introduces losses with increasing frequency which likewise act to reduce the resistance.

3. *Humidity Characteristics.* The effect of humidity in general is to cause a rise of resistance. This effect may sometimes be reduced by suitable treatment.

4. *Noise.* These types of resistors all show, in varying degree, the presence of microphonic noise. The degree of noise is a function of the load, size of the resistor, and the nature of the materials used in the resistor. In general, for a given set of materials in the resistor, the noise level increases with increasing resistance and decreasing size of the resistor. Figures 7*a* and 7*b* show typical noise-level curves for two makes of resistors. The change in each curve or the point of discontinuity shows where a change of mix or materials was made. The curves also show the increase in noise for a given value as the resistor size decreases. Noise measurements were made in accordance with the method described in Art. 25.

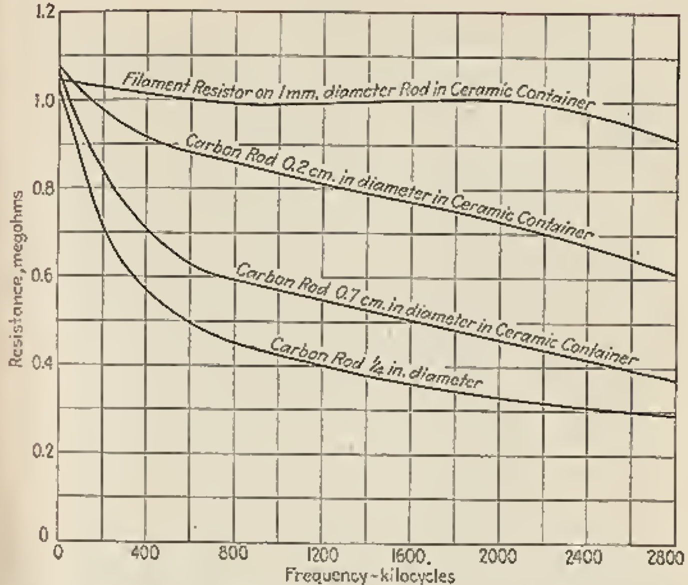


FIG. 6*a*.—Resistance-frequency characteristics of various types of 1-megohm resistors up to 3 megacycles (University of Wisconsin CWA project E-16-5).

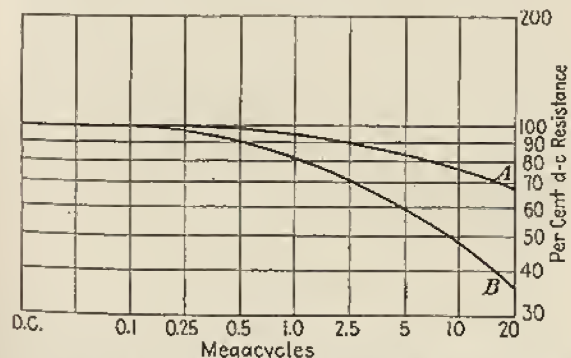


FIG. 6*b*.—Characteristic for filament-type resistor, carried to 20 megacycles, in two different insulating housings.

22. Rating Composition-type Resistors. The rating of composition-type resistors is a more complicated matter. The temperature coefficient of this type of resistor being larger, it is possible for a resistance change to become quite appreciable before a temperature limitation is exceeded.

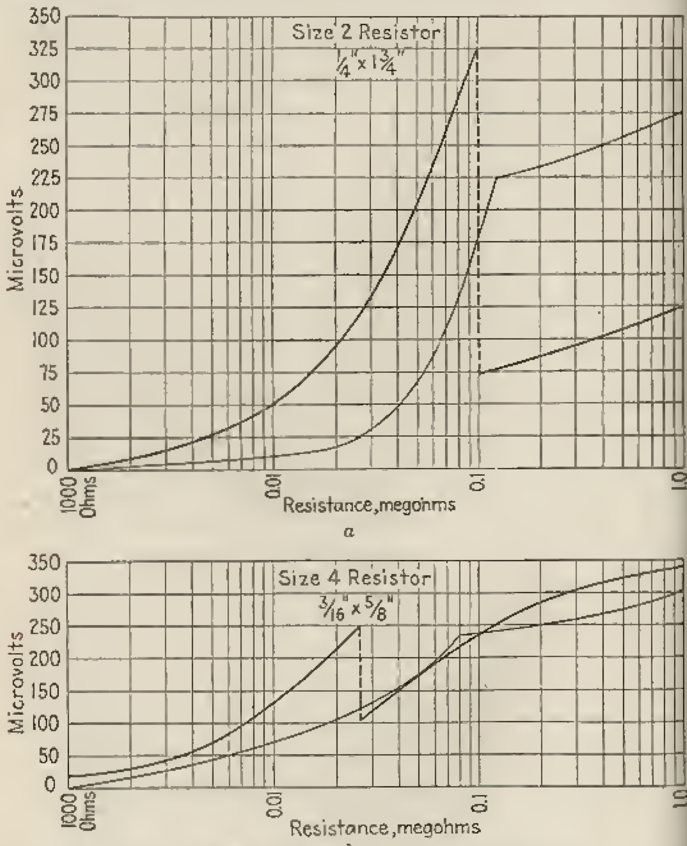


FIG. 7.—Noise-level curves of typical resistors.

Furthermore, with the higher ranges, such as 0.25 megohm and over, in which the power dissipation may be very low, the voltage characteristic may be a determining factor instead of the load-carrying characteristic. It is therefore customary to rate this type of unit on the basis of the maximum load it can carry, or the maximum voltage which can be applied to it, without exceeding prescribed resistance changes. The prescribed

changes generally accepted are 5 per cent for intermittent rated-load operation and 10 per cent for 50 per cent overload operation.

As a result of recent developments, notably the development of insulated resistors in which the resistance element is molded in bakelite and also the development of new mixes, it has been possible to increase the rating of given sizes of resistors. It has especially been possible to increase the rating of the smaller sizes of insulated resistors. This was made possible by the discovery that the temperature rise of the very short resistors is appreciably lower for a given power dissipation than would be expected from its reduced cooling surface. The reason for this is that the metal end terminals, because of shortness of the unit, cover a substantial portion of the entire resistor and are very close to the center hot section, and therefore cool the resistor by conducting the heat away.

The following table gives the most generally adopted standard ratings and sizes of insulated resistors, as these resistors are definitely replacing the non-insulated type:

Rating, watts	Diameter of resistor, inches	Over-all length of resistor, inches
$\frac{1}{4}$	$\frac{1}{8}$	$\frac{3}{8}$
$\frac{1}{2}$	$\frac{1}{8}$ to $\frac{3}{16}$	$\frac{3}{8}$ to $\frac{5}{8}$
1	$\frac{1}{4}$	$\frac{7}{8}$ to $1\frac{1}{4}$
2	$\frac{5}{16}$	$1\frac{3}{4}$

23. Composition of Resistors. Radio resistors of the carbon and filament types generally employ a conducting material of high specific resistance mixed with a filler and binder. The most widely used conducting material is some form of carbon or graphite. The fillers and binders employed vary with the type of resistor. Examples of these are clay, rubber, and bakelite. The filler, binder, and conductor are mixed in various proportions to obtain resistors having different ranges. The method of making the resistor varies also with its type. The solid-body types are generally either molded or extruded. The filament resistor is made by baking the resistance material on a glass rod which is sealed in a ceramic or bakelite container.

24. R.M.A. Color Code. The use of resistors has increased to such an extent and so many are employed in a radio set that it has become desirable to identify each resistor for range in a quick and simple manner. Such identification simplifies assembly of these units in radio sets and helps in servicing. A color code has therefore been adopted by the Radio Manufacturers' Association.

This color code takes into consideration the fact that composition resistors are made with leads coming out at right angles to the axis of the resistor and also with leads brought out of the ends axially. The color code also enables the tolerance of the resistor to be identified. The color code standard follows:

Ten colors shall be assigned to the figures as shown in the table below in which cable designations indicate the color shades as shown on the Standard Color Card of America, 8th ed., 1928, issued by the Textile Color Card Association of the United States.

Figure	Color	Color to be equivalent to
0	Black	
1	Brown	Cable 60113
2	Red	Cable 60149
3	Orange	Cable 60041
4	Yellow	Cable 60187
5	Green	Cable 60105
6	Blue	Cable 60102
7	Violet	Cable 60010
8	Gray	Cable 60034
9	White	

It shall be standard in fixed-composition resistors with radial leads to indicate the nominal resistance value of the resistor in accordance with the following system and diagram:

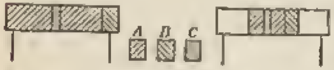


FIG. 8.—Standard resistor of R.M.A.

The body *A* of the resistor shall be colored to represent the first figure of the resistance value. One end *B* of the resistor shall be colored to represent the second figure. A band, or dot, *C* of color, representing the number of ciphers following the first two figures, shall be located within the body color. Two diagrams (Fig. 8) illustrate two interpretations of this standard, both of which are deemed to be in accordance with the standard.

Examples illustrating the standard are as follows:

Ohms	*A*	*B*	*C*
10	Brown 1	Black 0	Black, no ciphers
200	Red 2	Black 0	Brown, one cipher
3,000	Orange 3	Black 0	Red, two ciphers
3,400	Orange 3	Yellow 4	Red, two ciphers
40,000	Yellow 4	Black 0	Orange, three ciphers
44,000	Yellow 4	Yellow 4	Orange, three ciphers
43,000	Yellow 4	Orange 3	Orange, three ciphers

It shall be standard in making fixed composition resistors with axial leads to indicate the nominal resistance value of the resistors by bands of color around the body of the resistor, in accordance with the following system:

Three or more bands of color shall provide indications as follows:

Band *A* shall indicate the first significant figure of the resistance of the resistor.

Band *B* shall indicate the second significant figure.

Band *C* shall indicate the decimal multiplier.

Band *D*, if any, shall indicate the tolerance limits about the nominal resistance value.

FIG. 8*a*.—Fixed composition resistor with axial leads.

It shall be standard to indicate the significant figures of the resistance value, the decimal multiplier, and the tolerance with the modifications and extensions of the Standard R.M.A. Color Code M4-213 as given below:

Color	Decimal multipliers			Tolerance per cent
	Significant figure	Power of 10	Multiplying value	
Black	0	10^0	1	
Brown	1	10^1	10	
Red	2	10^2	100	
Orange	3	10^3	1,000	
Yellow	4	10^4	10,000	
Green	5	10^5	100,000	
Blue	6	10^6	1,000,000	
Violet	7	10^7	10,000,000	
Gray	8	10^8	100,000,000	
White	9	10^9	1,000,000,000	
Gold	.	10^{-1}	0.1	± 5
Silver	.	10^{-2}	0.01	±10
No color	..	....	....	±20

25. Test Specifications. Over the last few years, a series of tests have been developed which are designed to establish the performance merit of composition resistors. While these tests have not been established as standard, they have gradually been adopted by the leading manufacturers as the basis of specifications for composition resistors. These tests are as follows:

Resistance Measurements. Unless otherwise specified it shall be standard to measure the resistance under the same voltage drop as normally exists across the resistor in the application for which it is intended.

The readings are to be made as quickly as possible at 20°C., preferably with a limit-bridge circuit arrangement so that the resistors do not have an opportunity of undergoing an appreciable temperature rise due to the current passing through them under the conditions of the test.

Normal-load Life Test. It shall be standard to make normal-load life tests by placing the resistors on load intermittently 1½ hr. on and ½ hr. off at an ambient temperature of 40°C., for 1,000 cycles or 2,000 hr. at the voltage representing the rating of the resistor as specified by the resistor manufacturer. Any readings taken should be made by uniform method at the end of ½-hr. off period. The results of this test shall be plotted, showing the per cent permanent change in resistance *versus* time in hours.

Either direct or alternating voltage may be used in the foregoing tests depending on how the resistors are intended to be used.

It shall be standard for the resistor manufacturer to state the rated potential in direct voltage with a supplementary rating on alternating voltage when requested.

Load Characteristics. It shall be standard to plot these characteristics, showing the per cent change in resistance values *versus* loads in watts, making readings at 10 per cent intervals up to 100 per cent overload value or up to the maximum rated voltage as specified by the resistor manufacturer, conducting the tests at an ambient temperature of 40°C., and allowing a minimum of 15 min. at constant load immediately preceding each reading, so that the resistor comes up to equilibrium temperature conditions after each change in load. The resistors are to be exposed 1 hr. at 40°C., before starting the test. Each reading is to be made under steady-state hot conditions at the voltage drop existing for the particular wattage setting.

Voltage Characteristics. It shall be standard to plot voltage-characteristic curves, making readings with uniform voltage increments up to a maximum voltage representing 100 per cent overload in watts on the resistor or up to the maximum voltage rating of the resistor. The resistors are to be at 40°C. for 1 hr. before starting the test, and readings are to be made as quickly as possible so that the resistors do not have an opportunity to heat under the conditions of the test. The resistors are to be connected in the circuit only during a period of time sufficient for making resistance determinations.

Humidity Test. It shall be standard to expose resistors to a relative humidity of 32 per cent at an ambient temperature of 40°C. for 150 hr., at which time the resistance value is recorded. The resistors then are to be exposed to a relative humidity of 90 per cent for 300 hr. with an ambient temperature of 40°C., and the final resistance value is to be recorded. Finally, the resistors are again subjected for 150 hr. to a relative humidity of 32 per cent at 40°C., and a final reading taken at the end of this period. The readings are to be made at 20°C. by uniform method not later than 30 min. and not less than 15 min. after the resistors have been removed from the humidity chamber.

It is recommended that the resistors be suspended in an enclosed chamber over a saturated solution of cupric chloride or sodium tartrate for the 90 per cent relative humidity condition and over a saturated solution of magnesium chloride for the 32 per cent relative humidity condition.

On account of the difficulty in obtaining quantitative results on humidity tests, it is recommended that the various resistors involved should be tested together at the same time under exactly the same conditions.

Overload Tests. It shall be standard to make overload tests with a 50 per cent overload on the resistors for 100 hr. at an ambient temperature of 40°C. Resistance measurements are to be made by uniform method before commencing the overload test but after the resistors have been at 40°C. for ½ hr. Resistance measurements are again to be made, under the same conditions, ½ hr. following the completion of the test. The differences between the initial readings and final readings are to be expressed as per cent permanent change in resistance.

Aging Tests. It shall be standard to make an aging test wherein the resistors are kept under standard conditions of 40°C. ambient temperature and 32 per cent relative humidity for a period of 90 days. Readings are to be taken at intervals by uniform method so that a curve can be plotted showing the per cent change in resistance *versus* time in days.

It is recommended that the standard conditions in the foregoing be attained by means of an enclosed chamber containing a saturated solution of magnesium chloride, further, that the resistors be suspended over the solution as specified under humidity test.

If shelf tests are made, it shall be standard to test all the resistors together under identical conditions. Results of one test should not be compared with another unless the time, temperature, and humidity cycles are precisely the same.

Noise Test. It shall be standard to test resistors for noise, using resistors having the same value tested under the voltage drop normally existing in the application for which they are intended. A resistance-type amplifier is to be used with a resistance input circuit, the entire combination to be as independent of frequency as is possible. A visual instrument, such as an r-m-s vacuum-tube voltmeter, shall be used on the output of the amplifier. An aural test, using a loud-speaker on the output of the amplifier, should also be used in conjunction with the foregoing.

A circuit arrangement, such as shown in Fig. 9, shall be used. In this circuit arrangement E represents an adjustable voltage source of constant value; C a large by-pass condenser; R represents an adjustable, standard, quiet resistor, such as a laboratory decade box; X represents the unknown under test; R_1 is a calibrated potentiometer; S is a source of a-c supply of 1,000 cycles; and V in both cases represents an indicating voltmeter. In operation it shall be standard to first connect the resistor as shown, adjusting R to have approximately the same resistance value as the unknown under test. E is then adjusted until the voltage normally existing across the resistor, in the application for which it is used, is placed across the terminals of X. This voltage is, of course, one-half that shown on the voltmeter when R is adjusted to be exactly the same as X. The switch on the output of the amplifier is placed on the tube voltmeter setting, and the switch on the input is connected across the unknown resistor. The gain of the amplifier is adjusted to obtain a definite deflection on the vacuum-tube voltmeter, after which it is not changed. The input switch is then thrown to the calibrated potentiometer setting, and the setting of the potentiometer is adjusted until the reading of the tube voltmeter on the output of the amplifier is the same as before. The setting of the calibrated potentiometer, which is calibrated in microvolts, shows the equivalent r-m-s voltage variation existing across the particular unknown resistor being tested. It can then be stated that the noise of the resistor is equivalent to so many microvolts r.m.s. for the particular voltage drop existing across the same.

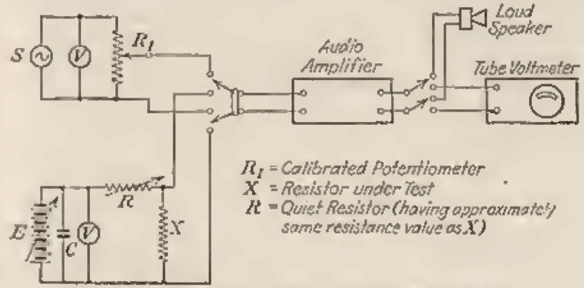


FIG. 9.—Circuit for resistor-noise measurement.

26. Acceptable Performance. On the basis of these specified tests the following is considered acceptable performance:

1. Life Test	5 per cent change or less
2. Overload	10 per cent change or less
3. Humidity	10 per cent change or less
4. Noise	500 μv or less

27. Representative Values of Resistors Employed in Radio Sets. The range of resistors usually employed in radio sets extends from 1 ohm up to 20 megohms. These resistors are used for various purposes, such as providing grid bias to radio, audio, and detector tubes; plate coupling; voltage dividers; and filters. Typical values employed for these various applications are enumerated below:

1. Detector bias resistors	5,000 to 50,000 ohms
2. Power bias resistors	200 to 3,000 ohms
3. Voltage dividers	1,000 to 100,000 ohms
4. Plate-coupling resistors	50,000 to 250,000 ohms
5. Grid leaks	100,000 to 20 megohms
6. Filter resistors	100 to 100,000 ohms

28. Variable Carbon-type Resistors. In numerous radio applications high variable resistors are required, *e.g.*, for controlling the sensitivity of a receiver by varying the C bias on the r-f tubes a variable resistor up to 50,000 ohms maximum is commonly employed. For adjusting the audio signal level in automatic volume control sets a variable resistor up to 2.5 megohm is not uncommon. From the point of view of cost, wire-wound resistors of this order of magnitude are prohibitive. Further more it is desirable to have a non-uniform rate of change resistance with respect to angular rotation, which is very difficult to secure with wire-wound resistors. Carbon or graphitic types of variable resistor that can be made to meet these requirements at reasonable cost are therefore widely used. Such resistors generally consist of a resistive solution applied to some flat form, such as paper, bakelite, or ceramic and baked on. A rotating slider or some other form of contact travels over this resistive element producing a continuous variation of resistance. Since the resistor is essentially painted on the form, its geometrical form may be varied by design. Also different concentrations of the

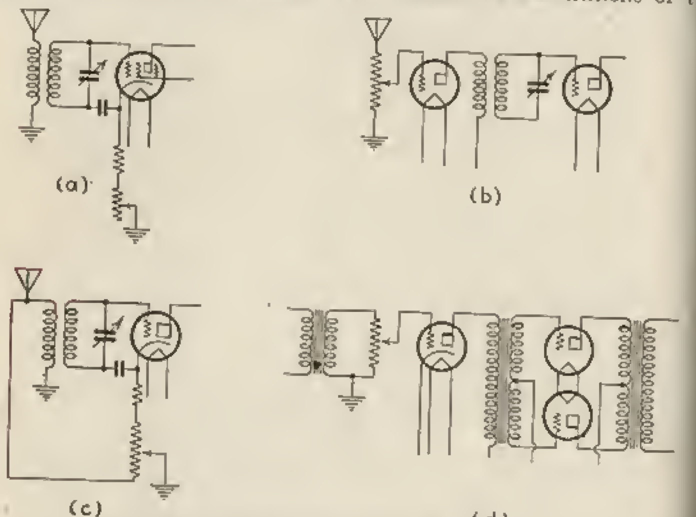


FIG. 10.—Typical uses of variable resistors.

resistor ink or paint may be employed at different positions of the resistor element. By the use of these two expedients the resistor may be designed to give any variation of resistance desired.

29. Uses for Variable Carbon Resistors. Within their power limitation these resistors may be used wherever a continuously variable resistor is required. They may be used as either potentiometers or rheostats. They find their widest use as volume controls and tone controls in radio receivers. Some of their specific uses are here listed, and the basic circuits illustrating these uses are shown in Fig. 10.

1. Sensitivity control for radio receivers, by varying control-grid or screen-grid potentials of r-f tubes (Fig. 10*a*).
2. Antenna control for varying r-f input to antenna tube (Fig. 10*b*).
3. Sensitivity and antenna input control, combination of Figs. 10*a* and 10*b* (Fig. 10*c*).
4. Audio-level control (Fig. 10*d*).

5. Combination load-resistor and audio-level control in diode rectifier circuit.
6. Tapped volume control for acoustic compensation at low levels. Tuned circuits are shunted across one or more taps to produce varying degrees of a-f compensation at different levels.
7. Gain controls and faders for phonograph and a-f amplifiers.
8. Tone control in a-f amplifiers for varying a-f frequency characteristics.
9. High-frequency variable resistor when non-reactive feature is essential, as in signal generator attenuators.
10. Television controls, such as brightness, contrast, focusing.

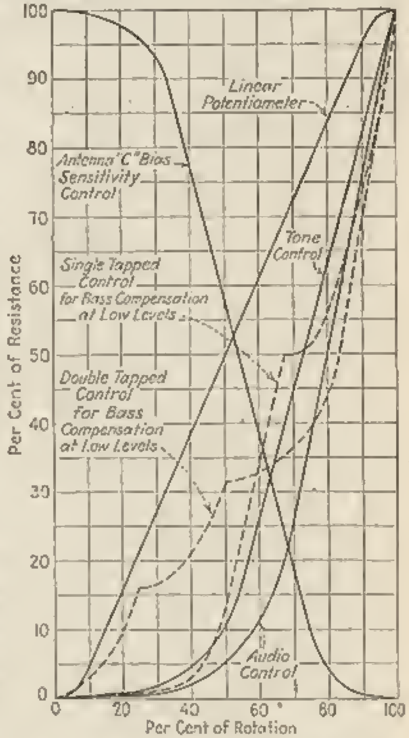


FIG. 11.—Taper curves of variable resistors.

30. Tapers. The circuit considerations involved in these applications are discussed elsewhere in this handbook, particularly in the section on Receiving Systems. However, each of these applications calls for a resistance curve, or *taper* as it is termed, which is most suitable for it. This taper defines the law of resistance changes *versus* angular rotation of the variable arm. Some widely used curves are given in Fig. 11.

A suitable specification defining the taper should include:

1. Curve showing resistance variation against active angular rotation of the contactor. Where a switch is incorporated in the variable resistor, the angle taken up for operation of the switch is considered inactive. Curve should indicate whether resistance increases with clockwise or counterclockwise rotation.
2. Resistance at extreme counterclockwise end between variable arm and left terminal; this is generally called *left terminal minimum* and is specified as "less than so many ohms."
3. Resistance at extreme clockwise end between contactor and right terminal; this is generally called *right terminal minimum* and is specified as "less than so many ohms."
4. When a tap is specified, the angular location and resistance of the tap should be given. The resistance between the tap terminal and the variable arm, when located at the tap, is sometimes specified.

31. Choice of Volume-control-resistance Curve.[1] In an audio amplifier in which the maximum output is 40 db above the minimum output,

[1] By the editor.

the volume control should be so made that each $\frac{1}{40}$ of the rotation should correspond to an attenuation of 1 db. If the volume control has a total attenuation of 80 db, more than is necessary on this particular amplifier, each $\frac{1}{40}$ of the rotation will correspond to 2 db attenuation since only half of the total rotation can be used. In the second case the control should be more critical than in the first case.

In a radio receiver the design of the volume control differs widely depending upon whether the receiver has automatic volume control or not. If not, the entire voltage gain of the receiver must be under control, perhaps 120 db. The tendency for the volume control to become

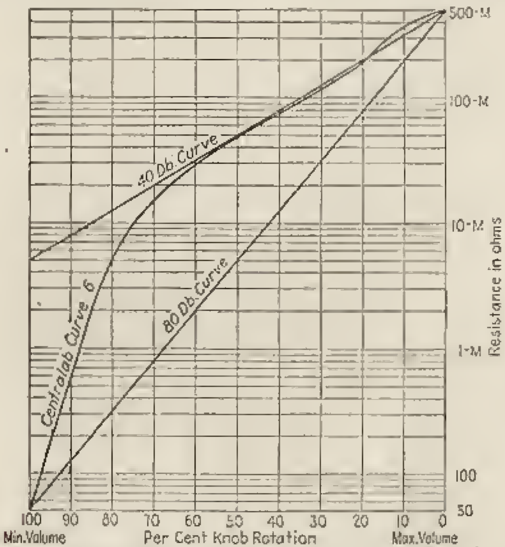


FIG. 12.—Advantage of special taper for volume control.

noisy or to be difficult to adjust without producing violent jumps of volume change increases with the total gain that must be controlled.

The fact that a-v-c systems cannot deliver a uniform voltage to the audio detector because of the wide variations of input voltage (ranging from a microvolt to several volts) makes necessary a different shape of attenuation curve than would be used on an audio amplifier used by itself. A type of curve (Centralab) useful in the a-v-c receiver is shown. Here, approximately uniform attenuation of 40 db is secured in 80 per cent rotation from the maximum volume. This is the range most often used. The departure from linearity in the first 15 per cent of rotation is to keep the resistance gradient within limits representing low noise.

Between 80 and 100 per cent rotation, the curve changes rapidly to provide a total attenuation of 80 db. Rapid attenuation in this region is accomplished without noise because the resistance change per decibel

is small. Such a curve is much more satisfactory than a straight logarithmic line (note the 80-db curve). In addition they are simpler to build. A tapered resistance curve such that equal increments in rotation produce equal increments in attenuation (a straight line when plotted against the logarithm of the resistance) requires that a change of 300,000 ohms take place in the first 10 per cent, 120,000 ohms in the second 10 per cent, and so on till the last 10 per cent rotation produces a change of only 75 ohms. This is true of a 500,000-ohm control with a total attenuation of 80 db.

32. Wear Characteristics. Variable carbon resistors necessarily have the same general electrical characteristics as fixed carbon resistors. In addition, owing to the motion of the slider on the resistance element, there is a certain amount of wear on the resistance element. This produces a change in resistance value and noise. Factors influencing these changes are as follows:

1. Hardness of resistance element which determines ability to withstand abrasion.
2. Pressure of moving contact on resistance element.
3. Smoothness of moving contact surface.

33. Specifications for Variable Resistors. No standard specifications have been established for variable resistor performance. A typical specification, however, representative of acceptable performance is here given.

1. *Endurance or Wear.* Life test: Units shall not fail before 10,000 complete operations when operated without electrical load. The unit shall be operated over its full range including operation of switch at a rate of approximately 1,000 operations per hour. Failure shall be considered as a change in resistance of greater than 15 per cent of the initial resistance or mechanical fracture of the switch.
2. *Noise.* Units shall be of such a nature as to produce no audible sound in the loud-speaker of the apparatus in which the unit is used.
3. *Humidity.* The resistance of units shall not show a temporary change of more than 25 per cent when conditioned 100 hr. at a temperature of 40°C. and a relative humidity of 90 per cent. Units shall be conditioned 24 hr. in a desiccator before placing in the humidity chamber.
4. *Resistance Curve.* The resistance curve and permissible variations over the entire resistance range of effective electrical rotation shall be in accordance with the drawing (as supplied by the purchaser). These curves shall be within the required limits and must conform in general shape to the nominal curve of the drawing.

References

Bur. Standards Circ. 74.
DRIVER-HARRIS COMPANY: *Bull.* R-40.
KENNELLY, LAWS, and PIERCE: Experimental Research in Skin Effect in Conductors, *Trans. A.I.E.E.*, 1915.
MORECROFT: "Principles of Radio Communication."
PENDER: "Handbook for Electrical Engineers."
WILBUR B. DRIVER COMPANY: "Resistance Handbook."

SECTION 4

INDUCTANCE

By Gomer L. Davies, B.S.[1]

1. Magnetic Flux. The property of electrical circuits called *inductance* depends upon the magnetic effects associated with a flow of electric current. In a magnetic system the magnitude of the force of magnetic attraction or repulsion is proportional to the product of the strengths of the poles and inversely proportional to the square of the distance between them. A *unit magnetic pole* is defined as that pole which repels a similar pole at a distance of 1 cm with a force of 1 dyne. The force between two poles acts along the line joining the poles. Consequently a unit north pole in the vicinity of a magnet is acted upon by two forces: one of repulsion, due to the north pole of the magnet; and one of attraction, due to the south pole. The resultant is the total force exerted by the magnet upon the unit pole. Thus the magnet is surrounded by a *field of force* or *magnetic field* whose direction and magnitude at any point are defined as the direction and magnitude of the force acting upon a unit north pole at that point.

If a unit north pole is allowed to move freely in a magnetic field, it will move in the direction of the field at each point and will trace out a path which is called a *line of force*. The total field is considered to be made up of a large number of such lines. In any region of space the total of all the lines of force in that region is called the *magnetic flux* in that region, and the number of lines of force passing through a unit area of a surface perpendicular to the direction of the field is the *flux density* and is determined by the strength of the field.

2. Magnetic Effects of Current-carrying Conductors. Magnetic effects are exhibited not only by magnets but also by wires carrying electric currents. The magnetic field near a straight current-carrying conductor consists of circular lines of force surrounding the conductor; the flux density at any point outside the wire is proportional to the current and inversely proportional to the distance of the point from the axis of the conductor. If the wire carrying the current is wound in one or more layers on a cylindrical form, the field inside of this coil is parallel to the axis of the cylinder and is proportional to the product of the current and the number of turns on the coil. This product of current (in amperes) and number of turns is called the *ampere-turns* of the coil. The flux density along the axis of the coil may be expressed as the product of the ampere-turns by a constant. If the winding is of infinite length, this constant is 4π.

[1] Engineer, Washington Institute of Technology, Washington, D. C.

3. Inductance—Definition and Units.[1] When the current in a circuit varies Ohm's law in the form in which it is stated for constant-current circuits, no longer serves to define the current.

The magnetic flux associated with the circuit varies with the current and induces a voltage in the circuit which is given by the equation

$$e = -\frac{d\phi}{dt} \tag{1}$$

where e is the induced voltage, ϕ the flux, and t the time. As the flux is proportional to the current, it may be written

$$\phi = Li \tag{2}$$

where L is a constant and i the current. Then

$$e = -\frac{d}{dt}(Li) = -L\frac{di}{dt} \tag{3}$$

If the current is increasing, the induced e.m.f. opposes the current, and work must be done to overcome this e.m.f. If the work is W,

$$\frac{dW}{dt} = ei = -Li\frac{di}{dt} \tag{4}$$

and

$$W = -\int_0^{i_0} Li\,di = -\frac{Li_0^2}{2} \tag{5}$$

i_0 being the final value of the current, the initial value being taken as zero.

The quantity L in these equations is the *coefficient of self-induction, self-inductance*, or simply *inductance* of the circuit. It may be defined in three ways: from Eq. (2), as *the flux associated with the circuit when unit current is flowing in it;* from Eq. (3), as *the back e.m.f. in the circuit caused by unit rate of change of current;* and from Eq. (5), as *twice the work done in establishing the magnetic flux associated with unit current in the circuit.* These three definitions give identical and constant values of L provided there is no material of variable permeability near the circuit, and provided the current does not change so rapidly that its distribution in the conductors differs materially from that of a constant current. If these conditions do not hold, L is not constant and the values obtained from the three definitions will in general be different.

The units used for inductance must conform to the units used for the other quantities used in the defining equations. The practical unit is the *henry*, which is the inductance of a circuit when a back e.m.f. of 1 volt is induced in the circuit by a current changing at the rate of 1 amp. per second. The relations between units are as follows:

$$\begin{aligned} 1 \text{ henry} &= 10^9 \text{ e.m.u.} \\ &= 1.1124 \times 10^{-12} \text{ e.s.u.} \end{aligned}$$

The henry is subdivided into two smaller units, the millihenry and the microhenry. The millihenry is one-thousandth of a henry, and the microhenry is one-millionth of a henry. The millihenry and microhenry are abbreviated mh and μh respectively. Thus

$$1 \text{ henry} = 1{,}000 \text{ mh} = 1{,}000{,}000\ \mu\text{h}$$

[1] Starling, S. G., "Electricity and Magnetism," Chap. XI, 1926.

The term "inductance" refers to a property of an electrical circuit or piece of apparatus but not to any material object. A piece of apparatus used to introduce inductance into a circuit is properly called an *inductor* or coil.

4. Current in Circuits Containing Inductance. If a circuit containing a source of constant e.m.f. and pure resistance only is closed, the current rises instantly to its full value as determined by Ohm's law. If the circuit contains inductance, a back e.m.f. of the value $L\frac{di}{dt}$ acts during the time the current is changing, so that, if the e.m.f. of the source is E, the actual e.m.f. available to force current through the resistance is $E - L\frac{di}{dt}$.

The equation for the current in the circuit is

$$E - L\frac{di}{dt} = Ri \tag{6}$$

or

$$L\frac{di}{dt} + Ri = E \tag{7}$$

The solution of this equation is

$$i = \frac{E}{R}\left(1 - \epsilon^{-\frac{Rt}{L}}\right) \tag{8}$$

The time t is reckoned from the instant at which the switch is closed, and ϵ is the base of natural logarithms.

At a time $t = L/R$ after the circuit is closed, the current has a value equal to $I_0\left(1 - \frac{1}{\epsilon}\right)$, or about 63 per cent of its final value. The quantity L/R is called the *time constant* of the circuit. The time constant, or the time required for the current to rise to a value of $1 - \frac{1}{\epsilon}$ times its final value, does not depend upon the actual values of inductance and resistance but only upon their ratio.

The current in such a circuit is shown in Fig. 1 for several values of L/R. Theoretically the current does not reach its maximum value I_0

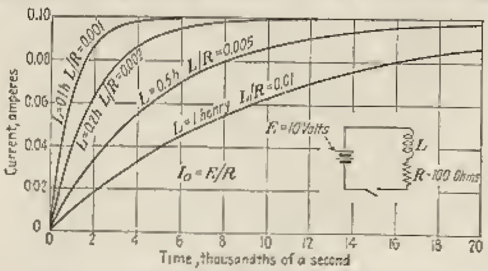


FIG. 1.—Rise of current in inductive circuit.

except at an infinite time after the circuit is closed, but practically the difference between the actual current and the value I_0 becomes negligible after a relatively short time.

If, after the steady current I_0 has been established in the circuit, the source of the e.m.f. is short-circuited, the current does not fall to zero instantly but decreases according to the equation

$$i = \frac{E}{R}\epsilon^{-\frac{Rt}{L}} \tag{9}$$

This equation is plotted in Fig. 2 for the same values of the circuit constants as were used in Fig. 1. In this case the time constant L/R

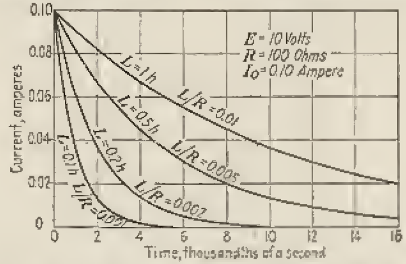


FIG. 2.—Fall of current in inductive circuit.

represents the time required for the current to fall to $1/\epsilon$ or about 37 per cent of its initial value.

If, instead of the source of e.m.f. being short-circuited, the circuit is opened, the resistance becomes extremely large and the current falls to zero almost instantly. As a result of this rapid change of current, a large e.m.f. is induced in the circuit, causing a spark or arc at the point at which the circuit is opened.

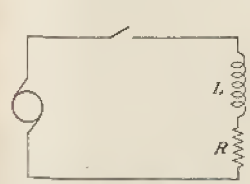


FIG. 3.—Series circuit containing resistance and inductance.

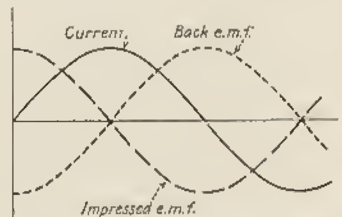


FIG. 4.—Phase relations in inductive circuit.

When the current in an inductive circuit is changing, a back e.m.f. other than that due to resistance acts in the circuit. This back e.m.f. is proportional to the current and to the quantity ωL, which is called the *inductive reactance* and usually written X_L. Also, the phase of the back e.m.f. is 90 deg. behind that of the current. To force a current through a pure inductance, therefore, requires an impressed e.m.f. 180 deg. out of phase with the back e.m.f., or one leading the current by 90 deg. (Fig. 4).

Now, if a sinusoidal e.m.f. is impressed on a circuit containing resistance and inductance in series (Fig. 3), the current in the circuit will also be sinusoidal, provided the resistance and inductance are independent of the current. The portion of the impressed e.m.f. required to force current through the resistance will be in phase with the current, while the portion required to force current through the inductance will lead the current by 90 deg. The resultant phase of the impressed e.m.f. with respect to the current will have some value between zero and 90 deg., depending upon the values of resistance and inductance in the circuit.

To determine mathematically the behavior of the circuit described above, it is necessary to set up and solve the differential equation for the circuit. This equation will have the same form as Eq. (7) with E replaced by $E_M \sin \omega t$; that is,

$$L\frac{di}{dt} + Ri = E_M \sin \omega t \tag{10}$$

The solution is

$$i = \frac{E_M}{\sqrt{R^2 + \omega^2 L^2}} \sin (\omega t - \phi) + c\epsilon^{-\frac{Rt}{L}} \tag{11}$$

where $\tan \phi = \omega L/R$, and c is a constant to be determined. The first term is the only one of importance after the current has been flowing for a short time. Thus the current has a peak or maximum amplitude of $E_M/\sqrt{R^2 + \omega^2 L^2}$ and lags the impressed e.m.f. by the phase angle ϕ whose tangent is $\omega L/R$. The quantity $\sqrt{R^2+\omega^2L^2}$ is called the *impedance* of the circuit and is denoted by Z. In terms of the effective values of current and e.m.f. I and E, the equation for the current may be written

$$I = \frac{E}{Z} \text{ or } I_M = \frac{E_M}{Z} \tag{12}$$

In complex notation this form is

$$i = \frac{E_M \sin \omega t}{R + j\omega L} \tag{13}$$

or, in terms of the instantaneous e.m.f.,

$$i = \frac{e}{R + j\omega L} = \frac{e}{z} \tag{14}$$

The quantity z is called the *complex* or *vector impedance.* It is a vector with a magnitude $\sqrt{R^2 + \omega^2 L^2}$ or Z, and an angle ϕ whose tangent is $\omega L/R$. A vector diagram showing these relations is given in Fig. 5. Thus the relation between current and e.m.f. in an a-c circuit containing resistance and inductance in series may be expressed in the same form as Ohm's law for d-c circuits, provided instantaneous values of current and voltage and vector impedance are used [Eq. (14)]. A similar relation may be written using effective values of current and voltage and the magnitude of the vector impedance. Both the vector impedance z and its magnitude Z

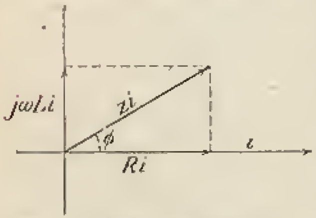


FIG. 5.—Vector relations of inductive circuit.

are generally referred to simply as impedance, the context usually indicating which quantity is meant.

The impedance Z increases as the frequency is increased. Consequently, for constant values of E, R, and L, the current I will decrease

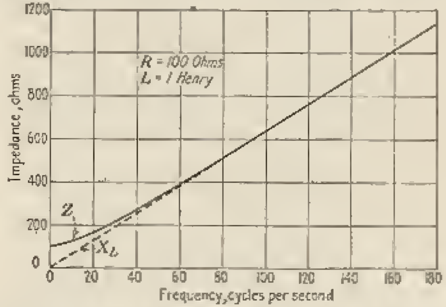


FIG. 6.—Impedance of inductive circuit with frequency.

as the frequency increases. Figure 6 shows values of Z plotted against frequency, and Fig. 7 shows how the current in the circuit of Fig. 3 varies with the frequency of the impressed voltage.

Consider Eq. (11). After the switch has been closed for some time, the values of current and voltage bear a definite relation to each other at each instant during a cycle, and this series of relations is repeated during every cycle. The circuit is now said to be in the steady-state condition, and the first term of the right-hand side of Eq. (11) completely defines the current in terms of the voltage and impedance. However, for a short interval of time after the switch is closed, the second or transient term generally has an appreciable value and must be considered. By comparison with Eq. (9) it is seen that this transient current has the form shown in Fig. 2. It is evident that the duration of the transient current will depend upon the time constant L/R. The initial value of the current, which is equal to the constant c, must, however, be determined. Now the current must be zero at the instant the switch is closed (since it cannot rise to some finite value instantaneously because of the inductance in the circuit) and, therefore, if t is taken as zero at the instant of closing the switch, the value of c may be found mathematically to be defined by the equation

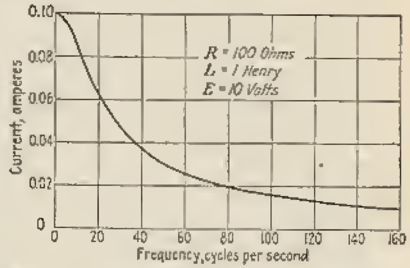


FIG. 7.—Current vs. frequency in inductive circuit.

$$c = \frac{E_M}{Z} \sin \phi = I_M \sin \phi \tag{15}$$

The physical significance of this equation is most readily seen by reference to Fig. 8.[1] In *a* of this figure, the curve *e* represents the voltage impressed upon the circuit and the curve marked "Steady-state current" indicates the value the current would have if the switch had been closed at a time much earlier than the time represented in the figure. Accordingly, at the instant of closing the switch, the current should have the value given by the intersection of the steady-state current curve with the vertical axis in the figure. But the actual current must be zero at this instant; therefore, the transient current must have the value *c*, just neutralizing the fictitious steady-state current. This transient current then decreases according to the curve labeled "transient current," and the actual current is the sum of the steady-state current and the transient current. If the switch should be closed at an instant at which the steady-state current would be zero, as in Fig. 8*b*, the constant *c* would be equal to zero and there would be no transient term. Consequently the quantity ϕ in Eq. (15) represents the phase angle of the instant of closing the switch with reference to the nearest time at which the steady-state current crosses the zero axis in passing from negative to positive values. In Fig. 8*a*, the switch was assumed to be closed shortly after the steady-state current passed through such a zero value; therefore in this case, the so-called "phase angle" is a lag angle, and sin ϕ is negative, making *c* negative as shown.

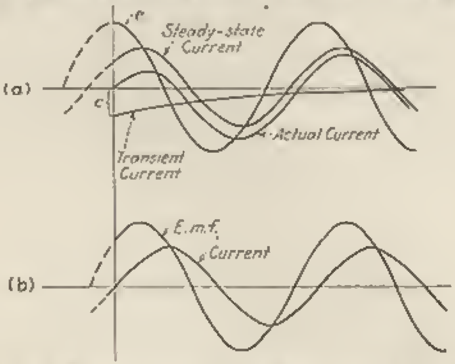


FIG. 8.—Effect on transient current of closing circuit at different times in the cycle.

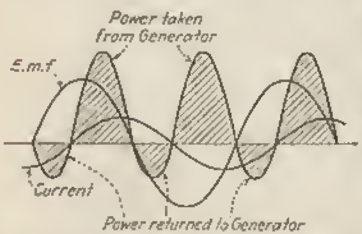


FIG. 9.—Power in inductive circuit.

5. Power in Inductive Circuit. The instantaneous power used in the circuit of Fig. 3 is the product of the instantaneous values of current and voltage. Figure 9[2] shows this power at times to be negative because the current and voltage have opposite signs. Such negative power represents a restoration to the source of some of the energy stored in the magnetic field. In a circuit containing inductance only, the current and voltage are 90 deg. out of phase and the negative loops of the instantaneous-power curve are exactly equal to the positive loops, so that the average power taken by the inductance is zero.

[1] MORECROFT, J. H., "Principles of Radio Communication," 2d ed., 1927.
[2] *Ibid.*

In general, the instantaneous power is given by[1]

$$\begin{aligned} p &= E_M \sin \omega t \times I_M \sin (\omega t - \phi) \\ &= E_M I_M(\sin^2 \omega t \cos \phi - \sin \omega t \cos \omega t \sin \phi) \\ &= \frac{E_M I_M}{2} (\cos \phi - \cos 2\omega t \cos \phi - \sin 2\omega t \sin \phi) \qquad (16) \end{aligned}$$

The average value of the second and third terms in the last parenthesis is zero, so that the average power taken by the circuit is that expressed by the first term, or

$$P = \frac{E_M I_M}{2} \cos \phi = EI \cos \phi \qquad (17)$$

where, as before, E_M and I_M are maximum values, and E and I are effective values of the voltage and current. Since

$$E = IZ$$

and

$$\cos \phi = \frac{R}{Z}$$

$$P = IZ \times I \times \frac{R}{Z} = I^2R \qquad (18)$$

This last equation is often used to define the effective resistance of an a-c circuit.

As a consequence of Eq. (17), the power in an a-c circuit containing inductance and resistance cannot be determined by measuring the current and voltage unless the value of the phase angle ϕ can also be measured. As this is usually difficult, the power must generally be measured with a wattmeter.

The quantity cos ϕ is called the *power factor* of the circuit. In a circuit containing only resistance, the power factor is unity; in a circuit containing only inductance, the power factor would be zero. As applied to a coil used as an inductor, the power factor at a given frequency gives the ratio of the resistance of the coil to its impedance and may be used as a figure of merit for the coil. As the ideal inductor would have zero power factor, a good coil should have a very small power factor.

6. Measurements of Inductance at Low Frequencies. The measurement of the inductance of air-core coils at low frequencies is relatively simple, as the inductance is sensibly constant with change in frequency and current. Iron-core inductors, for reasons which will be examined in detail later, do not have a fixed inductance under all conditions, and measurements on them must be made under conditions which duplicate as nearly as possible the conditions under which the inductor is used.

[1] *Ibid.*

A simple method of approximate measurement uses the circuit of Fig. 10. An a-c voltage of known frequency is applied at E, and the current and voltage read on the meters. The voltmeter reading divided by the ammeter reading gives the impedance and, if the resistance is measured by a d-c-bridge or voltmeter-ammeter method,

$$L = \sqrt{\frac{z^2 - R^2}{4\pi^2 f^2}} = \frac{0.159\sqrt{z^2 - R^2}}{f} \qquad (19$$

The method is usable for iron-core coils that carry a.c. only, provided the measuring current is adjusted to the value that the coil carries in use.

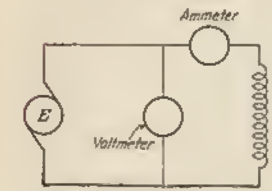


FIG. 10.—Circuit for measurement of inductance.

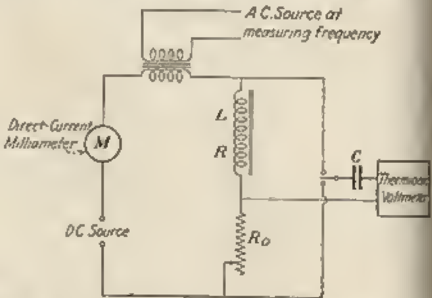


FIG. 11.—Measurement of iron core coil carrying a.c. and d.c.

If measurements are made at a number of current values, the curve of inductance against current may be plotted. The results obtained by this method are generally slightly larger than the true values of inductance because the a-c resistance, particularly in iron-core coils, is greater than the d-c resistance.

7. Measurement of Inductance of Iron-core Coils. When an iron-core coil must carry relatively large d.c. upon which is superimposed a small value of a.c., its inductance is dependent upon the magnitudes of the two currents flowing through it, and other methods must be used.

The impedance of an iron-core coil carrying d.c. and a.c. may be measured by the circuit of Fig. 11. The d.c. through the circuit is adjusted to the value carried by the coil during operation, and the a-c source adjusted to impress a voltage across the coil (measured by the thermionic voltmeter) equal to the a-c voltage across it under operating conditions. The resistance R_0 is then varied until the alternating voltage across it is equal to that across the coil, as measured by the thermionic voltmeter. Then the impedance of the coil at the measuring frequency is equal to R_0. Readjustments of the impressed direct and alternating voltages may be necessary as R_0 is changed. The condenser C prevents the direct voltages across the coil and resistor from affecting the thermionic voltmeter. From the impedance and the resistance of the coil the inductance may be calculated by Eq. (19).

In Fig. 12 is a simple method of arriving at the impedance of an iron-core coil based on the supposition that the inductance is high compared

to the resistance. The voltage across R and X is measured with a vacuum-tube voltmeter, for example. Then $E_r/R = I$ and $E_x/I = X = (E_x/E_r) \times R$, whence

$$X = R/E_r \qquad (20)$$

In the general case in which M represents the total losses of the coil, the power factor of the inductance is $\cos\theta$ and

$$\cos\theta = \frac{E_{2-3} - E_x - E_r}{2E_xE_r} \qquad (21)$$

and the total losses in the core and winding may be thus obtained.

Once the impedance, reactance, and inductance of a coil have been determined, the permeability and finally the magnetizing force and flux density of an iron-core coil may be obtained. Thus the a-c flux density

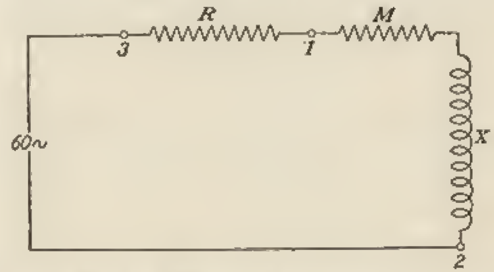


FIG. 12.—Circuit for determining inductance of iron-core coil.

$$B_{\max.} = \frac{E_{\text{eff.}} \times 10^8}{4.44 \times f \times N \times A \times K} \text{ gausses} \qquad (21a)$$

where $E_{\text{eff.}}$ = r.m.s. voltage across the coil
f = frequency in cycles per second
N = number of turns in the winding
A = cross section of the core in square centimeters
K = core-stacking factor (see Sec. 2, Art. 29).

The polarizing m.m.f. resulting from the d.c. in the winding, in gilberts per centimeter is given by

$$H_0 = \frac{1.256NI}{l} \qquad (21b)$$

where N = number of turns in the winding
I = d.c. in amperes
l = length of magnetic circuit in centimeters.

To get m.m.f. in ampere-turns per inch, multiply H_0 by 2.032.

The following table (Allegheny Steel Company) gives values of $B_{\max.}$ and H_0 found in practice.

Coil	B_{max}, gausses	H_0, gilberts/cm
Detector-stage audio transformer	0.5 to 10	0.6 to 1.2
Second-stage a-f transformer	250	1.5
Push-pull output transformer with two primaries	7,000	0
Polarized output transformer	4,200	6.7
Heavy-duty filter reactor (80 ma)	300	27

8. Turner Constant-impedance Method. For measurements involving a.c. only, the constant-impedance method (of Turner[1]), shown in Fig. 13, is used. The method is based upon the fact that, when $1 - \omega^2 LC = 0$, the impedance of the parallel circuit is equal to ωC and is independent of the resistance in the inductive branch. Consequently the line current will have the same magnitude with the switch open or closed. To measure any value of inductance, then, it is only necessary to adjust the capacity so that the reading of the ammeter A is the same for both positions of the switch. Then

$$L = 1/(2\omega^2 C) \tag{22}$$

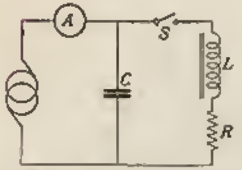


FIG. 13.—Turner constant-impedance method.

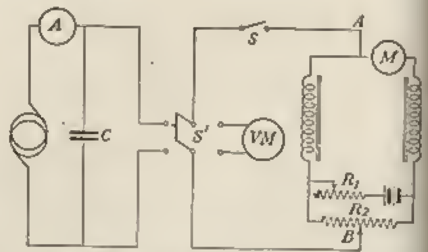


FIG. 14.—Measuring circuit for coils carrying a.c. and d.c.

When the coil must carry d.c. as well as a.c., the circuit of Fig. 14 may be used for the inductance measurement. Two similar inductors are used, the d.c. through them being adjusted to the proper value by means of the resistor R_1 and measured by means of the d-c ammeter M. The switch S' is then thrown to the right and the resistor R_2 adjusted to make the constant-potential difference between the points A and B zero. Then, with S' thrown to the left, the inductance measurement may be carried out in the manner already described. The result is the inductance of the two coils in parallel, which is one-half the inductance of one coil.

9. Measurements of Inductance at High Frequencies. Very often the low-frequency inductance of a coil, determined by one of the methods already given, may also be used as the high-frequency inductance. In some instances it is desirable to determine the inductance at the operating frequency. Bridge methods are not suitable for measurements at high frequencies. Two other methods are commonly used: comparison of the coil with a standard, and measurement of the capacity required to tune the coil to resonance with a known frequency, from which the inductance may be calculated. Both methods give the apparent inductance.

[1] TURNER, H. M., Constant Impedance Method for Measuring Inductance of Choke Coils, *Proc. I.R.E.*, **16**, 1559, 1928.

In the comparison method, a standard inductor, having an apparent inductance L_s at the measuring frequency, is connected in parallel with a calibrated variable condenser, coupled to an oscillator and the coil-condenser circuit tuned to resonance, the capacity C_s of the condenser being noted at the resonance setting. The coil to be measured, whose inductance is denoted by L_x, is then substituted for L_s, the circuit retuned, and the condenser capacity C_x again observed. Since the frequency is the same in both cases,

$$L_x C_x = L_s C_s \tag{23}$$

If the low-frequency inductance L_0 and internal capacity C_0 of the standard coil are known,

$$L_x C_x = L_0(C_s + C_0) \tag{24}$$

In the second method, it is necessary to determine accurately the frequency of the source. The coil to be measured is connected to a calibrated variable condenser, coupled loosely to the generator and tuned to resonance. If f is the frequency of the source, L_x the apparent inductance of the coil, and C_x the condenser capacity at resonance,

$$L_x = \frac{1}{39.48 f^2 C_x} = \frac{0.02533}{f^2 C_x} \tag{25}$$

In this equation, L_x is expressed in henrys and C_x in farads. For L_x in μh and C_x in $\mu\mu$f, the equation becomes

$$L_x = \frac{25.33 \times 10^{15}}{f^2 C_x} \tag{26}$$

If the capacity necessary to tune the coil to resonance at a number of different frequencies is determined, a graph of the squares of the wave lengths corresponding to the several measuring frequencies against the measured values of capacity will be a straight line whose slope is the pure inductance and whose intercept with the negative-capacity axis is the internal capacity of the coil. This is illustrated in Fig. 15.

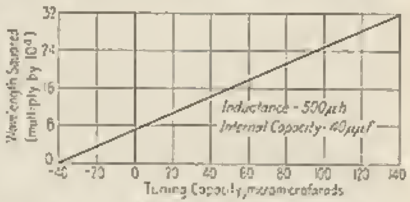


FIG. 15.—Method of determining inductance and distributed capacity of a coil.

10. Inductance of Iron-core Coils. Iron-core coils are mainly useful at relatively low frequencies, and their use is generally confined to circuits carrying currents within the a-f range. (But see Art. 16.)

The inductance of a circuit is not constant if any material of variable permeability is within the magnetic field of the circuit. Consequently, when a coil is wound on an iron core, its inductance is dependent upon the circumstances under which it is used. Accordingly, to use iron-core coils most advantageously, it is necessary to study their characteristics under varying conditions. Three important cases must be distinguished: the current through the coil is a.c. of single frequency; the current consists of a d-c component upon which is superimposed a single-frequency a-c component; the current is comprised of two a-c components of different frequencies.

The average inductance of an iron-core coil carrying a.c. of single frequency is dependent upon the magnitude of the current. Also, the a-c resistance of such a coil is higher than that of an air core coil with an identical winding. Therefore all inductance measurements of iron-core coils should be made with the measuring current equal to the current which will flow through the coil in operation, or the inductance may be measured for a number of different currents and a curve of inductance against current plotted.

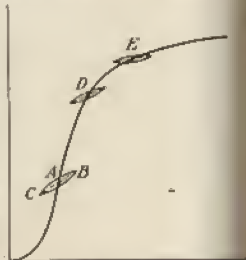


Fig. 16.—Characteristics of coil carrying large value of d.c. and small value of a.c.

In many radio applications a coil carries a relatively large d.c. with a small a-c component superimposed. The inductance of an iron-core coil under such conditions is a function of the magnitudes of the d-c and a-c components of the current. This is illustrated by Fig. 16. The constant magnetizing force (due to the d.c.) may be such as to cause the core to be magnetized to the point A. The alternating component of the magnetizing force (due to the a.c.) will then carry the iron through the small hysteresis loop CB whose slope is not the same as the slope of the magnetization curve. The permeability represented by the slope of this small hysteresis loop is called the *incremental permeability*. As the constant component of the magnetizing force or current is increased, the point A moves farther up the magnetization curve and the incremental permeability decreases, as indicated by the small loops at D and E. As saturation of the core is approached, the incremental permeability, and hence the inductance, becomes very small. As the magnitude of the a-c component is increased, the slope of the hysteresis loop, and accordingly the incremental permeability, increases, thus increasing the inductance. Consequently the inductance of an iron-core coil under these conditions decreases with increase of the d-c component of the current, and increases with increase of the a-c component. Figure 17 shows the decrease in inductance with increase in constant magnetizing force.

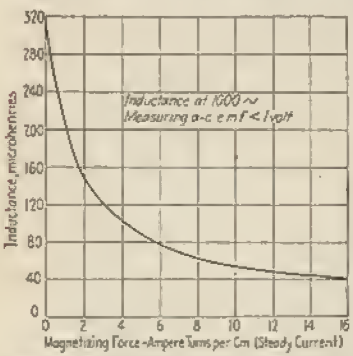


Fig. 17.—Effect of d.c. on inductance of coil.

If an *air gap* is introduced in the magnetic circuit of an iron-core coil, the inductance of the coil is generally diminished. If, however, the coil is carrying both d.c. and a.c., the air gap may so decrease the constant flux that the incremental permeability is actually increased, so that the effective inductance for the a-c component is increased. The effective resistance of the inductor is also decreased by the introduction of an air gap. These effects are illustrated in Fig. 18.[1] As a consequence of these characteristics, iron-core inductors that are intended for use in circuits where they must carry d.c. as well as a.c. are usually made with an air gap in the magnetic circuit of the core.

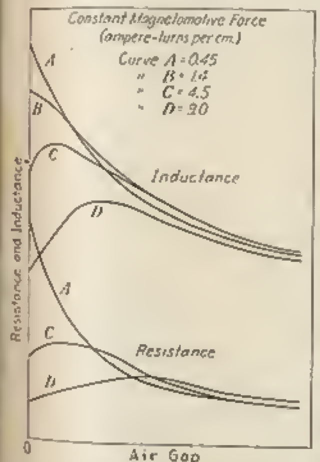


Fig. 18.—Effect of air gap on coil characteristics.

When the inductor carries two alternating currents of different frequencies, the effects of the variable permeability of the iron are somewhat more complicated and of relatively less practical importance than in the cases already treated.[2]

11. Inductors at Radio Frequencies. When inductors are used at radio frequencies, many factors affecting their performance come into prominence. The h-f resistance of a coil is much larger than its d-c resistance because of a number of losses which come into existence with the operation of the coil in h-f circuits. The factors causing this increase are skin effect, eddy currents, dielectric losses, and internal capacity.

When the wire is wound into a coil, the effect of the magnetic field of the coil is such as to concentrate the current on the inner surfaces of the turns. Figure 19 illustrates this effect, the depth of shading indicating the current density. This concentration of current causes a further increase in the effective resistance of the coil, and also causes a decrease in the inductance as the frequency increases. However, the variation of inductance with frequency is generally small in comparison with the variation caused by internal capacity.

Eddy currents in the conductors composing the coil constitute a serious source of loss at frequencies over 3,000 kc. These losses are minimized by the use of wire as small as possible without unduly increasing the conductor resistance, or by the use of tubing instead of wire. Because of these losses at frequencies higher than 3,000 kc there is an optimum wire size giving a minimum resistance in inductance coils.

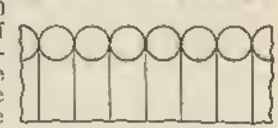


Fig. 19.—Concentration of current at surface at high frequencies.

Any dielectric in the field of the coil also introduces losses which become important at

[1] Morecroft, J. H., "Principles of Radio Communication," 2d ed., 1927.
[2] Turner, H. M., Inductance as Affected by Initial Magnetic State, Air Gap, and Superposed Currents, *Proc. I.R.E.*, 17, 1822, 1929.

these frequencies, so that the type and amount of dielectric within the field of the coil must be carefully regulated. The dielectric should be of the best quality and its volume must be kept at a minimum. The conductor of the coil should, in general, come in contact with the dielectric as little as possible. Coils are often wound upon skeleton or ribbed winding forms so that each turn touches the supporting insulating material at only a few points and is surrounded for the greater part of its length solely by air.

12. Effect of Coil Capacity. Every inductor behaves not as a pure inductance and resistance in series but as an inductance and resistance shunted by a small capacity. This behavior is caused by the self- or internal capacity of the coil. The resistance and inductance of the equivalent parallel circuit at any frequency are called the *apparent resistance* and *apparent inductance* of the coil at that frequency. The apparent resistance is given approximately[1] by the equation

$$R_A = \frac{R}{(1 - \omega^2 LC_0)^2} \tag{27}$$

and the apparent inductance by

$$L_A = \frac{L}{1 - \omega^2 LC_0} \tag{28}$$

where R and L are the resistance and inductance the coil would have at the frequency $\omega/2\pi$ if the internal capacity C_0 were absent. These equations do not hold for frequencies near the natural frequency of the coil; that is, the frequency for which $1 - \omega^2 LC_0 = 0$. These equations are derived on the assumption that the e.m.f. in the circuit is introduced in some manner other than by induction in the coil itself. If the e.m.f. is induced in the coil, the internal capacity is merely added to any other capacity which may be connected in parallel with the coil. Since a coil is practically always used at frequencies for which $1 - \omega^2 LC_0$ is positive, the apparent resistance and inductance of the coil will increase as the frequency increases, the apparent resistance becoming very large as $1 - \omega^2 LC_0$ approaches zero. The percentage change in resistance for a given change in frequency is about twice as great as the change in inductance. At frequencies for which $1 - \omega^2 LC_0$ is negative, the coil behaves as a capacity rather than an inductance.

It has been found[2] that the internal capacity of a single-layer coil is roughly proportional to the radius and practically independent of the number of turns and the length. For a closely wound solenoid, the internal capacity in $\mu\mu$f is very approximately equal to six-tenths of the radius in centimeters.

13. Types of Inductors. A straight wire has a certain amount of inductance, but to make inductors small enough to be convenient it is necessary to wind the wire in the form of a coil thus utilizing a great length of wire in a small space and also increasing the interlinkages of flux and wire.

The simplest inductor consists of a single square turn of wire. The inductance of this arrangement may be calculated accurately, but it has

[1] Radio Instruments and Measurements, *Bur. Standards Circ.* 74.

[2] Howe, G. W. O., *Jour. I.E.E.* (London), **60**, 63, 1922; also Moullin, E. B., "Radio Frequency Measurements," p. 340, 1931.

few other advantages. This type is sometimes used as a fundamental standard.

The *single-layer solenoid* consists of one layer of wire on a cylindrical form, the turns either adjacent to one another or spaced. Sometimes the coil is made self-supporting by means of a binder, such as collodion, and the form removed after winding.

Multilayer coils must be used when a single-layer coil of the required inductance would be inconveniently large. The multilayer coil may take one of three forms: *layer wound, bank wound,* and *honeycomb or duolateral.*

The *layer-wound coil* is useful only at low frequencies because of its high internal capacity caused by the proximity of turns of greatly differing potentials. The wire is wound on the coil in layers, each layer being completed before another is begun. Iron-core coils are usually wound in this manner. If a very large number of turns must be used, it is better for the whole coil to be made up of a number of "pies," each pie being a short layer-wound coil. The pies are assembled side by side to form the complete coil. Insulation is greatly facilitated by this type of construction, and the internal capacity is somewhat reduced.

Fig. 20.—Bank winding.

Bank winding is one result of the attempt to devise a multilayer coil with relatively low internal capacity. The turns are wound in the order shown by the cross-sectional view in Fig. 20.

Honeycomb and *duolateral* windings are further results of the same effort. The wire zigzags back and forth from one side of the winding space to the other, adjacent turns of the same layer being spaced from each other by several times the wire diameter. The effect of this type of winding is to cause turns of adjacent layers to cross each other at an angle and to separate parallel turns by at least the diameter of the wire. A coil of this type is self-supporting and quite compact.

Basket-weave and *spider-web* windings were developed also to minimize the internal capacity. In the basket-weave coil the wire is wound in and out of a number of pegs set in a circle. Adjacent turns cross at an angle. The pegs are usually removed after the winding is completed and the coil is self-supporting. This is essentially a single-layer coil. The spider web, on the other hand, is primarily a multilayer coil of one turn per layer. The wire is wound back and forth between a series of pegs fastened radially in a circular form. This coil may also be self-supporting.

The *toroidal* coil is wound around a doughnut-shaped form. Its field is almost entirely internal, so that it may be placed close to other coils and apparatus.

The *flat spiral* type of coil is self-explanatory—the wire being wound in the form of a spiral, each turn having a greater radius than the preceding one.

14. Variable Inductors. Any of the previous types of coils may be tapped and the number of turns in circuit varied with a tap switch or clip. This method gives only a step-by-step variation, and considerable loss may be introduced by the unused portions of the coil.

A continuously variable inductor may be made by connecting in series or parallel two coils having a variable mutual inductance. The coils may be single-layer or multilayer solenoids and their mutual inductance may be varied by changing the distance between the coils or by rotating one with respect to the other. The most common form of variable inductor, however, is the arrangement commonly called a *variometer*, a cross section of which is

Fig. 21. Variable inductor.

shown in Fig. 21. The inner coil is rotatable about the axis A, which i perpendicular to the plane of the figure. The two coils may be connecte in either series or parallel, thus increasing the range of the instrumen considerably. The mutual inductance between the coils may be increas by winding the outer coil upon the interior of a spherical surface, instea of using the cylindrical form shown.

If a slight increase of resistance of a coil is not objectionable, and th desired range of inductance variation is small, a copper disk slightl smaller than the inside of the coil form may be mounted on a shaf perpendicular to the axis of the coil. The inductance of the coil will l appreciably decreased when the plane of the disk is perpendicul to the coil axis, the decrease of inductance becoming less as the disk rotated away from this position.

15. Design of Inductance Coils. It is desirable that the inductan should be as large as possible, while the resistance is kept at a minimu There are some cases in which a relatively high resistance is permissib or even desirable. Choke coils for use at high frequencies must have high impedance with a minimum internal capacity.

To determine a basis for comparison between coils of different cha acteristics, a factor of merit for an inductor must be defined. Coi for use at frequencies above 300 or 400 kc are usually small in size, that volume is relatively unimportant and the desirable characteristi are high inductance (and, therefore, high reactance) and low resistan The ratio of inductance (or reactance) to resistance may then be take as a factor of merit, the ideal coil having a large ratio. Sometimes th power factor of the coil, which is equal to the ratio of resistance impedance, is taken as a factor of merit, an ideal coil having zero pow factor. The ratio of reactance to resistance ($L\omega/R$) is sometimes calle the Q of the coil. (See Table I, Sec. 6.)

A coil to be used at frequencies below 300 kc is likely to be somewh large if wound in a manner that would be entirely appropriate at high frequencies. Consequently the factor of merit for coils designed for us at the lower radio frequencies should include the volume of the induct and may be defined as the inductance-resistance ratio divided by th volume of the coil.

For a given length of wire, maximum inductance is obtained when th wire is wound as compactly as possible; that is, in a bank-wound coi with a winding cross section as nearly square as possible. The bank wound type is mentioned because the simple multilayer coil is practical useless at radio frequencies because of its high internal capacity. A closely wound single-layer coil made up of the same length of wire has considerably lower inductance than the bank-wound coil. However, a radio frequencies, the resistance of the single-layer coil is so much low than that of the multilayer coil that the L/R ratio of the former is muc larger than that of the latter. In view of its simplicity of constructio the single-layer solenoid wound with solid wire would appear to be th most desirable coil type at medium and high radio frequencies, eve though within certain ranges of frequency some other types have certai advantages. At high frequencies (above 3,000 kc), the single-lay solenoid, either closely wound or spaced, is used almost exclusively.

For a given wire length, this type of coil has a maximum inductan when the ratio of diameter to length of coil is 2.46,[1] although this val

[1] Radio Instruments and Measurements, *Bur. Standards Circ.* 74.

is not critical. The inductance decreases somewhat rapidly as this ratio becomes much smaller than 2.46, while the decrease is only slight for larger values of the ratio. Since the internal capacity of the coil is approximately proportional to the diameter, it is advantageous to use a ratio of diameter to length somewhat smaller than 2.46, provided that the coil is to be used under such conditions that the decrease in internal capacity effected in this way more than compensates for the slightly lower inductance-resistance ratio.

A multilayer coil has a maximum inductance when the cross section of the winding is a square. It has also been shown[1] that, with a square cross section given, the inductance of this type of coil is maximum when the mean diameter is 3.02 times the depth of the winding.

Below 300 kc the volume of the coil must be included in the factor of merit. In these circumstances, the honeycomb and bank-wound coils outstrip all others, the honeycomb type being somewhat superior to the bank wound. Table I gives the characteristics of honeycomb coils.

TABLE I.—HONEYCOMB-COIL DATA

Turns on coil	Size of wire, B. & S. gage	Inductance, mh	Distributed capacity, $\mu\mu f$	Natural wave length, meters	Wave lengths with the following shunt-condenser capacities, μf			
					0.001	0.0005	0.00025	0.0001
25	24	0.038	26.8	60	372	267	193	131
35	24	0.076	30.8	91	528	378	277	188
50	24	0.150	36.4	139	743	534	391	270
75	24	0.315	28.6	179	1,007	770	560	379
100	24	0.585	36.1	274	1,470	1,055	771	532
150	24	1.29	21.3	313	2,160	1,546	1,110	746
200	25	2.27	18.9	391	2,870	2,050	1,470	980
250	25	4.20	22.9	585	3,910	2,800	2,020	1,355
300	25	6.60	19.0	669	4,900	3,490	2,510	1,670
400	25	10.5	17.4	806	6,160	4,400	3,100	2,095
500	25	18.0	17.3	1,052	8,070	5,750	4,140	2,740
600	28	37.5	19.2	1,600	11,600	8,300	5,980	3,980
750	28	49.0	18.3	1,785	13,300	9,500	6,830	4,540
1,000	28	85.3	16.8	2,260	17,600	12,500	9,000	5,950
1,250	28	112.0	15.5	2,490	20,100	14,300	10,250	6,780
1,500	28	161.5	15.8	3,000	24,200	17,200	12,350	8,150

16. Coils for Various Frequency Ranges. A study of the characteristics of various types of inductors in the frequency range of 300 to 1,500 kc has been made by Hund and De Groot.[2] Their results show that in this frequency band the single-layer solenoid and the loose basket-weave coils have the highest inductance-resistance ratios of the coils wound with solid wire, with the radial basket weave or spider web a close third. Coils wound with 32–38 Litz wire were found to be somewhat better in all respects than solid-wire coils. Contrary to a somewhat generally accepted belief, a few broken strands in the Litz wire made only a slight difference in the r-f resistance of a coil.

[1] Radio Instruments and Measurements, *Bur. Standards Circ.* 74.
[2] HUND, AUGUST, and H. B. DE GROOT, Radio Frequency Resistance and Inductance of Coils Used in Broadcast Reception, *Bur. Standards Tech. Paper* 298, Vol. 19, p. 651, 1925.

In solid-wire coils, little is gained by using a wire size larger than No. 24-AWG, although No. 16 gives a slightly lower resistance between 30 and 1,200 kc. Spacing the turns does not decrease the resistance appre ciably—not enough to compensate for the extra length necessary. A number of binders were tried on single-layer coils, all of them causing a slight increase in the r-f resistance of the coil. Collodion appeared to be the best of these binders.

At frequencies above 3,000 kc, dielectric losses, eddy currents, and internal capacity are important. The first two cause relatively larg increases in the coil resistance. The third increases both the resistanc

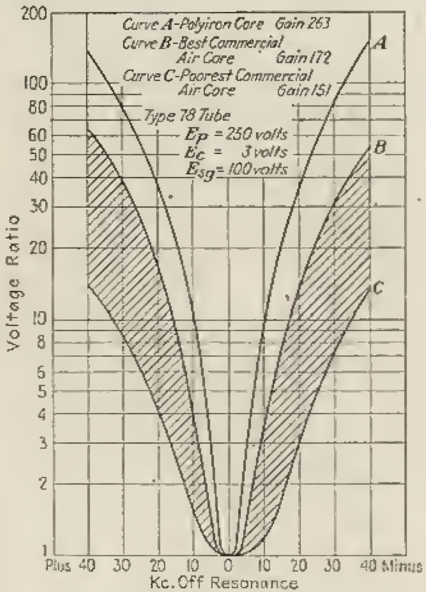


FIG. 22.—Iron-core coil characteristics.

and inductance of the coil if the voltage in the circuit is not induced in the coil itself. If the circuit e.m.f. is introduced by induction in the coil, the internal capacity, acting as a parallel condenser, determines the highest frequency to which the coil can be tuned. As the upper limit of parallel tuning capacity is not very large (in order that the L/C ratio be not too small), a large internal capacity seriously restricts the range over which the coil may be tuned efficiently. It is for these reasons that the single-layer solenoid is used almost exclusively at such frequencies.

Coils for Short-wave Receivers. A considerable study of coils of various sizes made from wire of various sizes and for use at frequencies of the order of 15 Mc was made by W. S. Barden and David Grimes.

[1] *Electronics*, June. 1934, p. 174. (This material and that on iron-core inductances are by the Editor.)

It was determined that maximum value of Q for such coils, of the order of 1 μh inductance, could be realized when wire diameter and spacing between turns were of the same order of magnitude. Very large wire (long coils) was not superior to medium-size wire, say No. 20 or No. 22. Using wire of No. 14 size, 1-in.-diameter coils were superior to ½-in.-diameter coils for any winding length.

It was determined that shielding the coil does not reduce the Q to a serious extent, provided proper spacing is observed. In reasonable practice Q need not be decreased by more than 10 per cent or L by more than 15 per cent. Bakelite winding forms have some effect upon Q. Thus a 1-μh coil of No. 10 wire (0.104 in.) was wound on a 2-in. length of 1.5-in.-diameter bakelite having a 0.125-in. wall. This coil had 0.333-in. winding pitch. At 15 mc, $Q = 212$. Upon removing the winding form it remained self-supporting, and Q increased to 229.

Coils made of No. 14 wire on a 1-in.-diameter form with 0.111 in. between turns (0.88 μh, 5¾ turns) were found to be good compromise coils. These would have a Q of 184. Coils made on 0.5-in. forms wound with small wire, say No. 24, have values of Q in the region from 75 to 100.

Iron-core R-f Inductances. From 1931 to 1935 considerable headway was made in the use of ferro inductors at broadcast and intermediate frequencies. The advantages offered by iron coils over air coils are the small size and high Q. They have been especially useful where it is necessary to get high gain, or high selectivity, in small space, or with a minimum number of tuned circuits. Some attempt has been made to use coils with variable iron cores so that in tuning a circuit the inductance would be varied instead of the capacity.

One such material (Polyiron) has an iron content of 95 per cent. The remainder of the pressed core is bakelite and insulating varnish. Permeability measured with toroidal cores is of the order of 12; its specific gravity is 4.8 against 7.0 for solid iron; its conductivity is 100 mhos per cubic centimeter against 10^{-5} for solid iron. Permeability remains constant from 50 to 2,000,000 cycles. Variation of magnetic force from 0.01 to 10 gauss makes no appreciable change.[1]

Another iron which has come into use in this country is Ferrocart, already widely used in Europe. Intermediate-frequency transformers for 456, 370, 360 and 175 kc have been designed from Ferrocart and Polyiron as have transformers coupling an i-f stage to a diode detector. For automobile and other receivers where high initial gain is required, to reduce the noise to signal ratio, iron coils seem to offer considerable advantages.

In a typical receiver of the characteristics given below, the table shows the advantages to be gained by using iron instead of air-core coils.

This receiver was a six-tube a-c export tube, employing 370-kc i-f transformers. It used a type 57 first detector, type 27 oscillator, a type 58 i-f amplifier, a type 2A6 diode-triode, a type 2A5 output tube and a type 80 rectifier. The high impedance of the plate-cathode circuit of the first detector is partially responsible for the excellent selectivity of the receiver.

[1] LANGLEY, RALPH H., Tuning by Permeability Variation, *Electronics*, July, 1931; CROSSLEY, ALFRED, Iron Core Intermediate Frequency Transformers, *Electronics*, November, 1933; POLYDOROFF, W. J., Further Notes on Iron-core Coils, *Electronics*, January, 1934; and Ferro-inductors and Permeability Tuning, *Proc. I.R.E.*, May 1933; FILL, J. V., Ferrocart and Its Applications, *Electronics*, November, 1934.

Care was taken to align the receiver properly at each frequency i
order that each test be made under the best conditions.

WITH AIR-CORE TRANSFORMERS

Frequency, kilocycles	Band width 10 times, kilocycles	Band width 100 times, kilocycles	Band width 1,000 times, kilocycles	Sensitivity, microvolts
1,400	18	37	62	5
1,000	13	28	46	4
600	13	26	42	5
WITH IRON-CORE UNITS				
1,400	7	16	31	5
1,000	7	15	27	4
600	7	14	26	6

The advantages from the standpoint of gain are as follows.

In a five-tube a-c d-c set of the better type employing 456-kc i-f tran-formers, the tube complement was as follows: 6C6, 6D6, 75, 43, an 25Z5. The type 6C6 was employed as a composite oscillator-fir detector. In this receiver the two i-f transformers and also the antenn coupler were replaced with iron-core units. The sensitivity at 1000 k increased from 100 to 20 μv.

17. Calculation of Inductance of Air-core Coils. The inductance o many types of air-core coils may be calculated by means of formula involving the dimensions of the coil and the number of turns.[1] Sever formulas from *Circular* 74 of the Bureau of Standards are given her Few of the available corrections to inductance formulas are include since they apply only to the calculation of the l-f inductance. The h-inductance of a coil cannot be calculated with a high degree of accurac because of the skin effect and coil capacity.

In the following formulas all dimensions are expressed in centimeters an the inductance is in microhenrys.

18. Straight Round Wire. If l is the length of the wire, d is the diamete of the cross section, and μ is the permeability of the material of the wire,

$$L_0 = 0.002l\left[\log_\epsilon \frac{4l}{d} - 1 + \frac{\mu}{4}\right] \quad (29)$$

$$= 0.002l\left[2.303 \log_{10} \frac{4l}{d} - 1 + \frac{\mu}{4}\right] \quad (30)$$

If $\mu = 1$ (for all materials except iron),

$$L_0 = 0.002l\left[2.303 \log_{10} \frac{4l}{d} - 0.75\right] \quad (31)$$

The return conductor is assumed to be remote. These formulas give the low-frequency inductance.

[1] ROSA, E. B., and F. W. GROVER, *Bur. Standards Sci. Paper* 169; GROVER, F. W., *Bur. Standards Sci. Papers* 320, 1917; 455, 1922; 468, 1923. See for coil design and calculation, especially at low frequencies, MORGAN BROOKS and H. M. TURNER, Inductance of Coils, *Bull.* 53, *Univ. Ill. Eng. Exper. Sta.*, Jan. 8, 1912.

As the frequency increases, the inductance decreases, its value at infinite frequency being

$$L_\infty = 0.002l\left[2.303 \log_{10} \frac{4l}{d} - 1\right] \quad (32)$$

A general expression for the inductance at any frequency is

$$L = 0.002l\left[2.303 \log_{10} \frac{4l}{d} - 1 + \mu\delta\right] \quad (33)$$

The quantity δ is obtained from the table below, as a function of the argument x, where

$$x = 0.1405d\sqrt{\frac{\mu f}{\rho}} \quad (34)$$

and f is the frequency and ρ is the volume resistivity of the wire in microhm-centimeters. For copper at 20°C.,

$$x_c = 0.1071d\sqrt{f}$$

This quantity δ will be used in several of the following formulas without further definition.

VALUE OF δ IN INDUCTANCE FORMULAS

x	δ	x	δ	x	δ	x	δ	x	δ	x	δ
0	0.250	2.5	0.228	6.0	0.116	12.0	0.059	25.0	0.028	70.0	0.010
0.5	0.250	3.0	0.211	7.0	0.100	14.0	0.050	30.0	0.024	80.0	0.009
1.0	0.249	3.5	0.191	8.0	0.088	16.0	0.044	40.0	0.0175	90.0	0.008
1.5	0.247	4.0	0.1715	9.0	0.078	18.0	0.039	50.0	0.014	100.0	0.007
2.0	0.240	5.0	0.139	10.0	0.070	20.0	0.035	60.0	0.012	∞	0.000

19. Two Parallel Round Wires—Return Circuit. The current is assumed to flow in opposite directions in two parallel wires of length l and diameter d, the distance between centers of wires being D. Then

$$L = 0.004l\left[2.303 \log_{10} \frac{2D}{d} - \frac{D}{l} + \mu\delta\right] \quad (35)$$

This neglects the inductance of the wires connecting the two main wires. If these wires are long, their inductance may be calculated by Eq. (33) and added to the result from Eq. (35), or the whole system may be treated as a rectangle and the inductance calculated by Eq. (37).

20. Square of Round Wire. The length of one side of the square is denoted by a; other letters have already been defined.

$$L = 0.008a\left[2.303 \log_{10} \frac{2a}{d} + \frac{d}{2a} - 0.774 + \mu\delta\right] \quad (36)$$

21. Rectangle of Round Wire. The sides of the rectangle are a and a_1 and the diagonal $g = \sqrt{a^2 + a_1^2}$. Then

$$L = 0.00921\left[(a + a_1)\log_{10} \frac{4aa_1}{d} - a\log_{10}(a + g) - a_1\log_{10}(a_1 + g)\right]$$
$$+0.004\left[\mu\delta(a + a_1) + 2\left(g + \frac{d}{2}\right) - 2(a + a_1)\right] \quad (37)$$

22. Grounded Horizontal Wire. The wire is assumed to be parallel to the earth which acts as the return circuit. In addition to symbols already used, h denotes the height of the wire above ground. Then

$$L = 0.004605l\left[\log_{10}\frac{4h}{d} + \log_{10}\left\{\frac{l+\sqrt{l^2+\frac{d^2}{4}}}{l+\sqrt{l^2+4h^2}}\right\}\right] + 0.002\left[\sqrt{l^2+4h^2} - \sqrt{l^2+\frac{d^2}{4}} + \mu l\delta - 2h + \frac{d}{2}\right] \quad (38)$$

23. Circular Ring of Circular Section. If a is the mean radius of the ring,

$$L = 0.01257a\left[2.303\log_{10}\frac{16a}{d} - 2 + \mu\delta\right] \quad (39)$$

provided that $d/2a \leqslant 0.2$.

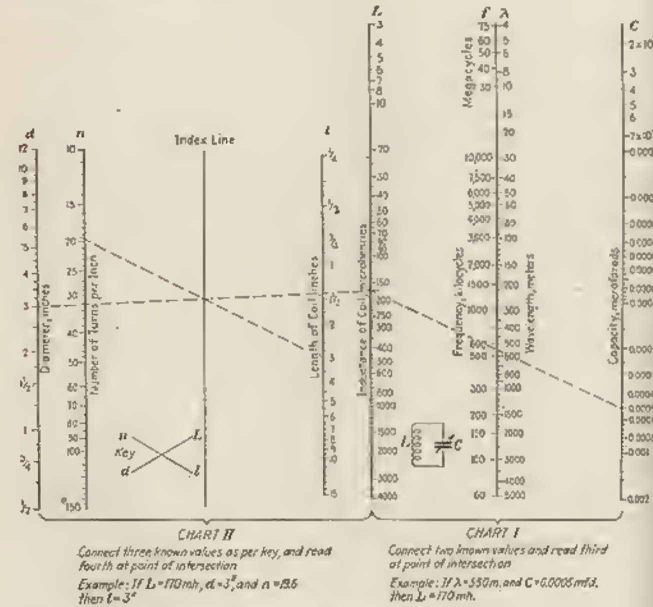


Fig. 23. — Inductance-design chart.

24. Single-layer Coil or Solenoid.

$$L = \frac{0.0395a^2n^2}{b}K \quad (40)$$

where n is the number of turns, a is the radius of the coil measured from the axis to the center of the wire, b is the length of the coil, and K is a function of $2a/b$, the value of which may be determined by means of the table below

Value of K in Formula 40

Diameter to length	K	Difference	Diameter to length	K	Difference	Diameter to length	K	Difference
0.00	1.0000	−0.0209	2.00	0.5255	−0.0118	7.00	0.2584	−0.0047
.05	.9791	203	2.10	.5137	112	7.20	.2537	45
.10	.9588	197	2.20	.5025	107	7.40	.2491	43
.15	.9391	190	2.30	.4918	102	7.60	.2448	42
.20	.9201	185	2.40	.4816	97	7.80	.2406	40
0.25	0.9016	−0.0178	2.50	0.4719	−0.0093	8.00	0.2366	−0.0094
.30	.8838	173	2.60	.4626	89	8.50	.2272	86
.35	.8665	167	2.70	.4537	85	9.00	.2185	79
.40	.8499	162	2.80	.4452	82	9.50	.2106	73
.45	.8337	156	2.90	.4370	78	10.00	.2033	
0.50	0.8181	−0.0150	3.00	0.4292	−0.0075	10.0	0.2033	−0.0133
.55	.8031	146	3.10	.4217	72	11.0	.1903	113
.60	.7885	140	3.20	.4145	70	12.0	.1790	98
.65	.7745	136	3.30	.4075	67	13.0	.1692	87
.70	.7609	131	3.40	.4008	64	14.0	.1605	78
0.75	0.7478	−0.0127	3.50	0.3944	−0.0062	15.0	0.1527	−0.0070
.80	.7351	123	3.60	.3882	60	16.0	.1457	63
.85	.7228	118	3.70	.3822	58	17.0	.1394	58
.90	.7110	115	3.80	.3764	56	18.0	.1336	52
.95	.6995	111	3.90	.3708	54	19.0	.1284	48
1.00	0.6884	−0.0107	4.00	0.3654	−0.0052	20.0	0.1236	−0.0085
1.05	.6777	104	4.10	.3602	51	22.0	.1151	73
1.10	.6673	100	4.20	.3551	49	24.0	.1078	63
1.15	.6573	98	4.30	.3502	47	26.0	.1015	56
1.20	.6475	94	4.40	.3455	46	28.0	.0959	49
1.25	0.6381	−0.0091	4.50	0.3409	−0.0045	30.0	0.0910	−0.0102
1.30	.6290	89	4.60	.3364	43	35.0	.0808	80
1.35	.6201	86	4.70	.3321	42	40.0	.0728	64
1.40	.6115	84	4.80	.3279	41	45.0	.0664	53
1.45	.6031	81	4.90	.3238	40	50.0	.0611	43
1.50	0.5950	−0.0079	5.00	0.3198	−0.0076	60.0	0.0528	−0.0061
1.55	.5871	76	5.20	.3122	72	70.0	.0467	48
1.60	.5795	74	5.40	.3050	69	80.0	.0419	38
1.65	.5721	72	5.60	.2981	65	90.0	.0381	31
1.70	.5649	70	5.80	.2916	62	100.0	.0350	
1.75	0.5579	−0.0068	6.00	0.2854	−0.0059			
1.80	.5511	67	6.20	.2795	56			
1.85	.5444	65	6.40	.2739	54			
1.90	.5370	63	6.60	.2685	52			
1.95	.5316	61	6.80	.2633	49			

25. Multilayer Coils: Circular Coils of Rectangular Cross Section. For long coils of a few layers, the following formula may be used:

$$L = L_s - \frac{0.0126n^2ac}{b}(0.693 + B_s) \quad (41)$$

where L_s is the inductance calculated by Eq. (40), n and b are the same as in Eq. (40), a is the radius of coil measured from axis to center of winding cross section, c is the radial depth of winding, and B_s is the correction given on p. 96.

VALUE OF B_s IN FORMULA 43

b/c	B_s	b/c	B_s	b/c	B_s	b/c	B_s	b/c	B_s	b/c	B_s
1	0.0000	6	0.2446	11	0.2844	16	0.3017	21	0.3116	26	0.318
2	0.1202	7	0.2563	12	0.2888	17	0.3041	22	0.3131	27	0.319
3	0.1753	8	0.2656	13	0.2927	18	0.3062	23	0.3145	28	0.320
4	0.2076	9	0.2730	14	0.2961	19	0.3082	24	0.3157	29	0.320
5	0.2292	10	0.2792	15	0.2991	20	0.3099	25	0.3169	30	0.321

For short multilayer coils, the dimensions shown in Fig. 24 are used. Two formulas are required, one for use when $b > c$, and the other for use when $b < c$. In the first case:

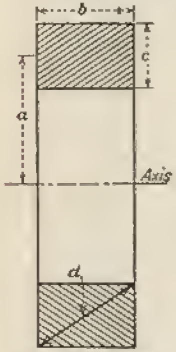


FIG. 24.—Multilayer coil.

$$L = 0.01257an^2\left[\left(1 + \frac{b^2}{32a^2} + \frac{c^2}{96a^2}\right)\log_e \frac{8a}{d} - y_1 + \frac{b^2}{16a^2}y_2\right]$$

$$= 0.01257an^2\left[2.303\left(1 + \frac{b^2}{32a^2} + \frac{c^2}{96a^2}\right)\log_{10} \frac{8a}{d} - y_1 + \frac{b^2}{16a^2}y_2\right] \quad (42)$$

When $b < c$:

$$L = 0.01257an^2\left[\left(1 + \frac{b^2}{32a^2} + \frac{c^2}{96a^2}\right)\log_e \frac{8a}{d} - y_1 + \frac{c^2}{16a^2}y_3\right]$$

$$= 0.01257an^2\left[2.303\left(1 + \frac{b^2}{32a^2} + \frac{c^2}{96a^2}\right)\log_{10} \frac{8a}{d} - y_1 + \frac{c^2}{16a^2}y_3\right] \quad (43)$$

y_1, y_2, and y_3 may be obtained from the table shown below. These formulas are quite accurate as long as the diagonal of the cross section (d Fig. 24) does not exceed the mean radius. The accuracy decreases considerably as b becomes large in comparison with a.

For very accurate results, a correction must be added if the insulation of the wire occupies a considerable percentage of the winding space. This correction is given by

$$\Delta L = 0.01257an\left[2.303 \log_{10} \frac{D}{d} + 0.155\right] \quad (44)$$

where D is the distance between the centers of adjacent wires, and d is the diameter of the bare wire.

26. Multilayer Square Coil. If n is the number of turns and a is the side of the square measured to the center of the rectangular cross section which has length b and depth c, then

$$L = 0.008an^2\left[2.303 \log_{10} \frac{a}{b + c} + 0.2235\frac{b + c}{a} + 0.726\right] \quad (45)$$

If the cross section is square ($b = c$), this becomes

$$L = 0.008an^2\left[2.303 \log_{10} \frac{a}{b} + 0.447\frac{b}{a} + 0.033\right] \quad (46)$$

VALUE OF CONSTANTS IN FORMULAS (42) AND (43)

b/c or c/b	y_1	c/b	y_2	b/c	y_3
0	0.5000	0	0.125	0	0.597
0.025	0.5253				
0.05	0.5490	0.05	0.127	0.05	0.599
0.10	0.5924	0.10	0.132	0.10	0.602
0.15	0.6310	0.15	0.142	0.15	0.608
0.20	0.6652	0.20	0.155	0.20	0.615
0.25	0.6953	0.25	0.171	0.25	0.624
0.30	0.7217	0.30	0.192	0.30	0.633
0.35	0.7447	0.35	0.215	0.35	0.643
0.40	0.7645	0.40	0.242	0.40	0.654
0.45	0.7816	0.45	0.273	0.45	0.665
0.50	0.7960	0.50	0.307	0.50	0.677
0.55	0.8081	0.55	0.344	0.55	0.690
0.60	0.8182	0.60	0.384	0.60	0.702
0.65	0.8265	0.65	0.427	0.65	0.715
0.70	0.8331	0.70	0.474	0.70	0.729
0.75	0.8383	0.75	0.523	0.75	0.742
0.80	0.8422	0.80	0.576	0.80	0.756
0.85	0.8451	0.85	0.632	0.85	0.771
0.90	0.8470	0.90	0.690	0.90	0.786
0.95	0.8480	0.95	0.752	0.95	0.801
1.00	0.8483	1.00	0.816	1.00	0.816

Formula (43) may be used to correct for insulation by replacing the factor 0.01257 by 0.008.

For a single-layer square coil,

$$L = 0.008an^2\left[2.303 \log_{10} \frac{a}{b} + 0.2231\frac{b}{a} + 0.726\right] - 0.008an(A + B) \quad (47)$$

A and B are given below, where d is the diameter of the bare wire and D is the distance between turns, measured to the centers of the wires.

VALUE OF A IN FORMULA (47)

d/D	A	d/D	A	d/D	A
1.00	0.557	0.40	−0.359	0.15	−1.340
0.95	0.506	0.38	−0.411	0.14	−1.409
0.90	0.452	0.36	−0.465	0.13	−1.483
0.85	0.394	0.34	−0.522	0.12	−1.563
0.80	0.334	0.32	−0.583	0.11	−1.650
0.75	0.269	0.30	−0.647	0.10	−1.746
0.70	0.200	0.28	−0.716	0.09	−1.851
0.65	0.126	0.26	−0.790	0.08	−1.969
0.60	0.046	0.24	−0.870	0.07	−2.102
0.55	−0.041	0.22	−0.957	0.06	−2.256
0.50	−0.136	0.20	−1.053	0.05	−2.439
0.48	−0.177	0.19	−1.104	0.04	−2.662
0.46	−0.220	0.18	−1.158	0.03	−2.950
0.44	−0.264	0.17	−1.215	0.02	−3.355
0.42	−0.311	0.16	−1.276	0.01	−4.048

VALUE OF B IN FORMULA (47)

Number of turns, n	B	Number of turns, n	B
1	0.000	40	0.315
2	0.114	45	0.317
3	0.166	50	0.319
4	0.197	60	0.322
5	0.218	70	0.324
6	0.233	80	0.326
7	0.244	90	0.327
8	0.253	100	0.328
9	0.260	150	0.331
10	0.266	200	0.333
15	0.286	300	0.334
20	0.297	400	0.335
25	0.304	500	0.336
30	0.308	700	0.336
35	0.312	1,000	0.336

27. Inductance Standards. Like all other standards, inductance standards must be rugged, permanent, and constant. The simplest fundamental standard is a single square turn of round wire. The inductance of such a standard can be calculated with great accuracy.

When a standard having a large value of inductance is desired, the single square turn becomes too large for use, and it is necessary to design some more compact form. The resistance and internal capacity must be kept to a minimum. Furthermore the turns must be held rigidly in place so they cannot change their relative positions. The dielectric in the field of the coil must have a minimum volume and be of such material that the losses in it are as small as possible.

These requirements are best met by a single-layer solenoid with a spaced winding. For a minimum conductor resistance, the ratio of diameter to length should be 2.46, but a somewhat smaller value of this ratio is desirable to reduce the internal capacity, this being proportional to the radius.

One excellent form of standard inductor is made by winding silk-covered Litz wire in slots in the edges of strips of hard rubber, the ends of which are supported by hard-rubber rings. With this skeleton type of winding form, the cross section of the coil is polygonal rather than circular. In order that the proper ratios of diameter to length may be maintained, the coils must be of large size, their diameters ranging from 10 to 40 cm. for inductance values that are necessary in the frequency range from 15 to 1,500 kc. Such a coil must be given relatively careful handling, however, since jolts might cause some of the wires to change their positions. A more rugged coil consists of bare wire wound upon a threaded cylindrical form, the turns being cemented in place with a very little cement, preferably collodion. The form should be as thin as is consistent with adequate strength. Glass forms may also be used, although it is then necessary to cement the turns more thoroughly than in the case of a threaded form.

With recent advances in the precision of frequency determination and improvement in standard condensers, the temperature coefficient of a

standard inductance may become an important factor. It is possible, in this case, to reduce the temperature coefficient by a special design of the winding form.

28. Mutual Inductance. As the changing magnetic field due to a varying current in a circuit induces an e.m.f. in the circuit itself, so may it induce an e.m.f. in any neighboring circuit. The e.m.f. induced in the first circuit depends upon the self-inductance of that circuit, and, in the same way, the e.m.f. induced in the second circuit depends upon the *mutual inductance* between the two circuits. Mutual inductance is defined in three ways exactly analogous to the three ways of defining self-inductance: (1) as the magnetic flux linking the second circuit when unit current flows in the first circuit; (2) as the e.m.f. induced in circuit 2 when the current in circuit 1 changes at the rate of one unit per second; (3) as twice the work done in establishing the magnetic flux, linking circuit 2, associated with unit current in circuit 1. These three definitions give constant and equal values for the mutual inductance if there is no material of variable permeability near the circuits and if the current does not vary so rapidly that its distribution in the cross section of the conductors differs greatly from a uniform one. The change in current distribution at high frequencies, however, has a very slight effect upon the mutual inductance.

The units of mutual inductance are the same as those of self-inductance: in the practical system they are the henry and its subdivisions, the millihenry (mh) and microhenry (μh).

29. Measurement of Mutual Inductance. When two inductors, having a mutual inductance, are connected in series so that their magnetic fields aid each other, the total inductance of the combination is

$$L' = L_1 + L_2 + 2M \tag{48}$$

where L' is the inductance of the combination, L_1 and L_2 are the inductances of the coils, and M is their mutual inductance. If the connections to one of the coils are reversed, the total inductance becomes

$$L'' = L_1 + L_2 - 2M \tag{49}$$

Then, from these two equations,

$$M = \frac{L' - L''}{4} \tag{50}$$

These relations furnish a convenient method for the measurement of mutual inductance. The inductance of the two coils connected in series is measured by any suitable method, the connections to one coil reversed, and the inductance again measured. The larger of the two measured values is then denoted by L' and the smaller by L'', and M is calculated by means of Eq. (50). This method is applicable at any frequency, provided the inductance-measurement method is appropriate at that frequency. It is not very accurate when M is small in comparison with the inductance of the larger of the two coils.

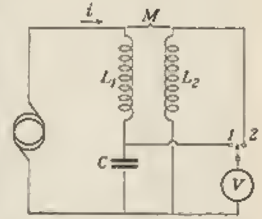


FIG. 25.—Circuit for measuring mutual inductance.

A method applicable for all values of M is illustrated in Fig. 25.[1] V represents a voltage-measuring device of high impedance, preferably a thermionic voltmeter. A voltage source of frequency $\omega/2\pi$ is connected to

[1] MOULLIN, E. B., "Radio Frequency Measurements," p. 383, 1932.

VALUES OF F FOR FORMULA 56

r_2/r_1	F	Difference
0	∞	
0.010	0.05016	−0.00120
.011	4897	109
.012	4787	100
0.013	4687	−0.00093
.014	4594	87
.015	4507	81
.016	4426	148
.018	4278	132
0.020	0.04146	−0.00119
.022	4027	109
.024	3918	100
.026	3818	93
.028	3725	86
0.030	3639	−0.00081
.032	3558	76
.034	3482	71
.036	3411	68
.038	3343	64
0.040	0.03279	−0.00061
.042	3218	58
.044	3160	55
.046	3105	53
.048	3052	51
0.050	0.03001	−0.00226
.060	2775	191
.070	2584	164
.080	2420	144
.090	2276	128
0.100	0.02148	−0.00116
.11	2032	104
.12	1928	96
.13	1832	89
.14	1743	82
0.15	0.01661	−0.00075
.16	1586	71
.17	1515	66
.18	1449	62
.19	1387	59
0.20	0.01328	−0.00055
.21	1273	52
.22	1221	50
.23	1171	47
.24	1124	45
0.25	0.010792	−0.000425
.26	10366	408
.27	0.009958	388
.28	9570	371
.29	9199	355

r_2/r_1	F	Difference
0.30	0.008844	−0.000341
.31	8503	328
.32	8175	314
.33	7861	302
.34	7559	290
0.35	0.007269	−0.000280
.36	6989	270
.37	6720	260
.38	6460	249
.39	6211	241
0.40	0.005970	−0.000232
.41	5738	225
.42	5514	217
.43	5297	210
.44	5087	202
0.45	0.004885	−0.000195
.46	4690	189
.47	4501	183
.48	4318	178
.49	4140	171
0.50	0.003969	−0.000166
.51	3803	160
.52	3643	156
.53	3487	150
.54	3337	146
0.55	0.003191	−0.000141
.56	3050	137
.57	2913	133
.58	2780	128
.59	2652	125
0.60	0.002527	−0.000120
.61	2407	117
.62	2290	113
.63	2177	109
.64	2068	106
0.65	0.001962	−0.000103
.66	1859	99
.67	1760	96
.68	1664	93
.69	1571	90
0.70	0.001481	−0.000087
.71	1394	84
.72	1310	81
.73	1228	78
.74	1150	76
0.75	0.0010741	−0.0000731
.76	10010	704
.77	9306	680
.78	8626	653
.79	7973	628

r_2/r_1	F	Difference
0.80	0.0007345	−0.000060[illegible]
.81	6741	570
.82	6162	555
.83	5607	531
.84	5076	507
0.85	0.0004569	−0.0000481
.86	4085	460
.87	3625	437
.88	3188	413
.89	2775	389
0.90	0.0002386	−0.0000365
.91	2021	341
.92	1680	316
.93	1364	290
.94	1074	263
0.95	0.00008107	−0.00002351
.96	5756	2046
.97	3710	1706
.98	2004	1301
.99	703	703
1.00	0	
0.950	0.00008170	−0.0000049[illegible]
.952	7613	[illegible]
.954	7131	[illegible]
.956	6661	[illegible]
.958	6202	444
0.960	0.00005756	−0.0000043[illegible]
.962	5320	421
.964	4899	409
.966	4490	397
.968	4093	383
0.970	0.00003710	−0.000003[illegible]
.972	3340	[illegible]
.974	2984	341
.976	2643	[illegible]
.978	2316	312
0.980	0.00002004	−0.00000[illegible]
.982	1708	[illegible]
.984	1430	[illegible]
.986	1168	[illegible]
.988	920	[illegible]
0.990	0.00000703	−0.0000020[illegible]
.992	502	[illegible]
.994	320	[illegible]
.996	177	115
.998	062	[illegible]

the terminals A and B, the current being denoted by i. When the switch is connected to point 1, the voltage measured is

$$e_1 = \frac{i}{\omega C} \tag{51}$$

With the switch on point 2, the measured voltage

$$e_2 = \omega M i = \omega^2 M C e_1 \tag{52}$$

Then

$$M = \frac{e_2}{e_1} \cdot \frac{1}{\omega^2 C} \tag{53}$$

The capacity C may be replaced by a resistance R. Then

$$M = \frac{e_2 R}{e_1 \omega} \tag{54}$$

If a variable standard of mutual inductance is available, any other mutual inductance whose value falls within the range of the standard may be readily measured. The primaries are connected in series to a voltage source, the secondaries in opposition to a telephone receiver or other indicating device, and the standard is varied until a null indication is obtained. The unknown mutual inductance then has the value indicated by the standard.

30. Calculation of Mutual Inductance.[1] The mutual inductance of two parallel coaxial circles may be calculated by the following method: first, calculate

$$\frac{r_2}{r_1} = \sqrt{\frac{\left(1 - \frac{a}{A}\right)^2 + \frac{D^2}{A^2}}{\left(1 + \frac{a}{A}\right)^2 + \frac{D^2}{A^2}}} \tag{55}$$

where a is the radius of the smaller circle, A the radius of the larger circle, and D the distance between the planes of the two circles. From the table shown on page 100 the value of F corresponding to the calculated value of r_2/r_1 is obtained. Then

$$M = F\sqrt{Aa} \tag{56}$$

The units are the same as in the formulas for self-inductance already given.

For two parallel coaxial multilayer coils of square or nearly square cross section, a good approximation is given by

$$M = n_1 n_2 M_0 \tag{57}$$

where n_1 and n_2 are the numbers of turns on the two coils, and M_0 is the mutual inductance of two circles located at the centers of the cross sections of the two coils.

The same formula may be used as a rough approximation for the mutual inductance of two coaxial single-layer solenoids.

References

BUTTERWORTH, S.: The High-frequency Copper Losses in Inductance Coils, *Exp. Wireless*, **2**, 613, 1925.

COURSEY, P. R.: Calculation and Design of Inductances, *Electrician*, **75**, 841, 1915.

GRIFFITHS, W. H. F.: Notes on Standard Inductances for Wave Meters and Other Radio-frequency Purposes, *Exp. Wireless and Wireless Eng.*, **6**, 543, 1929.

MAXWELL: "Electricity and Magnetism," Vol. II.

PERNOT, F. E.: Formulae and Tables for Design of Air-Core Inductance Coils, *Univ. Calif. Pub., Eng.*, **1**, 117, 1916.

WHITTEMORE, L. E., and G. BREIT: Inductance, Capacity, and Resistance of Coils at Radio Frequency, *Phys. Rev.*, **14**, 170, 1919.

WILMOTTE, R. M.: Parasitic Losses in Inductance Coils at Radio Frequency, *Exp. Wireless*, **2**, 451 and 477, 1925.

[1] ROSA, E. B., and F. W. GROVER, *Bur. Standards Sci. Paper* 169; GROVER, F. W., *Bur. Standards Sci. Papers* 320 and 498.

SECTION 5

CAPACITANCE

BY E. L. HALL,[1] E.E.

1. Capacitance. Capacitance is one of the three electrical quantities present in all radio circuits. The radio engineer endeavors to concentrate capacitance in definite well-known forms at definite points in the circuits, but capacitance exists between different conductors in the circuits and between the various conductors and the ground. Such capacitances, usually small, are ordinarily of no importance in the case of l-f or a-f currents but may be of great consequence in r-f circuits.

A *condenser* is an electrical device in which capacitances play the main role. While some inductance and some resistance are present, these quantities are usually of such minor importance that they are not considered.

A condenser has three essential parts, two of which are usually metal plates separated or insulated by the third part called the *dielectric*.

The amount of electricity which the condenser will hold depends on the voltage applied to the condenser. This may be expressed as $Q = C \times V$. The *capacitance* of the condenser is the ratio of the quantity of electricity and the potential difference or voltage, or $C = Q/V$ where Q is given in coulombs, C in farads, and V in volts. The capacitance of a condenser is dependent on the size and spacing of the plates and the kind of dielectric between the plates.

2. Units of Capacitance. The unit of capacitance is the *farad*. A condenser has a capacitance of one farad when one coulomb of electricity can be added to it by an applied voltage of one volt. This unit is too large for practical use so that a smaller unit, the microfarad, abbreviated μf, or one-millionth of a farad, is used. A condenser having a capacitance of one microfarad is much larger than is used in radio circuits. Condensers for such circuits usually have capacitances between a few thousandths and a few millionths of a microfarad. Another unit, the micromicrofarad, is often used. It is abbreviated $\mu\mu f$.

Another unit of capacitance sometimes used is the *centimeter*. The centimeter is equal to 1.1124 micromicrofarads.

3. Electrical Energy of Charged Condenser. Work is done in charging a condenser because the dielectric opposes the setting up of the electric strain or displacement of the electric field in the dielectric. The energy of the charging source is stored up as electrostatic energy in the dielectric.

The work done in placing a charge in the condenser is

$$W = \frac{1}{2}Q \times V = \frac{1}{2}CV^2 = \frac{Q^2}{2C}$$

[1] Radio Engineer, Radio Section, National Bureau of Standards.

where W is expressed in joules
Q is expressed in coulombs
V is expressed in volts.

The work done in charging the condenser is independent of the time taken to charge it.

4. Power Required to Charge Condenser. The average power required to charge a condenser is given by the equation

$$P = \frac{1}{2}\frac{CV^2}{t}$$

where P is expressed in watts
C is expressed in microfarads
V is expressed in volts
t is expressed in seconds.

If the condenser is charged and discharged N times per second the above equation becomes

$$P = \tfrac{1}{2}CV^2N$$

If an alternating e.m.f. of frequency f is used in charging the condenser, the equation may be written

$$P = CE_0^2f$$

where P = power in watts
C = capacitance in farads
E_0 = maximum value of voltage
f = frequency in cycles per second.

5. Dielectric Materials. The dielectric of a condenser is one of the three essential parts. It may be found in solid, liquid, or gaseous form or in combinations of these forms in a given condenser.

The simplest form of condenser consists of two electrodes or plates separated by air. This represents a condenser having a gaseous dielectric. If this imaginary condenser has the air between the plates replaced by a non-conducting liquid, such as transformer oil, and if the distance between the plates is the same as in the first case, it would be found that the capacitance was increased several times because the oil has a higher value of dielectric constant than air which is usually taken as 1.

If the space between the plates is occupied by a solid insulator, a condenser would result, which would be practical, as far as the possibility of constructing it is concerned. It would be found, in this case also, that the capacitance of the condenser was several times larger than when air was the dielectric.

The mechanical construction of either air or liquid dielectric condensers requires the use of a certain amount of solid dielectric for holding the two sets of plates.

There are a great many dielectric or insulating materials available from which the engineer may choose. It often is found that a material which is very good from the electrical standpoint is poor mechanically, or vice versa. Air is the gas generally used as a dielectric. Compressed air has been used in some high-voltage condensers, and compressed nitrogen and carbon dioxide are also in use.

Several kinds of oil have been used in condensers, such as castor oil, cottonseed oil, and transformer oil. More recently electrolytic condensers have come into use in radio equipment for use as filters and by-pass condensers where a large capacitance is required and either a d.c. or pulsating d.c. is applied.

Among the solids used as the condenser dielectric are mica, ceramic materials, and paper. Solid insulators used as mechanical supports in condensers include quartz, glass, Isolantite, porcelain, bakelite, mica, amber, hard rubber, Victron, Mycalex, etc.

6. Dielectric Properties of Insulating Materials. Such properties as surface and volume resistivity, dielectric strength or puncture voltage, dielectric constant, and absorption, are often considered in d-c and commercial-frequency applications. Such data are of little value if the insulating material is to be used at radio frequencies. For the latter application r-f measurements of various properties of the material are essential. A material which may be a satisfactory insulator for low frequencies may be worthless as an insulator at radio frequencies.

One of the most important properties of an insulator for radio frequencies is its power loss. This includes several factors which are difficult to separate but together indicate its suitability for radio purposes. The general idea of the imperfection of a condenser is brought out in several names such as "power loss," "power factor," and "phase difference," but they are not identical terms.

Dielectric constant is another important property of a material which has a definite bearing upon its use at radio frequencies.

Neither power loss nor dielectric constant alone can be used in selecting the best insulator for a particular application at radio frequencies. Some investigators have published results in which a product of the power loss and dielectric constant appears. This factor has no recognized name as yet but has certain merits for indicating more completely the suitability of an insulating material for radio uses.

7. Dielectric Constant. The *dielectric constant* K of an insulating material is the ratio of the capacitance C_x of a condenser using the material as the dielectric, to the capacitance C_a of the condenser using air as the dielectric, or $K = C_x/C_a$. This property of the material is sometimes called *inductivity* or *specific inductive capacity*.

The dielectric constant of a material is not a constant in the true sense of the word, but varies with the frequency, moisture content, temperature, voltage applied, and manner of applying it.

A table giving the dielectric constants of a large number of electrical insulating materials will be found in Art. 9.

8. Power Loss, Phase Difference, and Power Factor. Electrical insulating materials are not perfect in their insulating qualities, and there is a certain amount of power absorbed in them when used in an a-c circuit. A measurement of the power loss is the best single property that gives an indication of the suitability of an insulating material for use in radio circuits. Power loss can be expressed by a number of quantities, the most commonly used being resistance, power factor, phase difference and phase angle.

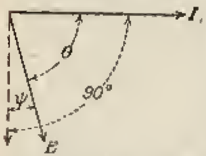


Fig. 1.—Phase in a capacitive circuit.

When a.c. flows in a condenser, the voltage across the condenser lags somewhat less than 90 deg. behind the current as shown by the angle θ (Fig. 1), called the *phase angle*. The complement ψ of the phase angle, is called the *phase difference*. The cosine of the phase angle is called the

power factor. The power loss in the insulating material is

$$P = EI\cos\theta$$

or

$$P = EI\sin\psi$$

where E = voltage across the condenser
I = current in amperes through the condenser
θ plus ψ = 90 deg. as shown in Fig. 1.

From the above, $\sin\psi = \cos\theta$, or the sine of the phase difference is equal to the power factor.

Fig. 2.—Condenser with dielectric losses.

When considering a condenser having dielectric losses, such as current leakage, brush discharge or corona, dielectric absorption or resistance in the plates, joints, contacts, leads, etc., it is customary to think of it as a perfect condenser C with a resistance R in series as shown in Fig. 2.

The voltage vectors may be shown as in Fig. 3, where the resultant voltage E flowing in the circuit is obtained by completing the vector diagram. The angle ψ is quite small for materials suitable for r-f insulators. For small angles the angle $\psi = \tan\psi$. In Fig. 3

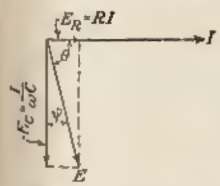


Fig. 3.—Vector relations in a condenser with dielectric losses.

$$\tan\psi = \frac{RI}{I/\omega C} = R\omega C = 2\pi fRC$$

If the resistance, capacitance, and frequency can be measured, the phase difference can be calculated from

$$\psi = 2\pi fRC$$

where ψ = phase difference in radians
f = frequency in cycles per second
R = resistance in ohms
C = capacitance in farads.

The following equation is sometimes convenient when wave length in meters is given

$$\psi = 0.1079\frac{RC}{\lambda}$$

where ψ = phase difference in degrees
R = resistance in ohms
C = capacitance in micromicrofarads
λ = wave length in meters.

For small angles, phase difference in radians is equal to power factor (nearly).

Power factor in per cent is 1.745 times phase difference in degrees. Power factor in per cent is given by the following equation:

$$\cos\theta = 2\pi fRC \times 10^{-7}$$

where $\cos\theta$ = power factor in per cent
f = frequency in kilocycles
R = resistance in ohms
C = capacitance in micromicrofarads

The leakage of electricity by conduction through the dielectric or along its surface contributes to the phase difference but is generally negligible at high frequencies. A condenser having leakage may be represented by a perfect condenser with a resistance in parallel as shown in Fig. 4. The current divides between the capacitance and the resistance, I_R through the resistance being in phase with the applied voltage E, and I_C through the capacitance leading E by 90 deg. as shown in Fig. 5. The resultant current I leads E by (90 deg. $-\psi$), where ψ is the phase difference. In Fig. 5

$$\tan \psi = \frac{E/R}{\omega CE} = \frac{1}{\omega RC}$$

or

$$\psi = \frac{1}{\omega RC}$$

Power factor is a term that involves all the power losses in a condenser. If the total power loss in a condenser is W watts, the voltage applied to it is V volts (r.m.s), and the current flowing through it is I amperes (r.m.s.): the

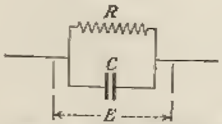


Fig. 4.—Equivalent of condenser with leakage.

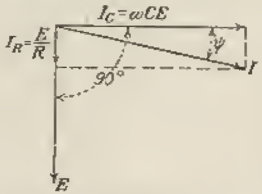


Fig. 5.—Vectors in condenser with leakage.

power factor of the condenser is W/VI. The relation between I (amperes) and V (volts) for a condenser of capacitance C (microfarads) operating at a frequency f is

$$I = \frac{2\pi fCV}{10^6} = \frac{\omega CV}{10^6}$$

The power factor of a condenser in per cent may be written

$$\cos \theta = \frac{W \times 10^6}{2\pi fCV^2} = \frac{W \times 10^6}{\omega CV^2}$$

Referring again to Fig. 2 showing the perfect condenser C and resistance R replacing the actual condenser, the value of R can be calculated from the equation $W = I^2R$. The quantity R is known as the *equivalent resistance* of the condenser at the given frequency.

The expression $W \times 10^6/\omega CV^2$ for power factor can be changed into the expression involving resistance, capacitance, and ω by substituting I^2R for W and then substituting $\omega CV/10^6$ for I, giving power factor equal to

$$RC\omega \times 10^{-6}$$

9. The following table gives dielectric constant and power factor at certain frequencies of a large number of electrical insulating materials, as obtained from the sources given at the end of the table. While in some cases data from different sources do not agree, differences in composition, method of making measurements, and condition of samples may account for such disagreements.

10. Dielectric Strength. The *dielectric strength* of an insulating material is the minimum value of electric field intensity required to rupture it. Dielectric strength is usually expressed in kilovolts per centimeter of dielectric thickness. The fall in insulation resistance with

Values of Dielectric Constant and Power Factor for Electrical Insulating Materials at Radio Frequencies

Material	Frequency, kilocycles	Dielectric constant	Power factor, per cent	Source
Alsimag No. 211	60,000	4.4	0.04	1
	120,000	4.4	0.05	
Alsimag No. 196	60,000	4.9	0.10	1
	120,000	5.0	0.13	
Amber	187.5	....	0.459	2
	300	....	0.476	
	429	....	0.478	
	600	....	0.495	
	1,000	....	0.513	
Calan	300	....	0.036	3
	1,000	....	0.032	
	3,000	....	0.028	
	10,000	....	0.026	
	50,000	....	0.025	
Calit	300	....	0.041	3
	1,000	....	0.038	
	3,000	....	0.037	
	10,000	....	0.034	
	50,000	....	0.032	
Celluloid	1,000	6.2	5 to 10	4
photographic film	A	6.7	4.2	5
Cellulose nitrate, laboratory product		3.8	2.8	
Cement, de Khotinski, medium hard		3.9	3.68	
Portland	1,000	0.018–0.029	6.8–8.0	6
Condensa	300	....	0.097	3
	1,000	....	0.08	
	3,000	....	0.072	
	10,000	....	0.061	
	50,000	....	0.057	
Condensa C	300	....	0.072	3
	1,000	....	0.06	
	3,000	....	0.041	
	10,000	....	0.032	
	50,000	....	0.028	
Fiber, black	A	7.6	4.35	5
red		4.8	4.89	
oil impregnated		5.8	3.68	
hard, dry	1,000	5.0	5.0	6
Frequenta	300	....	0.047	3
	1,000	....	0.038	
	3,000	....	0.030	
	10,000	....	0.028	
	50,000	....	0.026	
Frequenta D	1,000	....	0.038	3
	10,000	....	0.019	
	50,000	....	0.019	
Glass	30	5.1–7.9[a]	0.35–2.98	7
	600	....	0.04–0.65[b]	2
borosilicate No. 707	60,000	3.7	0.12	1
	120,000	3.7	0.12	
borosilicate	18,000	5.1	0.59	8
cobalt	500	7.3	0.70	9
electrical	100	5.7	0.4	10
heat resisting	A	5.7	0.61	5
photographic, with gelatin coating		7.5	1.00	
without gelatin coating		7.5	0.86	

VALUES OF DIELECTRIC CONSTANT AND POWER FACTOR FOR ELECTRICAL INSULATING MATERIALS AT RADIO FREQUENCIES.—(*Continued*)

Material	Frequency, kilocycles	Dielectric constant	Power factor, per cent	Source
Nonex No. 772	60,000	4.2	0.28	1
	120,000	4.2	0.25	
plate	14	...	0.97	11
	100	...	0.77	
	500	...	0.66	
	1,000	...	0.62	
	500	6.8	0.70	9
	1,000	8.4	1.0	6
American plate	*A*	7.6	0.93	5
Pyrex	14	...	0.88	11
	100	...	0.74	
	500	...	0.67	
	750	...	0.68	
	500	4.9	0.42	9
Pyrex No. 774	60,000	4.1	0.54	1
	120,000	4.2	0.51	
soda lime No. 008	60,000	6.1	1.06	
	120,000	6.2	1.03	
window	*A*	8.0	0.87	5
Hard rubber	210	3.0	0.88	9
	440	...	0.88	
	710	...	0.88	
	1,126	3.0	1.05	
	600	...	0.62	2
	1,000	...	0.68	
	1,000	3.0–5.0	0.015–0.02	6
	18,000	2.9	0.76	8
	300	...	0.65	3
	1,000	...	0.64	
	3,000	...	0.61	
	10,000	...	0.57	
	50,000	...	0.53	
	60,000	3.1	0.83	1
	120,000	3.1	0.84	
low-loss	60,000	2.9	0.57	
	120,000	3.0	0.50	
Isolantite	250–1,500	6.1	0.18	12
Italian lavite	60,000	4.7	0.15	1
	120,000	4.8	0.17	
Lucite	1,000	2.5–3.0		13
Marble, white	*A*	9.3	0.52	5
gray		11.6	4.2	
blue		9.4	1.22	
Mica	300–50,000	...	0.017	3
clear, muscovite	1,000	6.5–8.0	0.01–0.06	6
U. S. muscovite	100–1,000	8.69–6.57	0.04–0.01	14
India muscovite		7.90–7.07	0.02–0.01	
India	600	...	0.017	2
	A	0.4	...	5
amber	1,000	5.4–5.8	0.04–7.1	6
built-up, shellac binder	*A*	5.6	1.75	5
Mycalex	100	8.0	0.2	10
	300	...	0.19	3
	1,000–50,000	...	0.18	
	1,000	8.0	0.2	6
Mineral oil	1,000	2.7	0.08	

VALUES OF DIELECTRIC CONSTANT AND POWER FACTOR FOR ELECTRICAL INSULATING MATERIALS AT RADIO FREQUENCIES.—(*Continued*)

Material	Frequency, kilocycles	Dielectric constant	Power factor, per cent	Source
Phenolic insulation, laminated (bakelite)	190	5.4–5.8	3.85–7.35	9
	1,100	5.1–5.6	4.20–6.65	
black	18,000	4.7	6.0	8
natural brown		4.4	5.6	
Phenolic resin, laminated compound, highest grade, paper base	1,000	5.0	3.5	6
cloth base	1,000	5.0	4.5	
molded compound, wood-flour filler	1,000	5.5	3.5	
mica filler	1,000	6.0	1.0	
Polyindene	1,000	3.0	0.04	4
Polystyrene	60,000	2.6	0.05	1
	120,000	2.5	0.07	
Styron	35,000	2.64	0.01	4
Trolitul	750	2.6	0.02	
	3,000	...	0.03	
	6,500	...	0.058	
	13,600	...	0.125	
	60,000	...	0.07	
Porcelain	100	7.0	0.7	10
	300	...	0.70	3
	1,000	...	0.55	
	3,000	...	0.49	
	10,000	...	0.63	
	50,000	...	0.85	
wet process	1,000	6.5–7.0	0.6–0.8	6
Quartz	300–10,000	...	0.010	3
	50,000	...	0.011	
fused	1,000	4.1	0.02	6
clear	60,000	3.8	0.03	1
	120,000	3.8	0.05	
milky	60,000	3.5	0.03	
	120,000	3.5	0.05	
Rosin	1,000	3.3–4.7	0.26–0.37	6
Shellac	1,000	6.0	7.0	4
Shellac film	1,000	4.1	2.5	6
Slate	1,000	12.4–19.0	45–63	6
electrical	*A*	30.0	63	5
Steatite	800	...	0.20	15
	45,000	...	0.16	
	300	...	0.21	3
	1,000	...	0.20	
	3,000	...	0.18	
	10,000	...	0.17	
	50,000	...	0.15	
"commercial"	1,000	6.5	0.20	16
	10,000	6.2	0.18	
"low loss"	1,000	6.5	0.06	
	10,000	6.0	0.04	
Ultra-Calan	1,000–10,000	...	0.010	3
	50,000	...	0.011	
Ultra-Steatite	1,000	...	0.1	15
	50,000	...	0.06	
Urea resin, wood-flour filler	1,000	5.7	3.0	4
	10,000	5.5	3.8	
	35,000	5.3	4.2	

VALUES OF DIELECTRIC CONSTANT AND POWER FACTOR FOR ELECTRICAL INSULATING MATERIALS AT RADIO FREQUENCIES.—(*Continued*)

Material	Frequency, kilocycles	Dielectric constant	Power factor, per cent	Source
Varnish film, clear, linseed-oil	1,000	2.2	1.2	6
clear gum	1,000	3.2	1.1	6
black asphaltic	1,000	2.0	0.8	6
spar	*A*	5.5	3.15	5
insulating	*A*	4.8	5.25	5
Varnished cloth, yellow	1,000	2.5	3.0	6
black	1,000	2.0	2.0	6
Victron resin, clear	446–877	2.96	0.03	15
Vitrolex	1,100	6.4	0.3–0.4	15
Vulcanized rubber	18,000	3.9	2.9	8
Wax, beeswax	*A*	3.2	1.63	5
	1,000	2.9	2.5	6
ceresin	1,000	2.5–2.6	0.12–0.21	6
paraffin	1,000	2.5	0.97	6
Wood, basswood, dry	*A*	2.0	1.92	5
baywood, dry	*A*	2.4	2.45	5
cypress, dry	*A*	2.0	2.1	5
fir, dry	*A*	3.1	3.5	5
maple, dry	*A*	2.6	2.45	5
oak, dry	*A*	3.1	2.97	5
birch	500	5.2	6.48	9
maple	500	4.4	3.33	9
oak	300	....	3.68	9
	425	3.3	3.50	9
	635	....	3.85	9
	1,060	....	4.20	9
whitewood, dry	18,000	1.7	2.3	8

[a] Range of nine samples of various chemical compositions reported.
[b] Range of 27 samples of various chemical compositions reported.
[A] Measurements made between 80 and 1,875 kc.

[1] MILLER, J. M., and B. SALZBERG, Measurements of Admittances at Ultra-high Frequencies. *RCA Rev.*, **3**, 480–504, April, 1939.

[2] SCHOTT, ERICH, Hochfrequenzverluste von Gläsern und einigen anderen Dielektrics, *Jahrb. drahtlosen Tele. Tele.*, **18**, 82–122, August, 1921.

[3] HANDREK, H., Keramische Spezialmassen, *Archiv. für tech. Messen*, **44**, 28, 29, February, 1935.

[4] BLOOMFIELD, G. F., Insulating Materials for the Higher Frequencies, *T & R Bull.* **14**, 635–639, May, 1939; *Radio Tech. Digest*, No. 13, 23–32, September–October, 1939.

[5] PRESTON, J. L. and E. L. HALL, Radio-frequency Properties of Insulating Materials, *QST*, **9**, 26–28, February, 1925.

[6] General Electric Company.

[7] DECKER, WILLIAM C., Power Losses in Commercial Glasses, *Elec. World*, **89**, 601–603, Mar. 19, 1927.

[8] CHAFFEE, J. G., The Determination of Dielectric Properties at Very High Frequencies, *Proc. I.R.E.*, **22**, 1020, August, 1934.

[9] HOCH, E. T., Power Losses in Insulating Materials, *Bell System Tech. Jour.*, **1**, November, 1922.

[10] BROWN, W. W., Properties and Applications of Mycalex to Radio Apparatus, *Proc. I.R.E.*, **18**, 1307–1315, August, 1930.

[11] MACLEOD, H. J., Power Losses in Dielectrics, *Phys. Rev.*, **21**, 53–73, 1923.

[12] Isolantite circular.

[13] *The Neoprene Notebook*, No. 23, 95, January-February, 1940.

[14] LEWIS, A. B., E. L. HALL, and F. R. CALDWELL, Some Electrical Properties of Foreign and Domestic Micas and the Effect of Elevated Temperatures on Micas, *Bur. Standards Jour. Research*, **7**, 409, August, 1931.

[15] Dielectric Products Corp. circular.

[16] THURNAUER, H., Notes on Steatite-type High-frequency Insulation, *QST*, **21**, 33, November, 1937.

rise in temperature is a factor of great importance in connection with the breakdown of a dielectric under the applied voltage. Insulating materials are not strictly homogeneous. The current leak through an insulating material may perhaps be concentrated in a few small paths through the material, and the energy loss due to the leakage, while small, may be large compared with the area through which it is flowing. The paths of the current flowing through the dielectric become heated with a resulting lowering of the resistance of the path and an increase in the current leakage. The heating of the dielectric may lead to rapid deterioration, particularly if moisture is present, and ultimate breakdown. The length of time of the application of the voltage has a definite bearing upon the breakdown voltage. Most dielectrics will withstand for a very brief period a much higher voltage than they can when the voltage is applied for a longer period.

These effects have dictated two tests for condensers, a high flash-test voltage of very brief duration and the application of a much lower voltage for a longer period.

The dielectric strength of a material is usually found to be lower for r-f voltages than for a-f or d-c voltages. The rupturing voltage at radio frequencies depends on the rapidity with which the voltage is raised and is not nearly so definite a phenomenon as l-f puncture voltage. Dielectric strength of solid insulators is difficult to measure because of the complexity of the experimental effects. As the r-f currents flow in the material, heating, corona, flashover, and possible deterioration, blistering, or charring may result with consequent changing of voltage and current as the time of application elapses.

If high r-f voltages are applied to an air condenser, a corona discharge may be set up which appears as a visible glow around high-potential metal parts, points, and sharp edges and is usually distinctly audible. These corona effects represent a power loss in the condenser. Hence the construction of air condensers for high voltages requires the rounding of all edges and corners and the avoiding of sharp points which encourage the formation of corona and flashover.

11. Dielectric Absorption. When a condenser is connected to a d-c source of e.m.f., the instantaneous charge is followed by the flow of a small and steadily decreasing current into the condenser. The additional charge is absorbed by the dielectric. Similarly the instantaneous discharge of a condenser is followed by a continuously decreasing current. The condenser does not become fully charged immediately, nor does it completely discharge immediately when its terminals are shorted, but several discharges may be secured when the condenser possesses dielectric absorption. The maximum charge in a condenser cyclically charged and discharged varies with the frequency of charge.

If a condenser evidencing dielectric absorption is used at radio frequencies, a power loss occurs which appears as heat in the condenser. The existence of power loss indicates a component of e.m.f. in phase with the current as though a resistance were in series with the condenser as shown in Fig. 2. The effect of dielectric absorption can be measured along with other losses in the condenser, although dielectric absorption represents the chief power loss in solid dielectrics.

12. Calculation of Capacitance. Formulas are available for use in calculating the capacitance for a large number of geometrical shapes of conducting surfaces such as spheres and cylinders, either separated or concentric, and flat surfaces of various shapes. The usual type of condenser calculations are concerned with two or more flat conductors.

When two conducting plates are parallel, close together, and of large area the capacitance of the condenser is given by

$$C = 0.0885 \times \frac{KS}{t}$$

where C = capacitance in micromicrofarads
K = dielectric constant (which is 1 for air)
S = area of one plate in square centimeters
t = distance between plates in centimeters.

When more than two plates are used in the condenser, the formula becomes

$$C = 0.0885 \times \frac{KS(N - 1)}{t}$$

where N = number of plates.

The actual capacitance of a parallel plate condenser is slightly larger than the value as calculated from the above formula, because of the fringing of the electric lines of force beyond the space between the plates. A correction can be made for this fringing by slightly increasing the dimensions of the plates. A narrow strip of width w can be added to the actual plate dimensions. In the case of circular plates $w = 0.4413t$, and for plates with straight edges $w = 0.110t$, where t is the distance between the plates in centimeters.

13. Combinations of Condensers. Combinations of two or more condensers in a circuit are often arranged in either series or parallel. Condensers connected in parallel give a total capacitance equal to the sum of the capacitances of the individual condensers. Condensers connected in series give a resulting capacitance which may be calculated from the following:

$$C = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots}$$

This formula gives the following expression in the case of two condensers in series

$$C = \frac{C_1 \times C_2}{C_1 + C_2}$$

The various elements such as tubes, sockets, mountings, wiring, etc., in radio apparatus contain many small capacitances by virtue of the difference of potential existing between the numerous conductors insulated from one another. These small capacitances are known as *stray* capacitances. While they are unimportant in some kinds of work, in other types of work, such as in amplifier design, they must be taken into account. In the case of resistance-coupled amplifiers, for example, these capacitances reduce the amplification at the higher audio frequencies and make a flat-characteristic with high over-all gain impossible.

The effect of stray capacitances is eliminated in the case of condensers used as capacitance standards by shielding the insulated plates and

[1] COURSEY, PHILIP R., "Electrical Condensers," Sir Isaac Pitman & Sons, Ltd. London.

grounding the shield. In this manner a definite capacitance is always assured for a given scale setting.

14. Effect of Frequency on Condenser Capacitance. One of the most important considerations is the effect of frequency upon the capacitance value of a condenser. In the best condensers this effect is nil. In fact one of the criterions of a suitable condenser for a capacitance standard is that its capacitance shall be the same for two different sets of charging and discharging conditions. A variable air condenser, such as the Bureau of Standards type described on page 120 of the Bureau's *Circ.* 74, gives the same capacitance at 100 and at 1,000 charges and discharges per second. A condenser having considerable solid dielectric in its make-up will show a difference in capacitance with frequency. The quantity of electricity which flows into a condenser during a finite charging period is greater than would flow in during an infinitely short charging period. Consequently the measured or apparent capacitance with a.c. of any finite frequency is greater than the capacitance on infinite frequency, the latter being called the *geometric* capacitance. The capacitance of a condenser decreases as the frequency increases.

The length of the internal leads of a condenser should be kept as short and direct as possible to minimize the inductance of the leads which acts to give an apparent change of capacitance with frequency. The amount of this change can be calculated from $C_a = C(1 + \omega^2 CL \times 10^{-12})$, where C_a is the apparent or measured capacitance, C is in μf, and L in μh.

15. Types of Condensers. There are many ways in which condensers might be classified, by their construction, size, voltage rating, use, dielectric, or fixed or variable capacitance. The condensers used in various radio applications are found in innumerable sizes, shapes, and uses. The two simplest divisions into which condensers may be classified have to do with their capacitance, *i.e.*, whether it is fixed or variable.

16. Types of Fixed Condensers. Fixed condensers are available in all capacitance ranges from a few micromicrofarads to several microfarads, for any voltage rating up to 45,000 volts or higher, and in innumerable shapes and dimensions, all depending upon the use for which the condenser is intended.

Paper formerly was used as the dielectric for condensers for use on lower voltages, while mica was used in condensers for higher voltages. More recently as the art of condenser manufacture has progressed, an oil-impregnated paper dielectric is used in condensers for the higher voltages, the whole condenser being mounted within an oil-filled container.

For paper dielectric 100 per cent pure linen paper is used, which must meet severe requirements as to thickness, porosity, uniformity, width, freedom from conducting particles, alkalinity, and acidity. Two or more layers of paper are used between the metal foil plates, depending upon the voltage for which the condenser is designed. Paper condensers are available in hermetically sealed plug-in types to fit standard octal-type radio-tube sockets; both in wax-impregnated and oil-impregnated types for d-c working voltages up to 600.

Paper condensers are formed by winding two metal foil electrodes or ribbons in conjunction with the paper ribbons. There are two types of winding, *inductive* and *non-inductive*. The latter type is recommended for r-f and for the higher a-f work. The inductive type is satisfactory for l-f work.

In winding the inductive type of condenser, the foil used is narrower than the paper, and the contact is made with the foils by tinned copper strips inserted in the winding. The non-inductive type of winding is made with the foils about the same width as the paper. The foil is staggered so that the condenser plates project over the ends of the paper. The terminals are soldered to the extending foil at the opposite ends and thus make contact with every turn of the foil. The latter type of construction makes for minimum plate resistance and minimum power loss.

Mica has been used very extensively for condensers for use at radio frequencies. India mica has been used almost exclusively as it has been generally considered as of superior quality for radio use.

Selected mica is split into sheets of definite thickness, gaged, and tested for punctures or other defects. A condenser is built up of alternating mica and metal foil sheets, the sets of plates of opposite polarity being brought out at opposite ends where they are soldered together, forming the two terminals. The whole stack of plates is rigidly clamped together in such a way as to firmly grip the plates in the center and expel all dielectric other than mica. The condenser may be mounted in a suitable container.

During the last few years attention has been given by the manufacturers to the development of small condensers of great stability, or whose capacitance changes with temperature are a definite amount, positive or negative, as desired. The advent of push-button tuned receiving sets has required the use of small condensers which would maintain their capacitance as the receiver warmed up or would change their capacitance so as to compensate for changes in the coils. A type of condenser now available with positive, zero, or negative temperature coefficient employs a small ceramic tube as the dielectric, with silver plating inside and out followed by copper plating and solder forming the two electrodes, to which wire leads are soldered. Wax impregnation and moistureproof lacquer complete the condenser, which is said to be uneffected by changes in temperature and humidity. Condensers of this type have a d-c working voltage of 500 and can be obtained in sizes from 5 to 1000 $\mu\mu$f.

Another type of low-temperature coefficient condenser uses silver plating on mica and is mounted either in a ceramic or low-loss bakelite case. These condensers are wax-impregnated and sealed. They have small positive temperature coefficients.

If a condenser is to be used with higher voltages, the practice is to construct the condenser with two or more condenser sections in series, rather than to increase the thickness of the mica. The former method is more flexible than the latter, permitting the construction of condensers for 45,000 volts or higher.

It is customary to mount the large high-voltage condensers in steel tanks which are filled with a high flash-point insulating oil which serves to prevent access of dirt and moisture, prevents flashover along the condenser sections, insulates the condenser from the tank, and conducts heat away from the condenser elements.

17. Electrolytic Condensers.[1] Another type of fixed condenser which has important applications is known as the *electrolytic condenser*. Its advantages are low cost and high capacitance as compared with other types of fixed condensers. A unit of 8 μf, 500-volt d-c rating may be manufactured in a tubular assembly 7/8 in. in diameter by 1 11/16 in. long.

The electrolytic condenser consists of three essential components: the anode, the dielectric film, and the electrolyte. The anode is always

[1] Data supplied by S. H. Walters, Cornell-Dubilier Electric Corp.

made of aluminum of high purity and forms one plate on the condenser. The dielectric film is formed electrochemically on the anode and is very thin. The electrolyte may be either a liquid or a pastelike substance. It is the second plate of the condenser, insulated from the anode plate by virtue of the dielectric film formed on the latter.

Electrolytic condensers may be divided into two general classes:

1. *Dry* electrolytic condensers in which a pastelike form of electrolyte is used.
2. *Wet* electrolytic condensers in which a liquid or waterlike electrolyte is used.

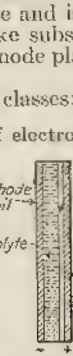


FIG. 6.—Electrolytic condenser construction.

The electrolyte in the case of dry electrolytic condensers is absorbed in porous paper and held in position adjacent to the anode foil by this paper. In addition another aluminum foil, generally called the *cathode foil* is incorporated for the purpose of making electrical contact to the electrolyte-saturated paper.

In the wet type the electrolyte is a dilute water solution. The anode member with adhering dielectric film is suspended in a can, generally of aluminum. The can is then filled with the electrolyte. The can acts as the electrical connection to the electrolyte similar to the cathode foil in the case of the dry electrolytic condenser.

For a given area of anode surface the capacitance in microfarads of the condenser is inversely proportional to the thickness of the dielectric film. The film thickness is proportional to the voltage during the electroformation of the film. Therefore, condensers with very low voltage ratings may be made with very high capacitances. The ordinary ranges are 500 to 6,000 μf in capacitance for voltage ratings of 6 to 60 volts d.c. and 2 to 100 μf with voltage ratings of 100 to 150 volts d.c.

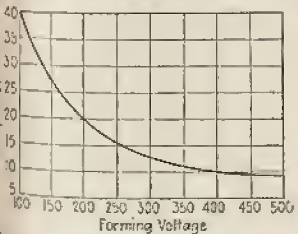


FIG. 7.—Electrolytic condenser characteristic.

18. Electrolytic Condenser Characteristics. The d-c voltage which an electrolytic condenser can withstand is governed by the voltage at which the original film is applied. It is necessary that the anode always be connected to the positive side of the voltage source. An electrolytic condenser connected in this manner will operate satisfactorily as long as the applied voltage is of correct polarity and does not exceed rated voltage for more than a few seconds at a time. A reversal of potential will cause the unit to draw considerable current even at low voltages. A d-c voltage in excess of rated causes the unit to draw an appreciable leakage current.

Dry electrolytic condensers have a definite breakdown voltage at which permanent failure occurs. Momentary surges less than this breakdown voltage but higher than operating voltage will ordinarily do no damage.

If the anode area is such as to give 8 μf when the working voltage is 500 volts d.c., then the same area at lower working voltages will yield a capacitance as indicated on the curve of Fig. 7.

Electrolytic condensers have a power factor which is considerably higher than other types of fixed condensers. This is due in part to the fact that one of the conducting plates is the electrolyte which has considerably higher resistance than the conventional metallic plates of the other types. In effect this places a resistance in series with the condenser and hence causes a high power factor of the entire unit. Dry electrolytic condensers have a power factor of about 6 per cent at 60 to 120 cycles. Power factors increase with frequency, and for this reason the use of electrolytic condensers is generally confined to the l-f application.

19. Etched-foil Types. Within recent years methods of treating the smooth foil surface in such a way as to make it extremely rough have been applied to electrolytic condensers. The anode then has an increased total area over and above the original smooth surface. The dielectric film follows the contours of the foil, and the result is a great increase in capacitance with no increase in volume for any given working voltage.

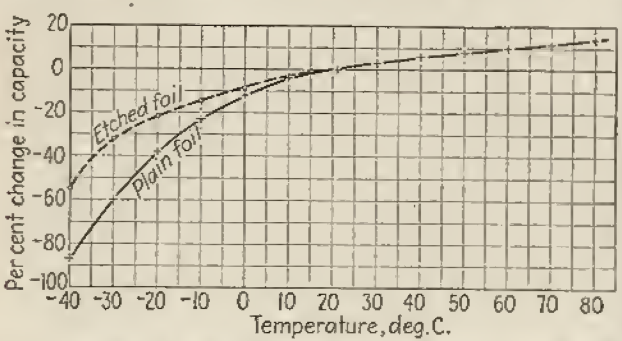


FIG. 8.—Comparison of etched- and plain-foil temperature characteristics.

Several methods have been evolved for formation of a roughened surface for the anode foil. Chief among these are as follows: (1) etching whereby the smooth foil surface is attacked either chemically or electrochemically; (2) mechanical roughening, whereby the surface is roughened by suitable abrasive; and (3) aluminum spraying, whereby the aluminum is sprayed in molten state on a suitable carrier medium.

20. Characteristics under Adverse Conditions. Electrolytic condensers operate best under normal condition of temperature. The limitations of the electrolyte and the film properties are the governing factors in the operation of this type of condenser.

Extremely high temperatures cause the electrolyte to dry out and increase in resistivity with consequent increase in power factor when normal temperature is again reached. Furthermore the increase in d-c leakage current with temperature must be considered since there is a danger of the start of a destructive cycle due to the generation of internal heat because of the increased d-c leakage.

Temperatures up to 140°F. are considered normal although temperatures up to 185°F. are not dangerous if the condenser is rated at 50 to 100 volts higher than the actual operating voltage.

Low temperature causes a decrease in capacitance and an increase in power factor. These changes are temporary and are restored to normal when normal temperatures are again reached.

Where high operating temperatures are to be experienced, the construction should be hermetically sealed in metal cans. This construction limits the loss of electrolyte to a minimum, and longer life is to be expected.

21. Applications. The nature of electrolytic condensers makes them particularly suitable for filter circuits in power supplies where a relatively high capacitance is required together with the ability to withstand a d-c potential and small superposed a-c ripple. Second only in importance is the use as a-f by-pass condensers across screen grids and cathode bias resistors. The use of a-c electrolytics wherein the cathode foil is replaced with a second anode is important in capacitor motor service. These latter units are divided into two classes, those for intermittent duty and those for continuous duty. The former are rated at from 30 to 500 μf at 110 volts a.c. and the latter at 10 to 50 μf at 25 volts a.c.

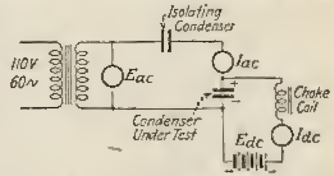


FIG. 9.—Production testing circuit for electrolytic condenser.

The intermittent-duty type functions only during the starting of a capacitor motor (capacitor start-induction run). The continuous-duty type functions in the smaller motors rated at about 1/100 hp. and is continuously on voltage during the operation of the motor.

22. Testing. The circuit of Fig. 9 is generally used to test electrolytics in production. E_{dc} supplies a polarizing voltage so that the voltage

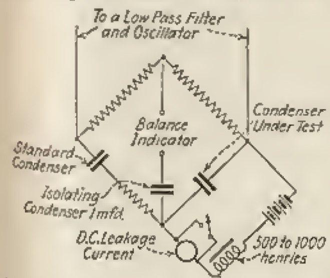


FIG. 10.—Circuit for measuring electrolytic condenser capacitance.

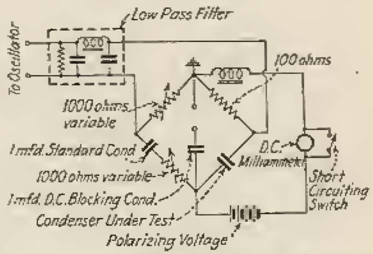


FIG. 11.—Capacitance and power factor measurement.

across the condenser will be pulsating d.c. The isolating condenser prevents short-circuiting the polarizing voltage. If E_{dc} is maintained at a constant value, the a-c milliammeter may be calibrated in terms of the capacitance of the condenser under test. I_{dc} reads the d-c leakage current through the condenser.

For the accurate measurement of capacitance and power-factor bridge systems such as those shown in Fig. 10 or 11 should be used. They are essentially standard bridge systems rearranged to permit the application of a polarizing voltage.

23. Types of Variable Condensers. The most common type of variable condenser consists of a series of parallel metal plates fastened to a shaft capable of rotation so that the moving plates intermesh with a set of fixed plates. Air is the main dielectric in such condensers, although some solid insulating material is required to ensure that the two sets of plates are correctly located with respect to each other. Many ways of insulating the plates from each other have been devised, using one or more pieces of the insulating material in sheet, rod, or bar form. Bakelite, hard rubber, Pyrex, porcelain, fused quartz, and Isolantite are some of the materials used for such insulators.

The most common use of a variable condenser is in association with a coil, the combination forming a circuit resonant to a band of radio frequencies depending upon the coil constants and the capacitance range of the condenser. For a number of applications it is more convenient to have the capacitance change in a different way than proportional to the angle of rotation of the plates. This first resulted in the "decremeter" plate and the straight-line wave-length plate. As the use of frequency

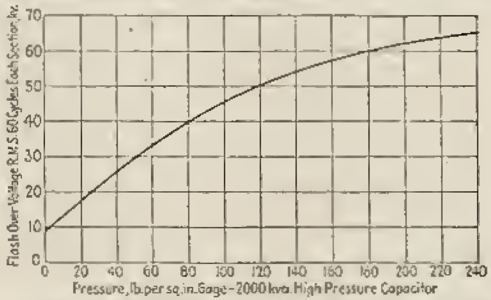


Fig. 12.—Flash-over voltage (60 cycles) of 2,000-kva. capacitor.

rather than wave length became common, the straight-line frequency plate came into use and later the "mid-line" plate. There are other possibilities such as straight-line percentage wave length and straight-line percentage frequency, the latter being of advantage in frequency measurements. In any of the above shapes or classifications, the movable plates formerly were so shaped as to give the desired frequency or wave-length curve. This resulted in an ill-shaped plate difficult to balance or to hold to a desired setting. In some cases semicircular rotating plates were used with the fixed plates cut away so as to obtain the desired curve. In any of the special forms of plates the plate shape may vary. The minimum and maximum capacitances of the condenser play a large part in determining the outline of the plate.

Brass or aluminum plates and steel shafts are ordinarily used. If the condenser is intended for use on high voltages, the spacing between opposite plates must be sufficient to avoid a flashover or arcing between plates. It is customary to round off all sharp edges and corners in such condensers to avoid flashover.

Condensers of the air type are often filled with oil, which increases the voltage that they can stand and increases the capacitance from two to five times depending on the dielectric constant of the oil used.

Compressed-air condensers were formerly used in some radio transmitting stations. The voltage which such a condenser will stand is increased without changing the capacitance.

Compressed gas condensers, utilizing nitrogen under pressure up to 2,000 lb. per square inch as the dielectric, are now being extensively used in broadcast transmitters. The advantages of low loss and permanent characteristics of this type of condenser have long been recognized, but it is only of recent date that any attempt has been made to offer a wide commercial selection of this type of condenser.

One manufacturer offers three lines of condensers with flashover ratings of 15, 20, and 30 kv r.m.s. at 1,000 kc, and capacitance ranges up to 1,000, 1,500, and 2,000 $\mu\mu$f, respectively. These are available in fixed, adjustable, or continuously variable types. Special units have been built with flashovers up to 60 kv r.m.s. and capacitances up to 20,000 $\mu\mu$f.

Construction varies somewhat with different manufacturers. One offers a completely non-magnetic assembly using heat-treated aluminum tank and end closures. As a typical example, a variable condenser having 30 kv r.m.s. flashover rating will have a height of 36 in., an over-all diameter of 12 in., and a weight of 90 lb.

Gases other than nitrogen have been used, some of which show considerable promise in increasing flashover voltage and reducing size and weight. These condensers are available in either fixed or variable capacitance types and in sizes from 100 to 2,000 $\mu\mu$f.

24. Gang Condensers. The single-dial control radio receiver brought problems to the designer in how to tune two to five circuits accurately using a corresponding number of similar coils and variable condensers operating on the same shaft. As it is practically impossible conveniently to manufacture two condensers exactly alike, to say nothing of three or four alike, so that their capacitances shall be exactly the same throughout the complete rotation of the condenser plates and accurately tune the condensers with the same number of similar coils which differ slightly in value, it has been customary to balance or equalize these tuned circuits by the addition of small paralleling condensers called *trimmer* or *padder* condensers. Such condensers can be obtained matched to one-half of 1 per cent. It is possible to obtain two to four condensers called *gang* condensers for radio receivers arranged with their shafts in line and operated by one dial, matched to one-half of 1 per cent. The individual condensers may be separated from one another by metal shields if desired.

25. Design Equations for Variable Air Condensers. The capacitance of a condenser made up of three plates as indicated in Fig. 13 can be obtained by determining the area of the overlapping plates, the distance between the adjacent plates, and substitution of these values in the general equation given in Art. 12. The area of the shaded portion of Fig. 13 is $\frac{1}{2}\pi(r_1^2 - r_2^2)$. The distance between the plates is $\frac{1}{2}(s - t)$. Substituting these values in the general equation, the capacitance of the condenser is given by

$$C = \frac{0.0885\frac{1}{2}\pi(r_1^2 - r_2^2) \times (3 - 1)}{\frac{1}{2}(s - t)}$$

The maximum capacitance of a condenser with N plates can be obtained by using a similar equation which may be written

$$C = \frac{0.278(r_1^2 - r_2^2)(N - 1)}{(s - t)}$$

In the above equations C is in micromicrofarads and the dimension r_1, r_2, s, and t in centimeters. These equations neglect the capacitance through the solid insulation which is used in the condenser and the fringing effect, the correction for which is in Art. 12. Many condensers are made to have as small a minimum capacitance as possible, giving a large ratio of maximum to minimum capacitance, but this is of doubtful advantage, as slight changes of capacitance due to warping of plates or wear in bearings will cause a relatively large error at the lower end of the scale but practically no noticeable effect at the maximum capacitance end of the scale.

A semicircular plate condenser gives a capacitance calibration curve similar to C shown in Fig. 14. With the exception of the portions near

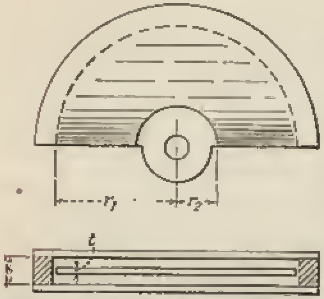


FIG. 13.—Dimensions useful in determining condenser capacitance.

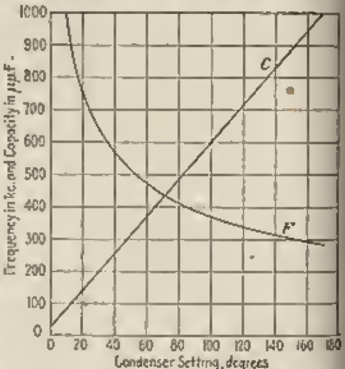


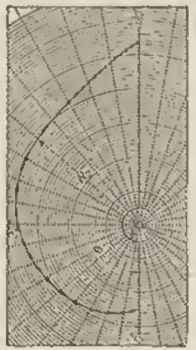
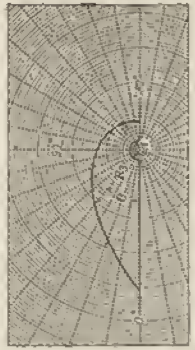
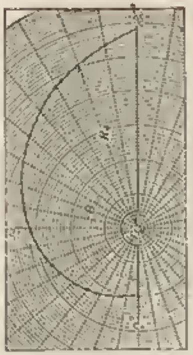
FIG. 14.—Semicircular plate condenser characteristic.

the ends of the curve, it is practically a straight line. In practice, the lower 10 and upper 5 or 10 deg. of a 180-deg. scale are not used, so as to avoid the curvature in the calibration curve in these regions. Zero setting does not give zero capacitance.

A curve for such a condenser used with a coil is shown at F in Fig. 14. The frequency changes very rapidly on the lower part of the scale. A slight capacitance change would make a large frequency change. Therefore, when using frequency meters having semicircular plate condensers which constitute the main capacitance of the circuit, the coils should be so designed as to give overlaps without resort to the low-capacitance end of the scale.

As the wave length λ of a wavemeter circuit is proportional to $\sqrt{LC}$, if L is assumed to be constant, $\lambda \propto \sqrt{C}$ and $\sqrt{C}$ is proportional to the square root of the setting θ. For a uniform wave-length condenser it is necessary to have C vary as the square of the setting θ, or $C \propto \theta^2$.

Again, it may be desirable that the percentage change in capacitance for a given angle of rotation of the plates be the same for all parts of the

Straight-line wave length	Straight-line frequency	Straight-line percentage wave length or frequency
		
$C_2 = (a_2\theta + b_2)^2$ $\lambda_2 = k[(a_2\theta + b_2)^2 - \text{resid. cap.}] + K\theta$ $R_2 = \left[114.6\{2ka_2(a_2\theta + b_2) + K\}\right]^{1/2}$ Constants: $a_2 = \frac{\sqrt{\text{max. cap.}} - \sqrt{\text{resid. cap.}}}{180}$ $b_2 = \sqrt{\text{resid. cap.}}$	$C_3 = \frac{1}{(a_3\theta + b_3)^2}$ $\lambda_3 = k\left\{\frac{1}{(a_3\theta + b_3)^2} - \text{resid. cap.}\right\} + K(180 - \theta)$ $R_3 = \left[114.6\left\{\frac{2ka_3}{(a_3\theta + b_3)^3} + K\right\}\right]^{1/2}$ Constants: $a_3 = \frac{1}{180}\left\{\frac{1}{\sqrt{\text{resid. cap.}}} - b_3\right\}$ $b_3 = \frac{1}{\sqrt{\text{max. cap.}}}$	$C_4 = a_4 e^{b_4\theta}$ $\lambda_4 = k\{a_4 e^{b_4\theta} - \text{resid. cap.}\} + K\theta$ $R_4 = \left[114.6\{ka_4b_4e^{b_4\theta} + K\}\right]^{1/2}$ Constants: $a_4 = \text{resid. cap.}$ $b_4 = \frac{\log\,(\text{max. cap}) - \log\,(\text{resid. cap.})}{78.174}$

Common constants $k = \frac{\text{total plate area} - 180\,K}{\text{max. cap.} - \text{resid. cap.}}$

$K = \frac{r^2}{114.6}$

scale as in the Kolster decremeter.[1] The polar equation for the boundary curve is

$$r = \sqrt{2C_0 a \epsilon^{a\theta} + r_2^2}$$

where C = capacitance when angle $\theta = 0$
a = constant = percentage change of capacitance per scale division
ϵ = 2.71828
r_2 = radius of cutout portion to clear washers separating variable plates.

The equations and tables on page 119 have been compiled by Griffiths.[2] The four types of plates given are for equivalent condensers having a capacitance at zero setting of 36 $\mu\mu$f and a maximum of 500 $\mu\mu$f, with a plate area of 20 sq. cm.

The paper mentioned above gives the following data for the radii at different angles for the condensers mentioned in the table of equations.

θ, degrees	Radius, centimeters		
	R_2	R_3	R_4
0	2.49	8.25	1.93
5	2.56		
10	2.60	6.70	2.02
20	2.76	5.62	2.13
30	2.89	4.80	2.24
40	....	4.17	2.36
60	3.18	3.32	2.64
80	....	2.75	2.98
90	3.56		
100	....	2.37	3.38
120	3.86	2.10	3.85
140	....	1.90	4.40
150	4.12	....	4.71
160	....	1.76	5.04
170	....	....	5.40
180	4.38	1.65	5.80

26. Effect of Putting Odd-shaped Plate Condensers in Series or Parallel. If any of the above condensers are placed in parallel or in series with another condenser, the straight-line calibration will be altered. If paralleling condensers are used, the plate shape would require recalculation, after which the plate would become more nearly semicircular. If a condenser is added in series, the calculation of the plate shape is more difficult. Griffiths[3] gives complete equations for a number of series combinations, the following table applying to the cases indicated where maximum capacitance of variable condenser = 500 $\mu\mu$f, minimum capacitance of variable condenser = 36 $\mu\mu$f, series fixed capacitance = 500$\mu\mu$f, total plate area = 20 sq. cm., r = radius of inactive semicircular area of moving plate = 1.2 cm.

[1] *Bur. Standards Circ.* **74**, p. 117. *Bur. Standards Sci. Paper* 235.
[2] GRIFFITHS, W. H. F., Notes on the Laws of Variable Air Condensers, *Exp. Wireless and Wireless Eng.*, **3**, 3–14, January, 1926.
[3] GRIFFITHS, W. H. F., Further Notes on the Laws of Variable Air Condensers *Exp. Wireless and Wireless Eng.*, **3**, 743–755, December, 1926.

θ, degrees	Radius, centimeters			
	R_5	R_6	R_7	R_8
0	2.74	2.16	9.25	1.82
10	2.80	....	6.95	
20	....	2.35	5.57	1.96
30	2.92	....	4.65	
40	....	2.56	....	2.15
50	3.06			
60	....	2.78	3.32	2.38
70	3.22			
90	3.40	....	2.42	2.85
100	....	3.37		
110	3.66			
120	....	....	2.02	3.57
130	3.88			
140	....	4.25		
150	4.18	....	1.78	4.74
160	....	4.85		
170	4.52			
180	4.73	5.66	1.62	7.16

R_5, straight-line capacitance with series fixed capacitance.
R_6, corrected square law of capacitance with series fixed capacitance.
R_7, inverse square law of capacitance with series fixed capacitance.
R_8, exponential law of capacitance with series fixed capacitance.

27. Important Considerations in Design. It is not difficult to find a large number of condensers on the market which will answer the needs of any condenser application in radio receivers. The manufacture of condensers for such use has been brought to a high state of development, both electrically and mechanically. The design problems here are simpler in that low power and low voltage are to be handled.

When condensers for radio transmitters are designed, provision must be made for handling high power and high voltage. The use of very high radio frequencies has added to the problem by requiring better insulating materials. Insulators which were satisfactory at low radio frequencies have been found to heat up and be unsuited for frequencies such as 30 to 100 Mc and higher.

The following classification shows how condensers for transmitting sets could be divided with respect to the voltages to which they are subjected:
Those subjected to steady d-c voltages only.
Those subjected to l-f voltages only.
Those subjected to damped r-f voltages only (obsolete).
Those subjected to steady cw r-f voltages only.
Those subjected to modulated cw r-f voltages only.
Those subjected to d-c voltages with superimposed r-f voltage.
Those subjected to l-f voltage and superimposed r-f voltage.
The last four of the above divisions could be further subdivided into those for use on frequencies up to about 3,000 kc, those for use on frequencies from 3,000 to about 25,000 kc, and those for use on frequencies of 30,000 kc and above. The two latter classes require special construction.

In specifying the rating of condensers for use in radio transmitters the following data should be given: capacitance, current, frequency, nature of voltage to be applied. A knowledge of the maximum r-f voltage and maximum current permissible is important. A condenser should never be operated at more than half the breakdown voltage. In the case of r-f voltages this fraction should be much smaller.

28. Standards of Capacitance. Fixed condensers using the best grade of mica or fixed-air condensers are used as capacitance standards for radio frequencies. For some work a variable air condenser is essential as a standard.

An important requirement of a standard condenser is that the capacitance remain constant, the prerequisite of which is rigidity of construction, which is more difficult to secure in a variable than in a fixed condenser. There should be no relative motion possible between the movable plates and the pointer. There should be no stops against which the pointer or movable plates may strike and thus destroy the calibration. The manner of insulating the two sets of plates is of great importance, not only in fulfilling the rigidity requirement, but in minimizing the power loss. An insulating material having a low temperature coefficient of expansion should be used, so that the capacitance will not change perceptibly with temperature. As small an amount of solid insulating material as possible should be employed, keeping it well out of the electric field. This field is quite intense near the high-potential post. All insulation should be avoided in the vicinity of that terminal if power factor is to be kept low.

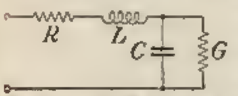


Fig. 14a.—Equivalent circuit of air condenser.

The condenser should be provided with a metal shield, which may be grounded during measurements if the capacitance is to remain constant. The leads inside the condenser should be as short and direct as possible. The resistance of leads, plates, and contacts should be kept to the minimum. Flexible connection to the moving plates should not be used in a standard.

While it has been customary and is permissible in some measurements to neglect the small resistance and inductance found in variable air condensers made for precision laboratory work, yet, as the frequency is increased to 5 Mc and higher, such omissions may result in considerable inaccuracy in the results. These small residual impedances, when taken into account, give an equivalent circuit for the variable air condenser[1] as shown in Fig. 14a, where C is the static capacitance of the condenser, R the resistance loss in the metal parts of the condenser, L the inductance of the leads and connections of stacks of plates, and G the conductance or losses in the solid dielectric parts of the condenser. The variations in these parameters with frequency and their effect upon the effective terminal capacitance for one type of laboratory condenser are treated in the paper.

High-grade mica condensers can be employed as standards after calibration as to capacitance and power factor over the range of frequencies at which they are to be used.

29. Methods of Measuring Capacitance. There are two general methods of capacitance measurement: (1) absolute measurements in terms of other electrical or physical units; (2) comparison methods, where a condenser of unknown capacitance is compared with a known calibrated condenser. The absolute methods are not carried out at radio frequencies.

[1] Field, R. F., and D. B. Sinclair, A Method for Determining the Residual Inductance and Resistance of a Variable Air Condenser at Radio Frequencies, *Proc. I.R.E.*, **24**, 255–274, February, 1936.

Approximate calibrations of condensers for r-f use can be obtained using some form of bridge operating at 1,000 cycles. A very convenient instrument for rapid checking work is found in the direct-reading microfarad meter which operates on 60-cycle current.

Condenser calibrations at radio frequencies are conveniently made by a substitution method in a resonance circuit. The standard used must be one which is constructed for use as a standard at radio frequencies. It should give the same calibration at two widely different charge and discharge rates, such as 100 and 1,000 charges and discharges per second. If it fills this requirement, it may be assumed to give the same calibration at radio frequencies.

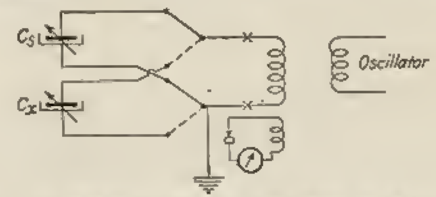


Fig. 15.—Measurement of condenser capacitance.

A simple tuned circuit consisting of a coil and the condenser under test is arranged with a double-throw switch so that the standard condenser may be readily substituted. Resonance may be indicated by a sensitive meter coupled to the main coil by a few turns of wire. A crystal detector and 1-ma d-c meter make a very convenient indicating device. Power is supplied electromagnetically by a small vacuum tube oscillator. The measurement circuit is shown in Fig. 15. The shielded side of the condenser should be grounded. It is essential that the leads connecting the switch points to each condenser be of the same length in each case as otherwise the circuits will not have the same amount of inductance when one condenser is substituted for the other, which will result in an error in the calibration. The coupling between the test circuit and the oscillators should be kept quite loose, which will be necessary if a sensitive resonance indicating instrument is used.

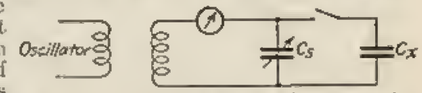


Fig. 16.—Simple scheme for measuring capacitance.

If in the circuit shown in Fig. 15 a fixed inductor is used, the calibration will be made at various frequencies depending upon the capacitance for the different condenser settings. A variable air condenser of suitable size could be connected across the coil at XX and used to keep the resonance frequency the same for any setting of C_x. If such a circuit is carefully set up, no errors will result if the two circuits connected to C_x and C_s are similar. The frequency at which the measurements are made can be measured with a frequency meter. The frequency or frequency range over which a calibration is made should always be stated.

For rougher calibration work, the circuit shown in Fig. 16 may be used where C_s is tuned both with and without C_x in the circuit. It should be noted that the leads and switch connecting C_x to the circuit will introduce errors in the calibration.

A method[1] of precision calibration of variable air condensers at a single frequency has been described in which the unknown condenser and the standard condenser are alternately made a part of the oscillator furnishing the power. The method also offers a very precise means of measuring the change in capacitance with frequency of a mica condenser.

30. Precautions in Measurement of Very Small Capacitances. It is difficult to get agreement between different laboratories in the measurement of capacitances of the order of 15 or 20 $\mu\mu f$ or less. The reasons for this are several and include differences in methods of measurement, different lengths of leads used, different sizes and spacing of leads, stray capacitances to neighboring objects, and differences of a few micromicrofarads in the capacitance standards of the various laboratories. Hence it is not unusual to find a disagreement as much as 30 per cent or more in the measurement of a capacitance of the order of 10 $\mu\mu f$.

For measurements of small capacitances it is essential to keep all connecting leads of minimum length and have them occupy definite positions, so the corrections for their inductance and capacitance can be applied if desired. Apparatus not actually needed should be kept away from the measuring circuit. A standard having a finely graduated scale is essential for such measurements. It should be capable of repeating its capacitance value for any given setting. Its capacitance curve should preferably be a straight line without any crooks in it, so that interpolations can be accurately made from calibrated points.

31. Methods of Measuring Condenser Resistance and Power Factor and Dielectric Constant of Insulating Materials at Radio Frequencies. Measurements of condenser resistance and power factor of insulating materials are made in practically the same manner as the sample of insulating material is prepared so as to form a condenser. Methods of measuring condenser resistance[2] and power factor of insulating materials[3] have been given in publications of the Bureau of Standards. The American Society for Testing Materials has one or more standard methods of testing electrical insulating materials for power factor and dielectric constant.[4]

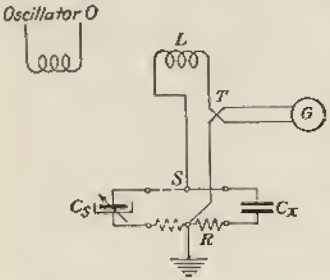


Fig. 17.—Circuit for measuring properties of insulators.

The circuit shown in Fig. 17 may be used for measurements of resistance, power factor, and dielectric constant. Assuming that the power factor of a sample of insulating material is to be measured, the sample in sheet form is made into a condenser of capacitance between 100 and 1,000 $\mu\mu f$, as represented by C_x (Fig. 17). The remainder of the circuit consists of the coil L, thermoelement T, and double-pole, double-throw switch S, in which resistors R may be inserted. The galvanometer G gives deflections which are proportional to the square of the current flowing in the circuit LTC_xR, as electromagnetically induced from the r-f oscillator O.

[1] Hall, E. L., and W. D. George, Precision Condenser Calibration at Radio Frequencies, *Electronics*, **7**, 318–320, 1934.

[2] Radio Instruments and Measurements, *Bur. Standards Circ.* 74, pp. 190–193.

[3] Methods of Measurement of Properties of Electrical Insulating Materials, *Bur. Standards Sci. Paper* 471.

[4] Tentative Methods of Test for Power Factor and Dielectric Constant of Electrical Insulating Materials, designation D150–39T; Tentative Method of Test for Power Factor and Dielectric Constant of Natural Mica, designation D351–39T; American Society for Testing Materials, 260 South Broad Street, Philadelphia, Pa.

The deflections of galvanometer G are noted for several values of inserted resistance R and for the case when R is a link of practically zero resistance. Using the "zero resistance" deflection and the deflection for a known value r of resistance inserted in switch S, the resistance R_T of the total circuit LTC_xR is given by

$$R_T = \frac{r}{\sqrt{\frac{d_0}{d_1}} - 1}$$

The average of the values of R_T calculated for various values of r should be taken as the resistance of the complete circuit. The resistance R_S of the circuit when C_S is substituted for C_X should be obtained in the same manner. The resistance R_x of the condenser C_x is then given by $R_x = R_T - R_S$. It is essential for this measurement that the two parts of the circuit which are interchanged should be as nearly identical as possible.

After the resistance R_x of the insulating material condenser is obtained, the power factor or phase difference can be calculated from the equations given above. The dielectric constant K can be calculated from the equation $K = Ct/0.0885S$, where C = capacitance of sample in micromicrofarads, t = thickness of sample in centimeters, and S = area of smaller plate in square centimeters. The capacitance is known, as given by C_S, and the area of one plate and the thickness of the sample can easily be measured.

The method described above operates satisfactorily at frequencies from 100 to 1,500 kc.

A bridge method is sometimes used for these measurements although the apparatus is considerably more complicated than that described above.

A comparative method for testing insulating materials at very high radio frequencies has been used by certain laboratories. In this method the insulating material sample is placed in an intense electric field produced by a 30-megacycle transmitter and the temperature rise in the sample measured for a definite time interval. While such results have not as yet been definitely tied up with power factor, dielectric constant, etc., yet they represent in a very practical manner a means for determining the suitability of different types of materials for use at very high radio frequencies. An insulator which is entirely satisfactory at lower radio frequencies such as 1,000 or 2,000 kc may prove to be unusable at 20 or 30 megacycles. Hence data on power factor and dielectric constant are meaningless without a statement of the frequency at which the data were obtained.

Some of the German technical periodicals[1] have reported the production of improved ceramic insulators in Germany. One type of material is claimed to have extremely low power loss at very high frequencies. Another type of material having moderate power loss possesses very high values of dielectric constant which can be made to have values

[1] Calit and Calan, Zwei neue hochwertige Isolierstoffe der Hochfrequenztechnik, *Hochfrequenztechnik und Elektroakustik*, **43**, 33, 34, January, 1934; Handrek, H., Neue Hochfrequenz-Isolierstoffe, *Hochfrequenztechnik und Elektroakustik*, **43**, 73–75, March, 1934.

as high as 100. The latter material would appear to have advantages in condenser manufacture for use at ultrahigh frequencies where very small parts and extremely short connections are required. These materials have several names and differ in their properties. The names are Calit, Ultra-Calit, Calan, Ultra-Calan, Frequentit, Frequenta, Condensa, and Condensa C. The last two materials have the high dielectric constants, and the ones with the prefix "Ultra" have very low losses and are intended for u-h-f work.

SECTION 6

COMBINED CIRCUITS OF L, C, AND R

BY W. F. LANTERMAN,[1] B.S.

1. Transient and Steady-state Currents. When a voltage is suddenly applied to a circuit, the current assumes a transient state for a brief interval, then gradually settles down to a steady-state condition which it maintains until the voltage is interrupted or changed. Relations for computing transient and steady-state currents in *LCR* circuits are given in the following paragraphs.

TRANSIENT CURRENTS IN LCR CIRCUITS

2. Symbols Used in Transient Expressions. In the transient expressions given in Arts. 3, 4, and 5, the following symbols will be used:

- L Inductance in henrys.
- C Capacitance in farads.
- R Resistance in ohms.
- T Time constants in second; time in seconds for current or voltage to reach $1/\epsilon$ or approximately 33 per cent of its initial value if decreasing; or $\left(1 - \frac{1}{\epsilon}\right)$, or approximately 67 per cent of its final value if increasing.
- i Instantaneous current in amperes at time t.
- e Instantaneous voltage in volts at time t.
- t Time in seconds after starting.
- I Steady-state d.c. in amperes.
- E Maximum value of a-c voltage in volts.
- V Steady-state d-c voltage in volts.
- Q Condenser charge in coulombs.
- Z A-c impedance in ohms.

$$= \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \text{ for } LCR \text{ circuit.}$$

$$= \sqrt{R^2 + (\omega L)^2} \text{ for } LR \text{ circuit.}$$

$$= \sqrt{R^2 + \left(\frac{1}{\omega C}\right)^2} \text{ for } RC \text{ circuit.}$$

- f Frequency of applied a.c. in cycles per second.
- ω Angular velocity of applied a.c. in radians per second $= 2\pi f$.
- f_1 Natural frequency of oscillatory circuit *LCR* in cycles per second.
- ω_1 Natural angular velocity of oscillatory circuit *LCR* in radians per sec. $= 2\pi f_1$.
- W_R Energy in joules dissipated in R during transient state.
- W_L Energy in joules stored by or lost by L during transient state.
- W_C Energy in joules stored by or lost by C during transient state.
- θ Phase angle of a-c voltage at $t = 0$, *i.e.*, when the switch is closed.
- ϕ Phase angle of impedance as defined for each case.
- ϵ 2.718 (base of natural logarithms).
- α and β (Defined in Art. 5a).

3. RC Circuit Transients.

NOTE. The following formulas for i in *RC* circuits are not true for extremely small values of t. For very small t, the L of the circuit, no

[1] National Broadcasting Co., Inc., Chicago, Ill.

matter how small, limits i, and the relations of Art. 5 for LCR circuits must be applied. This is especially important for short pulses or high frequencies, where small values of t are involved.

a. D-c Voltage V Suddenly Applied to Deenergized RC.

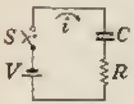


FIG. 1.

$$i = \frac{V}{R}\epsilon^{-\frac{t}{RC}}$$
$$T = RC$$
$$W_R = W_C = \frac{1}{2}CV^2 = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV$$

for fully charged C ($t = \infty$).

FIG. 2.

b. C Charged to Voltage V and Suddenly Discharged through R.

FIG. 3.

$$i = -\frac{V}{R}\epsilon^{-\frac{t}{RC}}$$
$$T = RC$$
$$W_R = \frac{1}{2}CV^2 = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV$$

for complete discharge of C ($t = \infty$).

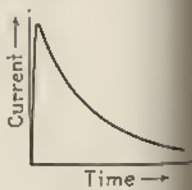


FIG. 4.

c. A-c Voltage e Suddenly Applied to Deenergized RC.

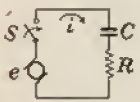


FIG. 5.

Applied voltage $e = E \sin (\omega t + \theta)$

$$i^* = \underline{\frac{E}{Z}\sin(\omega t + \theta - \phi)} - \frac{E}{Z}\cos(\theta - \phi)\frac{\epsilon^{-\frac{t}{RC}}}{\omega RC}$$
$$\phi = \cot^{-1}(\omega RC)$$

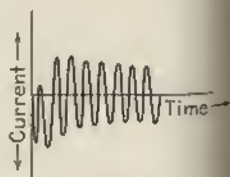


FIG. 6.

4. LR Circuit Transients. *a. D-c Voltage V Suddenly Applied to Deenergized LR.*

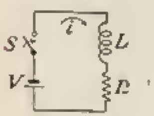


FIG. 7.

$$i^* = \frac{V}{R}(\underline{1} - \epsilon^{-\frac{Rt}{L}})$$
$$T = \frac{L}{R}$$
$$W_L = \frac{1}{2}LI^2 = \frac{1}{2}L\frac{V^2}{R^2}$$

for $t = \infty$.

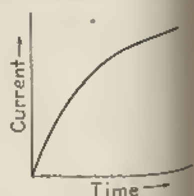


FIG. 8.

* Underscored terms represent steady-state currents; remaining term or terms are the transients.

b. LR Carrying Steady Direct Current I Suddenly Interrupted.

Steady current $I = \frac{V}{R}$

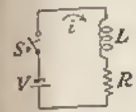


FIG. 9.

$$i = \frac{V}{R}\epsilon^{-\frac{Rt}{L}} = I\epsilon^{-\frac{Rt}{L}}$$
$$T = \frac{L}{R}$$
$$W_L = \frac{1}{2}LI^2 = \frac{1}{2}L\frac{V^2}{R^2}$$

for $t = \infty$.

FIG. 10.

c. A-c Voltage e Suddenly Applied to Deenergized LR.

Applied voltage $e = E \sin (\omega t + \theta)$

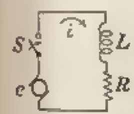


FIG. 11.

$$i^* = \underline{\frac{E}{Z}\sin(\omega t + \theta - \phi)} - \frac{E}{Z}\sin(\theta - \phi)\epsilon^{-\frac{Rt}{L}}$$
$$\phi = \tan^{-1}\left(\frac{\omega L}{R}\right)$$

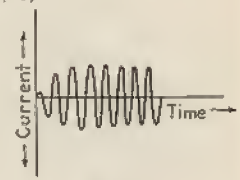


FIG. 12.

5. LCR Circuit Transients. *a. D-c Voltage V Suddenly Applied to Deenergized LCR.*

General Solution:

$$i = \frac{V}{2\beta L}\epsilon^{-\alpha t}(\epsilon^{\beta t} - \epsilon^{-\beta t}) = \frac{V}{\beta L}\epsilon^{-\alpha t}\sinh \beta t$$
$$\alpha = \frac{R}{2L} \qquad \beta = \sqrt{\alpha^2 - \frac{1}{LC}} = \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}$$

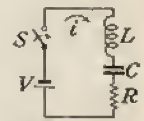


FIG. 13.

There are three special cases, depending upon the ratio $\alpha = R/2L$:

Case I. Aperiodic current, when $\alpha^2 = \frac{R^2}{4L^2} > \frac{1}{LC}$. ($\beta$ is real.)

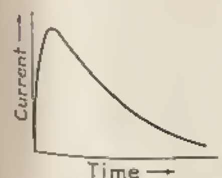


FIG. 14.

$$i = \frac{V}{R}\epsilon^{-\frac{t}{RC}} - \frac{V}{R}\epsilon^{-\frac{Rt}{L}}$$

i builds up to a maximum at

$$t = \left(\frac{1}{2\beta}\log_\epsilon \frac{\alpha + \beta}{\alpha - \beta}\right) \text{ sec.}$$

then slowly decays to zero.

* Underscored terms represent steady-state currents; remaining term or terms are the transients.

Case II. Critical damping, when $\alpha^2 = \frac{R^2}{4L^2} = \frac{1}{LC}$. ($\beta = 0$.)

$$i = \frac{V}{L}t\epsilon^{-\frac{Rt}{L}}$$

i builds up to a maximum at $t = 1/\alpha$ sec., then slowly decays to zero.

FIG. 15.

Case III. Oscillatory current, when $\alpha^2 = \frac{R^2}{4L^2} < \frac{1}{LC}$. ($\beta$ is imaginary.)

$$i = \frac{V}{\omega_1 L}\epsilon^{-\frac{Rt}{2L}} \sin \omega_1 t$$

$$\omega_1 = \sqrt{\frac{1}{LC} - \alpha^2} = 2\pi f_1$$

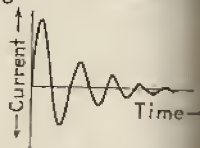


FIG. 16.

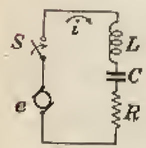
i builds up to a maximum at $t = 1/4f_1$ sec., then oscillates with amplitude slowly decreasing to zero.

For approximations

$$i = V\sqrt{\frac{C}{L}}\epsilon^{-\frac{Rt}{2L}}$$

$$\omega_1 = \frac{1}{\sqrt{LC}} \qquad f_1 = \frac{1}{2\pi\sqrt{LC}}$$

b. A-c Voltage e Suddenly Applied to Deenergized LCR.



Applied voltage $e = E \sin (\omega t + \theta)$

There are three special cases, depending upon the ratio

$$\alpha = \frac{R}{2L}$$

Case I. Aperiodic current, when $\alpha^2 = R^2/4L^2 > 1/LC$. ($\beta$ is real.)

FIG. 17.

$$i^* = \underline{\frac{E}{Z}\sin(\omega t + \theta - \phi)} + \frac{E}{Z}\frac{1}{2\beta L}\left[L(\alpha-\beta)\sin(\theta-\phi) - \frac{1}{\omega C}\cos(\theta-\phi)\right]\epsilon^{-(\alpha-\beta)t} - \frac{E}{Z}\frac{1}{2\beta L}\left[L(\alpha+\beta)\sin(\theta-\phi) - \frac{1}{\omega C}\cos(\theta-\phi)\right]\epsilon^{-(\alpha+\beta)t}$$

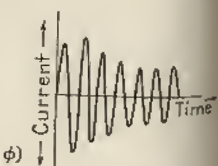


FIG. 18.

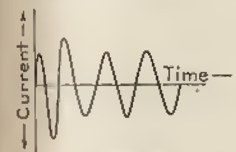
* Underscored terms represent steady-state currents; remaining term or terms are the transients.

Case II. Critical damping, when $\alpha^2 = R^2/4L^2 = 1/LC$. ($\beta = 0$.)

$$i^* = \underline{\frac{E}{Z}\sin(\omega t + \theta - \phi)} + \frac{E}{Z}\left[(\alpha t - 1)\sin(\theta-\phi) - \frac{1}{L\omega C}\cos(\theta-\phi)\right]\epsilon^{-\alpha t}$$

FIG. 19.

Case III. Oscillatory current, when $\alpha^2 = R^2/4L^2 < 1/LC$. ($\beta$ is imaginary.)



$$i^* = \underline{\frac{E}{Z}\sin(\omega t + \theta - \phi)} - \frac{E}{Z}\epsilon^{-\alpha t}\left[\sin(\theta-\phi)\cos\omega_1 t + \frac{1}{\omega_1 L\omega C}\cos(\theta-\phi)\sin\omega_1 t - \frac{R}{2\omega_1 L}\sin(\theta-\phi)\sin\omega_1 t\right]$$

FIG. 20.

STEADY-STATE CURRENTS IN LCR CIRCUITS

6. Impedance Relations. Steady-state currents are calculated by an expression similar to Ohm's law,

$$\mathbf{I} = \frac{\mathbf{E}}{\mathbf{Z}} \tag{1}$$

where **I** and **E** are vectors representing r-m-s values of current in amperes and voltage in volts. The impedance Z (expressed in ohms) is the vector sum of the a-c resistance R and the reactance X,

$$Z = R + jX \tag{2}$$

The factor j is an operator to indicate that X is 90 deg. out of phase with R. The magnitude of Z is $|Z| = \sqrt{R^2 + X^2}$, and its phase angle with R is

$$\phi = \tan^{-1}\frac{X}{R} = \cos^{-1}\frac{R}{Z} \tag{3}$$

7. Values of the Reactance X of Coils and Condensers. The X or reactance component of impedance is due to inductance or capacitance. If the reactance is a coil having an inductance of L henrys,

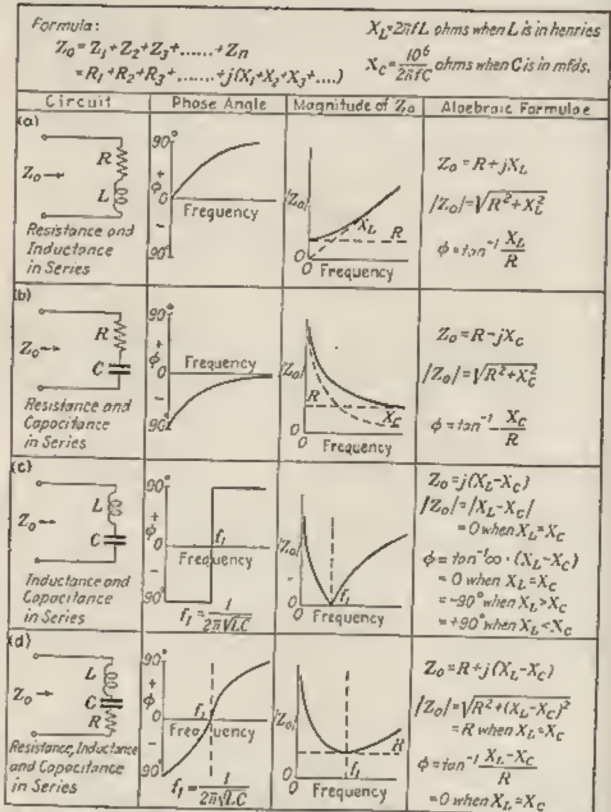
$$X_L = \omega L = 2\pi f L \text{ ohms}$$

where $\omega = 2\pi f$ and f = frequency in cycles per second; if it is a condenser of capacitance C farads, $X_C = -1/\omega C$ ohms; if it is composed of both L and C, $X = \left(\omega L - \frac{1}{\omega C}\right)$ ohms. Capacitance always has negative

* Underscored terms represent steady-state currents; remaining term or terms are the transients.

reactance, and inductance always has positive reactance. Thus Eq. (2) may be written

$$Z = R + j\left(\omega L - \frac{1}{\omega C}\right) \quad (4)$$



Equivalent impedances of series combinations of L, C, and R.

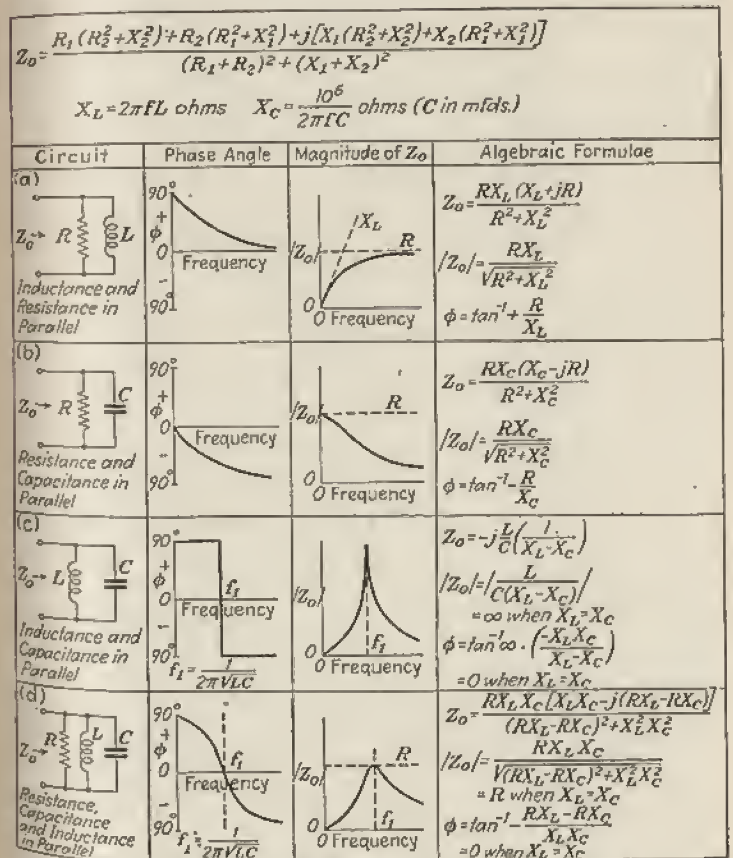
8. Equivalent or Total Impedance. Any network of impedances can be reduced to an equivalent impedance of the form $Z = R + jX$ by the following formulas:

1. *Impedances Z_1 and Z_2 in series:*

$$Z_0 = Z_1 + Z_2 = (R_1 + jX_1) + (R_2 + jX_2) = (R_1 + R_2) + j(X_1 + X_2) \quad (5)$$

2. *Impedances Z_1 and Z_2 in parallel:*

$$Z_0 = \frac{Z_1 Z_2}{Z_1 + Z_2} = \frac{(R_1 + jX_1)(R_2 + jX_2)}{(R_1 + jX_1) + (R_2 + jX_2)}$$
$$= \frac{R_1(R_2^2 + X_2^2) + R_2(R_1^2 + X_1^2) + j[X_1(R_2^2 + X_2^2) + X_2(R_1^2 + X_1^2)]}{(R_1 + R_2)^2 + (X_1 + X_2)^2} \quad (6)$$



Equivalent impedances of parallel combinations of L, C, and R.

This expression, while somewhat involved, is seen still to be of the form $Z = R + jX$. Charts showing impedance relations for some common LCR circuits are shown on these pages.

3. *Impedances in network.* By applying the foregoing formul[as] in a step-by-step process, any network can be reduced to a single equiva[lent] impedance. Thus, in Fig. 21,

FIG. 21.—Network with branch impedances.

$$Z_{34} = \frac{Z_3 Z_4}{Z_3 + Z_4}; \qquad Z_{234} = Z_2 + Z_{34}$$

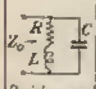
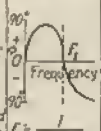
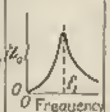
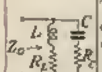
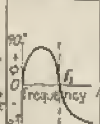
and finally,

$$Z_0 = \frac{Z_1(Z_2 + Z_{34})}{Z_1 + Z_2 + Z_{34}}$$

In a complex network, however, the number of terms will be large, an[d] the computations will be laborious.

$$Z_0 = \frac{R_1(R_2^2+X_2^2)+R_2(R_1^2+X_1^2)+j[X_1(R_2^2+X_2^2)+X_2(R_1^2+X_1^2)]}{(R_1+R_2)^2+(X_1+X_2)^2}$$

$X_L = 2\pi f L$ ohms $\qquad X_C = \dfrac{10^6}{2\pi f C}$ ohms (C in mfds)

Circuit	Phase Angle	Magnitude of Z_0	Algebraic Formulae
(e)  Resistance and Inductance in Series and in Parallel with Capacitance	 $f_1 \doteq \dfrac{1}{2\pi\sqrt{LC}}$ Max. + ϕ angle occurs when $f = 0.577 f_1$, $Q > 1$. Not valid for $Q \leq 1$		$Z_0 = \dfrac{RX_C^2 - j[\frac{L}{C}(X_L - X_C) + R^2X_C]}{R^2 + (X_L - X_C)^2}$ $\lvert Z_0 \rvert = X_C\sqrt{\dfrac{R^2 + X_L^2}{R^2 + (X_L - X_C)^2}}$ $\doteq \dfrac{L}{RC}$ when $X_L = X_C$ and R is small $\phi = \tan^{-1}\left(-\dfrac{X_L(X_L - X_C) + R^2}{RX_C}\right)$ $\doteq 0$ when $X_L = X_C$ and R is small
(f)  Resistance and Inductance in Series and in Parallel with Capacitance and Resistance in Series	 $f_1 \doteq \dfrac{1}{2\pi\sqrt{LC}}$ $Q > 1$ Not valid for $Q \leq 1$		$Z_0 = \dfrac{(R_LA + R_CB) + j(X_LA - X_CB)}{(R_L + R_C)^2 + (X_L - X_C)^2}$ $\lvert Z_0 \rvert = \dfrac{\sqrt{(R_LA + R_CB)^2 + (X_LA - X_CB)^2}}{(R_L + R_C)^2 + (X_L - X_C)^2}$ $A = (R_C^2 + X_C^2)$, $B = (R_L^2 + X_L^2)$ $\doteq \dfrac{L}{(R_L + R_C)C}$ when $X_L = X_C$ and R_L and R_C are small $\phi = \tan^{-1}\dfrac{X_LA - X_CB}{R_LA + R_CB}$ $\doteq 0$ when $X_L = X_C$ and R_L and R_C are small

Equivalent impedances of parallel combinations of L, C, and R.

9. Q of LCR Circuits. Every coil or condenser has some energ[y] losses, which at a given frequency may be represented as an equivalen[t] R in series with the reactance X_L or X_C. The ratios

$$Q_L = \frac{X_L}{R_L} = \frac{\omega L}{R_L} \quad \text{and} \quad Q_C = \frac{X_C}{R_C} = \frac{1}{\omega C R_C}$$

define Q, which is a figure of merit for the coil or condenser. As a general expression for Q in any circuit, we have

$$Q = \frac{\text{volt-amperes}}{\text{watts dissipated}} = \frac{VA}{W_d}$$

Although Q varies with frequency, it is nearly constant over narrow ranges of frequency, and its use therefore renders circuit computations somewhat simpler when losses have to be considered. Table I gives some representative values of Q which may be expected from ordinary coils and condensers.

TABLE I.—REPRESENTATIVE VALUES OF Q FOR VARIOUS COILS AND CONDENSERS

Frequency, cycles	Coils with powdered iron cores	Air-cored coils	Condensers with paper dielectric	Condensers with mica dielectric
100	25 to 50	3 to 10	1,000	
1,000	50 to 75	25 to 50	500	3,000
10,000	100 to 150	100 to 300	100 to 200	500
100,000	150 to 200	100 to 300	50 to 100	200 to 300
1,000,000	100 to 200	100 to 300	....	50 to 200

The following data are quoted from Franks:[1]

Item	Frequency, kilocycles	Q	Item	Frequency, kilocycles	Q
100 μμf molded bakelite fixed condenser........	1,000	40	Broadcast band bank-wound litz solenoid ¾ in. diam. in 1⅝ in. square shield can.......	1,000	110
Typical gang condenser; bakelite stator insulation	100 1,000 10,000	2,000 700 200	Broadcast band universal-wound litz coil with iron core in same can.......	1,000	185
Same with ceramic stator insulation.	100 1,000 10,000	8,000 3,000 1,000	Transmitter coil, 4½ in. diam. and 5 in. long, 11 turns of ¼ in. copper tubing...............	5,000	650
Single-section litz-wound universal coil; 456-kc intermediate frequency, in can..................	456	80	Transmitter coil 1¾ in. diameter and 1¾ in. long, 12 turns of No. 10 wire...............	10,000	400
Same but with powdered-iron core..............	456	145	Receiver coil, 1 in. diameter and ⅝ in. long, 5 turns of No. 14........	30,000	270

10. Loss Due to Inserting Series or Shunt Impedance in Audio Circuits. In audio circuits, attenuation-frequency characteristics are often purposely modified by the insertion of corrective impedances such as equalizers, "tone controls," and scratch filters. The following formulas give the insertion losses in such cases:

1. *Shunt Impedance.* The loss due to inserting a shunt impedance Z_s (Fig. 22a and b) is

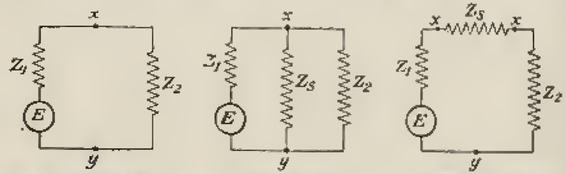
$$L = 20 \log_{10}\left(1 + \frac{Z_1 Z_2}{Z_s(Z_1 + Z_2)}\right) \text{ db} \tag{8}$$

The shunting impedance can usually be located at a point in the circuit where the impedances Z_1 and Z_2 are matched, and where each is substantially a pure resistance through the range of frequencies involved. Then, letting $Z_1 = Z_2 = R_0$, the loss is

[1] FRANKS, C. J., *Electronics*, p. 126, April, 1935.

$$L = 20 \log_{10} \left| \frac{2Z_s + R_0}{2Z_s} \right| \text{ db} = 20 \log_{10} \sqrt{1 + \frac{\cos \phi}{K} + \frac{1}{4K^2}} \text{ db} \quad (8a)$$

where $K = |Z_s|/R_0$ and ϕ is the phase angle of Z_s. For various values of K and ϕ the loss can be read from the curve (Fig. 23).



(a)-Circuit with Applied Voltage, E, Impedance of Source, Z_1, and Load Impedance, Z_2 (b)-Shunt Impedance Inserted (c)-Series Impedance Inserted

FIG. 22.—Shunt and series impedances inserted in audio-frequency circuits.

2. *Series Impedance.* The loss in decibels due to inserting a series impedance Z_s (Fig. 22*a* and *c*) is

$$L = 20 \log_{10} \left(\frac{Z_1 + Z_2 + Z_s}{Z_1 + Z_2} \right) \text{ db} \quad (9)$$

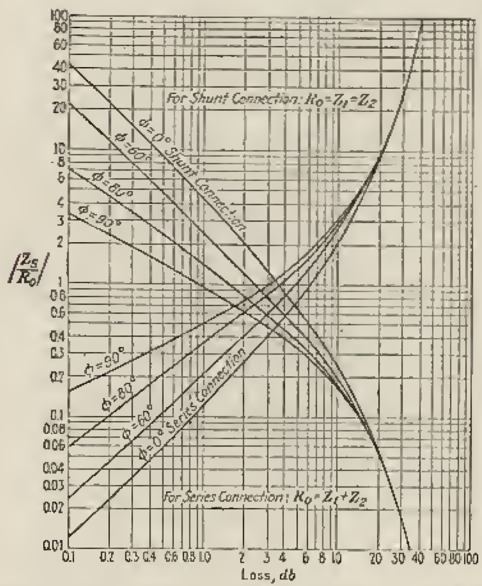


FIG. 23.—Transmission loss due to insertion of shunt or series impedance.

The series impedance can usually be inserted at a point in the circuit where the impedances Z_1 and Z_2 are matched and where each is substantially a pure

resistance through the range of frequencies involved. Then, letting $Z_1 + Z_2 = R_0$, the loss is

$$L = 20 \log_{10} \left| \frac{R_0 + Z_s}{R_0} \right| \text{ db}$$
$$= 20 \log_{10} \sqrt{1 + 2K \cos \phi + K^2} \text{ db} \quad (9a)$$

where $K = |Z_s|/R_0$ and ϕ is the phase angle of Z_s. The loss can be read from Fig. 23 for various values of K and ϕ.

SERIES RESONANCE

11. Definition. A series circuit containing LCR, such as that in Fig. 24, is *series resonant* when the line current $\mathbf{I}_0$ is in phase with the line voltage $\mathbf{E}_0$.

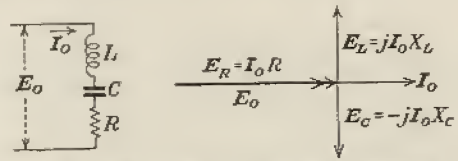


FIG. 24.—Series circuit and vector representing it.

12. Conditions for Series Resonance. The equivalent impedance of a series circuit is, by Eq. (4), $Z_0 = R + j(X_L - X_C)$. The total voltage drop in the circuit is $\mathbf{E}_0 = \mathbf{I}_0[R + j(X_L - X_C)]$. Resonance occurs when $X_L = X_C$. At resonance the total reactance is $X_L - X_C = 0$, and the current is $\mathbf{I}_0 = \mathbf{E}_0/R$.

There is only one frequency at which $X_L = X_C$; this gives the formula for resonant frequency:

$$f_r = \frac{1}{2\pi\sqrt{LC}} \quad (10)$$

where f_r is the resonant frequency in cycles per second, L is in henrys, C is in farads. The manner in which the reactances vary with frequency is shown in Fig. 25.

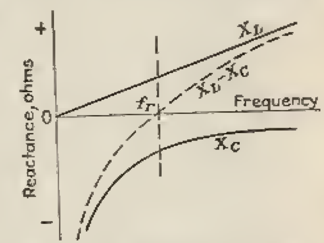


FIG. 25.—Series circuit reactance.

13. Properties of a Series Resonant Circuit. A series resonant circuit has the following properties at resonance: (1) the current is maximum; (2) the impedance is minimum; (3) the current is in phase with the impressed voltage; (4) the current is limited only by the total resistance, which is usually equal to the coil resistance; (5) if the coil resistance is the only resistance in the circuit, the voltage drop across the coil is greater than the impressed voltage; and (6) the voltage drop across the condenser exceeds the impressed voltage if X_C is greater than R.

Items 5 and 6 are of importance in practice because such high voltages may develop across C and L as to endanger their insulation, unless this is provided for in their design.

14. Impedance of Series Resonant Circuits. At resonance, the impedance is R, the total series resistance of the circuit. At any other frequency, f_1, there is also a reactance component

$$X_1 = 2\pi L\left(\frac{f_1^2 - f_r^2}{f_1}\right) \tag{11}$$

and the total impedance is

$$Z_1 = R + j2\pi L\left(\frac{f_1^2 - f_r^2}{f_1}\right) \tag{12}$$

The magnitude of Z_1 is

$$|Z_1| = \sqrt{R^2 + 4\pi^2L^2\left(\frac{f_1^2 - f_r^2}{f_1}\right)^2} \tag{13}$$

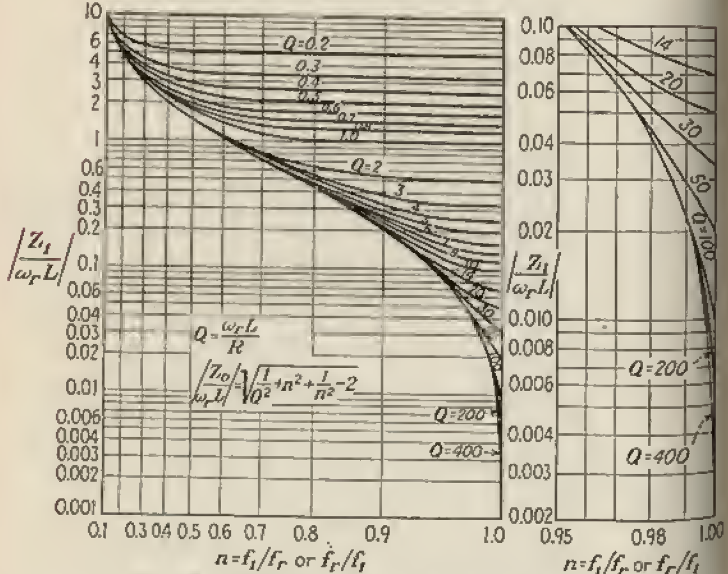


FIG. 26.—Design chart for series resonant circuit.

Circuit Constants Table. A table of LC products and other constants frequently used in calculations of resonant circuits and coil and condenser reactances is given in Sec. 1.

15. Series Resonance Curves—Z_0 and ϕ. The following useful relations for series resonant circuits are derived from Eq. (13):

$$\frac{|Z_1|}{\omega_r L} = \sqrt{\frac{1}{Q_r^2} + n^2 + \frac{1}{n^2} - 2} \tag{14}$$

and

$$L = \frac{1}{\omega_r}\sqrt{\frac{|Z_1|^2 - R^2}{\left(n^2 + \frac{1}{n^2} - 2\right)}} \tag{15}$$

where $|Z_1|$ = the absolute impedance at any frequency f_1
$\omega_r = 2\pi \times$ resonance frequency, f_r
$\omega_1 = 2\pi \times$ any other frequency, f_1
$n = \omega_1/\omega_r = f_1/f_r$ when $f_1 < f_r$; or $n = \omega_r/\omega_1 = f_r/f_1$ when $f_1 > f_r$
$Q_r = \omega_r L/R = Q$ at resonance.

The phase angle of Z_1 is

$$\phi = \tan^{-1}\frac{X_r}{R} = \tan^{-1} Q_r\left(n - \frac{1}{n}\right) \tag{16}$$

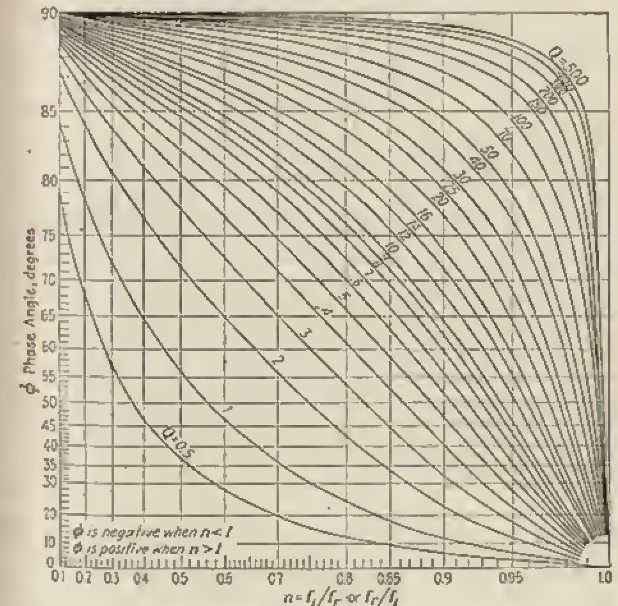


FIG. 27.—Phase angles in terms of n and Q, series circuits.

$$\phi = \tan^{-1}\left[Q\left(n - \frac{1}{n}\right)\right]$$

Using Eq. (14), the universal resonance curves of Fig. 26 were plotted in terms of $|Z_1|/\omega_r L$, Q, and n. Similar curves for ϕ in terms of Q and n are given in Fig. 27. In these curves the ratio n is to be taken as either f_1/f_r or f_r/f_1, whichever gives $n < 1$. From these universal curves, complete information about impedance and phase angle of a series resonant circuit can be read directly when the constants ω_r, L and Q or R are known.

16. Design of Series Resonant Circuit. To design a series resonant circuit, we have to determine values of L, C, and R to meet a given set of conditions. The given conditions must include values for three items: (1) resonance frequency, (2) impedance at resonance, and (3) impedance at some other frequency; otherwise there is no unique design.

Example. Assume, for example, that a series resonant circuit is to have an impedance of 100 ohms at a resonance frequency of 1,000 cycles, and an impedance of 500 ohms at 900 cycles. Then $\omega_r = 2\pi f_r = 6{,}280$. At resonance, *i.e.*, for $n = 1$, $|Z_1'| = R = 100$, and at 900 cycles, $n = 0.9$, $|Z_1| = 500$. Then, substituting in Eq. (15), we have

$$L = \frac{1}{6{,}280}\sqrt{\frac{500^2 - 100^2}{0.81 + 1.235 - 2}} = 0.369 \text{ henry}$$

By reference to the LC table, Sec. 1, page 9, we find that at $f_r = 1{,}000$, $LC = 2.533 \times 10^{-8}$, from which $C = LC/L = 0.0687 \times 10^{-6}$ farad. Thus we have $R = 100$ ohms, $L = 0.369$ henry, and $C = 0.0687 \times 10^{-6}$ farad as the constants of the circuit. Also,

$$Q = \frac{\omega_r L}{R} = \frac{6{,}280 \times 0.369}{100} = 23$$

and from this information we can select from Figs. 26 and 27 (by interpolation for $Q = 23$) the curves giving impedance and phase angles for frequencies above and below resonance.

17. Use of Series Resonant Circuit for Frequency Regulation. An application of a series resonant circuit is shown in Fig. 28. At resonance the excitation voltages applied to the grids are the reactance drops IX_C

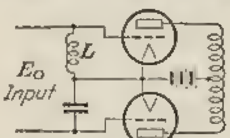


FIG. 28.—Use of series resonance circuit for frequency regulation.

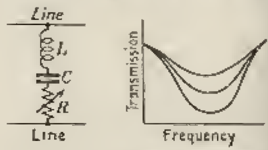


FIG. 29.—Series resonant equalizer.

and IX_L. The tubes are biased to the cutoff point so that rectification takes place. As long as the frequency of the applied voltage E_o is $f = 1/2\pi\sqrt{LC}$, the excitation voltages and therefore the plate currents of the two tubes will be equal, but if the frequency varies, the voltage drop across one reactance will increase and that across the other will decrease, causing the plate current of one tube to exceed the other. This difference in plate currents may be read on a meter to indicate the frequency of applied voltage, or it may be utilized through a differential relay to operate an automatic frequency controlling device.

17a. Series Resonant Circuits as Equalizers. Series resonant circuits are often used as equalizers where it is required to eliminate or attenuate a certain frequency or a small band of frequencies. The resonant circuit with a variable resistance in series is connected in shunt across the line

or terminals of the circuit to be equalized, and more or less readily by-passes currents of the resonant and adjacent frequencies, depending upon the adjustment of R (see Fig. 29).

18. Scratch Filters. A series resonant circuit is the simplest form of scratch-and-hiss filter for electric phonographs or carbon microphones. The resonance frequency is usually about 4,500 cycles; a typical filter is shown in Fig. 30 with its loss *versus* frequency characteristic. A low-pass filter with 5,000-cycle cutoff is much better for the purpose, however (see Art. 65).

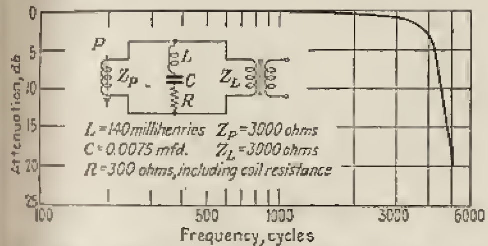


FIG. 30.—Transmission characteristic of scratch filter used with magnetic phonograph pickup.

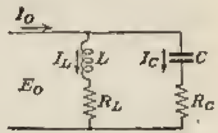


FIG. 31.—Parallel resonance.

19. Tone Control. A series circuit resonant at about 1,000 cycles is sometimes used as tone control in an a-f amplifier. It may have a variable resistance and be connected in shunt in a grid or plate circuit, or it may be shunted across part of a volume control. Such a tone control tends to compensate automatically for the frequency characteristic of the ear, which varies with sound volume.

PARALLEL RESONANCE

20. Definition. A parallel circuit containing L in one branch and C in the other (Fig. 31) is *parallel resonant* when the line current $\mathbf{I}_0$ is in phase with the line voltage $\mathbf{E}_0$. In this case there are two resistances to be considered. R_L is the resistance in the coil branch, and R_C the resistance in the condenser branch. The latter is usually small and often negligible as compared to R_L.

In some textbooks parallel resonance is defined as the condition of minimum $\mathbf{I}_0$ (or maximum Z_0, which is the same thing). On this basis a slightly different resonance frequency is obtained, depending upon whether L, C, or f is varied. For practical purposes, however, the difference is small enough to be neglected, and the results may be considered as being essentially identical, especially in view of the fact that nearly all tuned circuits require at least one variable reactance by which final tuning adjustments may be made on actual test.

21. Conditions for Parallel Resonance. The equivalent impedance of the parallel circuit of Fig. 31 is

$$Z_0 = \frac{R_L(R_C^2 + X_C^2) + R_C(R_L^2 + X_L^2) + j[X_L(R_C^2 + X_C^2) - X_C(R_L^2 + X_L^2)]}{(R_L + R_C)^2 + (X_L - X_C)^2} \tag{17}$$

where
$$X_L = 2\pi fL$$
$$X_C = \frac{1}{2\pi fC}.$$

$\mathbf{I}_0$ will be in phase with $\mathbf{E}_0$ when the j term equals zero,

$$X_0 = \frac{X_L(R_C^2 + X_C^2) - X_C(R_L^2 + X_L^2)}{(R_L + R_C)^2 + (X_L - X_C)^2} = 0 \qquad (18)$$

This condition exists if

$$\omega_r = \frac{1}{\sqrt{LC}}\sqrt{\frac{L - R_L^2C}{L - R_C^2C}} \quad \text{or} \quad f_r = \frac{1}{2\pi\sqrt{LC}}\sqrt{\frac{L - R_L^2C}{L - R_C^2C}} \qquad (19)$$

22. Approximate Formulas for Resonance Frequency When R_L and R_C Are Small or Equal. If the resistance R_C is negligible, Eq. (19) becomes

$$f_r \doteq \frac{1}{2\pi\sqrt{LC}}\sqrt{1 - \frac{R_L^2C}{L}} \qquad (20)$$

If $R_L = R_C$ or if R_L and R_C are both very small,

$$f_r \doteq \frac{1}{2\pi\sqrt{LC}} \qquad (21)$$

which is the same as the resonant frequency of a series circuit [Art. 1…, Eq. (10)].

23. Special Case Where $R_L = R_C = \sqrt{L/C}$. In this case Eq. (17) reduces to

$$Z_0 = R \qquad (22)$$

and the circuit is resonant at all frequencies with constant impedance equal to R. This special case is not so useful as might be expected, however, since L/C is usually large (on the order of 10^6) and R must therefore be too large for any normal application.

24. Properties of Parallel Resonant Circuits. At its resonant frequency a parallel circuit has the following properties: (1) the line current is essentially a minimum; (2) the impedance presented to the line is essentially a maximum; (3) the line current is in phase with the line voltage; and (4) the current circulating in the parallel circuit itself is usually much larger than the line current (the circulating current is Q times line current).

25. Absolute Value of Impedance at Resonance in Parallel Resonant Circuit. Letting $\omega = 1/\sqrt{LC}$ in Eq. (17) gives for the impedance of a parallel circuit at resonance

$$Z_0 \doteq \frac{(R_LR_C + X_LX_C) + j(R_CX_L - R_LX_C)}{R_L + R_C} \qquad (23)$$

The absolute value of this impedance is

$$|Z_0| \doteq \sqrt{\frac{(R_LR_C + \omega^2L^2)^2 + \omega^2L^2(R_CR_L)^2}{(R_L + R_C)^2}} \qquad (24)$$

26. Absolute Value of Impedance in General Parallel Circuit, with Negligible Resistance in Capacity Branch. In this case $R_C \doteq 0$, and from Eq. (17),

$$Z_0 \doteq X_C\left[\frac{R_LX_C - j[R_L^2 + X_L^2 - X_LX_C]}{R_L^2 + (X_L - X_C)^2}\right] \qquad (25)$$

The absolute magnitude of Z_0 is

$$|Z_0| \doteq \frac{X_C\sqrt{R_L^2 + X_L^2}}{\sqrt{R_L^2 + (X_L - X_C)^2}} \qquad (26)$$

If R_L is small compared with X_L,

$$|Z_0| \doteq \frac{X_LX_C}{\sqrt{R_L^2 + (X_L - X_C)^2}} \doteq \frac{L}{C}\,\frac{1}{\sqrt{R_L^2 + (X_L - X_C)^2}} \qquad (27)$$

At resonance $X_L = X_C$ (R_L and R_C being assumed negligible), and

$$|Z_0| \doteq \frac{L}{RC} \qquad (28)$$

The equivalent impedance of a low-resistance parallel circuit is therefore very nearly a pure resistance at the resonant frequency and has the value L/RC, approximately.

27. Parallel Resonance Curves—Z_0 and ϕ. The following useful relations for parallel resonant circuits (where R_C is considered negligible) are derived from Eq. (26):

$$\frac{|Z_0|}{\omega_rL} \doteq \left[\frac{1 + Q_r^2}{nQ_r^2}\right]\frac{\sqrt{\frac{1}{Q_r^2} + n^2}}{\sqrt{\frac{1}{Q_r^2} + \left(n - \frac{1 + Q_r^2}{nQ_r^2}\right)^2}} \qquad (29)$$

where $|Z_0|$ = the absolute impedance at any frequency f_1
ω_r = $2\pi \times$ resonance frequency
ω_1 = $2\pi \times$ any frequency
n = ω_1/ω_r
Q_r = $\omega_rL/R = Q$ at resonance.

For Q = 10 or larger, this is approximately

$$\frac{|Z_0|}{\omega_rL} \doteq \frac{\sqrt{\frac{1}{Q_r^2} + n^2}}{n\sqrt{\frac{1}{Q_r^2} + n^2 + \frac{1}{n^2} - 2}} \qquad (30)$$

and at resonance

$$\frac{|Z_0|}{\omega_rL} \doteq Q \qquad (31)$$

From Eq. (30) it can be shown that $|Z_0| = 1.414\sqrt{L/C}$ when $f_1 = 0.707f_r$. Hence the L/C ratio of a parallel resonant circuit may be expressed as a function of its impedance at 70.7 per cent of resonance frequency, or vice versa. This fact is useful in some design applications.

The ratio of $|Z_0|$ at $f_1 = 0.707f_r$ to $|Z_0|$ at resonance is

$$\frac{|Z_0| \text{ at } f_1}{|Z_0| \text{ at } f_r} = \frac{1.414}{Q} \qquad (32)$$

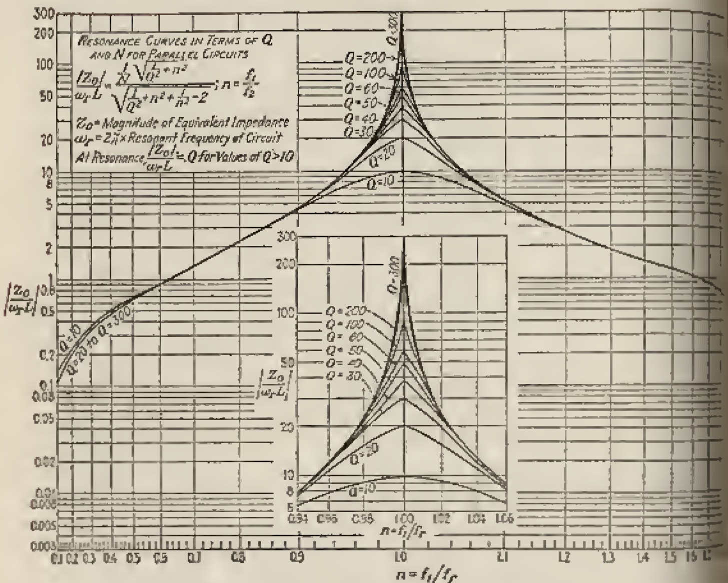


Fig. 32.—Parallel resonance curves.

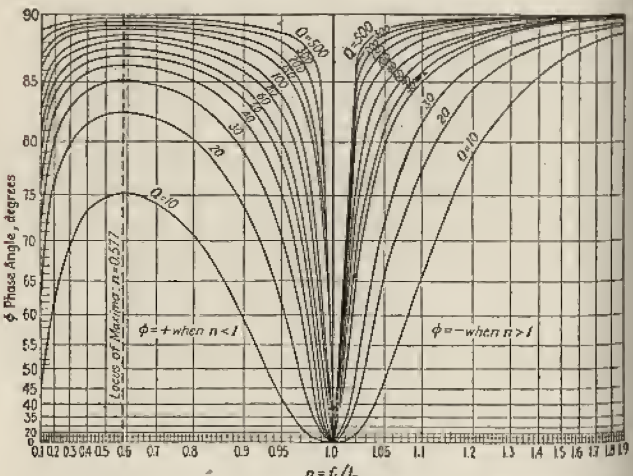


Fig. 33.—Phase angle of parallel LC circuit in terms of n and Q.

The phase angle of Z_0 is

$$\phi = \tan^{-1}\left[-nQ_r\left(\frac{1}{Q_r^2} + n^2 - 1\right)\right] \tag{33}$$

and for $Q = 10$ or larger is approximately

$$\phi = \tan^{-1}[-nQ_r(n^2 - 1)] \tag{34}$$

Using Eq. (29), the set of resonance curves of Fig. 32 has been plotted in terms of $|Z_0|/\omega_r L$, Q, and n. Similar curves for ϕ in terms of n and Q are given in Fig. 33. From these universal curves a complete resonance curve and its phase characteristic can be plotted when the constants ω_r, L, and Q or R are known.

28. Design of Parallel Resonant Circuits. To design a parallel resonant circuit, we have to determine values of L, C, R_L, and R_C to satisfy a given set of conditions. Values of $|Z_0|$ at resonance, the resonance frequency, and Q will first have to be determined by analysis of the intended use of the resonant circuit.

In a vacuum-tube oscillator, for example, f_r of the tuned circuit is known, and $|Z_0|$ at resonance is fixed by the permissible plate voltage swing. For Q (which includes the effect of the external load coupled to the tuned circuit, as well as the latter's ohmic resistance) a value of from 12 to 20 represents a good compromise between oscillator efficiency and frequency stability.

Another example of the factors involved in the choice of Q in an application is that of a tuned circuit for an r-f amplifier to pass a modulated carrier. In this case the LC circuit must have sufficient decrement to damp out its own natural oscillations between successive peaks of modulation; otherwise there is an effective decrease in modulation percentage with a corresponding loss of fidelity. If the carrier frequency is f_C and the modulation frequency f_m, the maximum decrement of the modulated carrier wave at 100 per cent modulation is approximately

$$\delta_1 = 2.303 \log_{10}\left(\frac{1}{1 - \pi\frac{f_m}{f_C}}\right) \tag{35}$$

The decrement δ_2 of the tuned circuit should be 10 to 20 times as large as δ_1 for faithful response. Then Q for the tuned circuit is

$$Q = \frac{\pi}{\delta_2} \tag{36}$$

The value of $|Z_0|$ at resonance will depend upon plate-load impedance requirements of the amplifier tube.

In some cases the ratio of volt-amperes circulating in LC to watts dissipated is the basis for the design of an LC circuit; in this case

$$Q = \frac{\text{volt-amperes}}{\text{watts dissipated}} \tag{37}$$

The effect of any load coupled to a tuned circuit must be taken into account as part of the total effective R of the circuit. If the power taken by the load is W_d watts and I_C is the circulating current in LCR, th total equivalent impedance of the circuit is, approximately,

$$R \doteq R_0 + \frac{W_d}{I_C^2} \qquad (38)$$

where R_0 is the ohmic resistance.

Examples of Design of Parallel Resonant Circuit. Assume that a parall circuit (Fig. 31) is to be resonant at 5,000 cycles, with an impedance of 4,00 ohms at resonance ($n = 1$) and an impedance of 100 ohms at 3,000 cycl ($n = 0.6$). From Fig. 32, $|Z_0|/\omega_r L = 0.9$ for all values of Q when $n = 0.6$ At resonance $|Z_0|/\omega_r L$ is to be $\frac{4,000}{100} \times 0.9 = 36$. From the curves i is found that $Q = 36$ gives $|Z_0|/\omega_r L = 36$ at $n = 1$ where $\omega_r = 31,41$ Then for $n = 1$,

$$Z_0 = 36\omega_r L = 4,000, \text{ or } L = \frac{4,000}{36 \times 31,416} = 0.00354 \text{ henry}$$

LC for 5,000 cycles $= 10.136 \times 10^{-10}$. Then $C = LC/L = 0.286 \times 10$ farad, and $R = \omega_r L/Q = 3.08$ ohms.

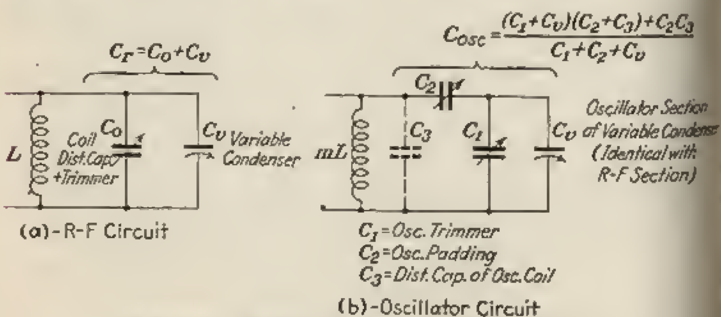


FIG. 34.—Oscillator circuits for superheterodyne.

As a second example suppose there is to be designed a tuned circuit for a r-f amplifier which requires a plate-load impedance of 10,000 ohms and whic is to amplify a 1,000-kc carrier with amplitude modulation up to 5,000 cycle From Eq. (35)

$$\delta_1 = 2.303 \log_{10} \frac{1}{1 - \pi \frac{5 \times 10^3}{10^6}} = 0.0159$$

Let

$$\delta_2 = 20\,\delta_1 = 0.318$$

$$Q = \frac{\pi}{\delta_2} = 9.85 \doteq \frac{|Z_0|}{\omega_r L} \text{ [from Eq. (31)]}$$

$$\omega_r L = \frac{|Z_0|}{Q} = \frac{10,000}{9.85} = 1,015$$

$$L = \frac{\omega_r L}{\omega_r} = \frac{1,015}{2\pi f_r} = 162\ \mu\text{h}$$

$$LC \text{ for } 1,000 \text{ kc} = 2.53 \times 10^{-10}$$

$$C = \frac{LC}{L} = \frac{2.53 \times 10^{-10}}{162 \times 10^{-6}} = 157\ \mu\mu\text{f}$$

$$R = \frac{\omega_r L}{Q} = \frac{1,015}{9.85} = 103 \text{ ohms.}$$

This consists of the ohmic resistance of LCR plus the equivalent R of the coupled load, as computed by Eq. (38).

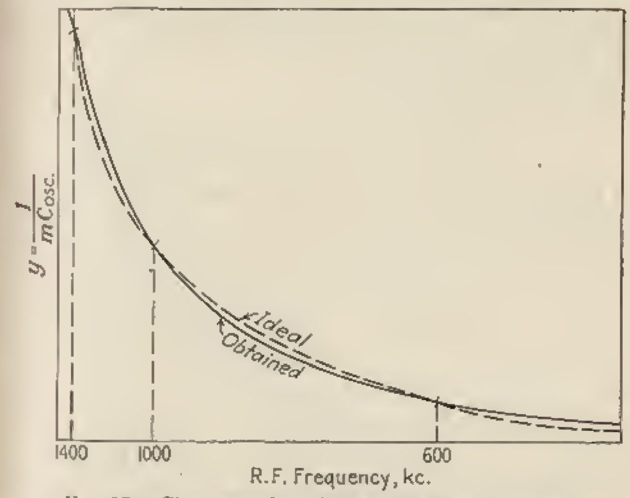


FIG. 35.—Closeness of tracking secured by formulas.

29. Design of Oscillator Tracking Circuits. In superheterodyne receivers with "ganged" condenser tuning, the most common method for tracking the oscillator-tuned circuit at a constant frequency difference from the r-f circuits is by means of a gang condenser with identical sections and with an adjustable padding condenser in series with the oscillator section. A typical oscillator circuit of this type is shown in Fig. 34. The tracking is approximate as no combination of C_1, C_2, and C_3 will give perfect alignment at more than three points on the dial shown in Fig. 35). These points are usually chosen near the ends and the middle of the frequency range—in a broadcast receiver, for example, at 1,400, 1,000, and 600 kc. Slight tracking errors will exist at all other frequencies in the band; these will be approximately proportional to the i-f frequency used. The maximum errors are at the ends of the band and amount to about 2 kc for a good design with an i.f. of 175 kc.

The values of C_1, C_2, and C_3 may be determined by calculation or by experimental methods. Either method involves a considerable amount of labor. The following design procedure, due to Roder,[1] is probably the most

[1] RODER, HANS, Oscillator Padding, *Radio Engineering*, March, 1935, p. 7.

direct method of solution (six-place logarithms or a calculating machine is recommended for all calculations):

Step 1. Known constants:

a. Three frequencies of perfect alignment ($= f_r$). (Usually 1,400, 1,000 and 600 kc for broadcast receivers.)
b. R-f circuit inductance ($= L$).
c. R-f circuit trimmer capacity ($= C_0$). (Including distributed capacity of r-f coil.)
d. Intermediate frequency ($= f_i$).
e. Distributed capacity of oscillator coil ($= C_3$).

Solution to yield: Values of C_1, C_2, and $L_{osc.}$

Units: All constants are measured in the following units:
f = frequency in kilocycles.
L = inductance in microhenries.
C = capacitance in micromicrofarads.

Step 2. Compute

$$x_n = \frac{253.3 \times 10^5}{Lf_r^2} \qquad \text{and} \qquad y_n = \frac{f_{osc.}^2 L}{253.3 \times 10^5}$$

for each alignment frequency.

Step 3. Compute

$$X = \frac{y_2 - y_3 + x_2 B - x_1 A}{B - A}; \qquad Y = \frac{y_1 B - y_3 A}{B - A}$$

where

$$A = \frac{y_1 - y_2}{x_2 - x_1} \qquad \text{and} \qquad B = \frac{y_2 - y_3}{x_3 - x_2}$$

Step 4. Compute

$$K = (x_1 - X)(y_1 - Y) = (x_2 - X)(y_2 - Y) = (x_3 - X)(y_3 - Y)$$

(The truth of these identities is a check on the accuracy of the calculations thus far.)

Step 5. Compute

$$m = \frac{1}{K}(1 - v)$$

where

$$v = 0.5u - 0.3125u^2 + 0.2188u^3$$

and

$$u = \frac{4C_3 Y}{K}$$

Step 6. Compute

$$C_1 = C_0 - X - \frac{Kv}{2Y}(1 + 0.75v + 0.625v^2 + 0.547v^3)$$

and

$$C_2 = \frac{1}{Y}\sqrt{\frac{K}{m}}$$

Step 7. Compute $L_{osc.} = mL$

Example:

Step 1. Let

f_{r_1} = 1,400 kc L = 200μh
f_{r_2} = 1,000 kc C_0 = 30 $\mu\mu$f
f_{r_3} = 600 kc C_3 = 15 $\mu\mu$f
f_i = 175 kc.

Step 2.

	f_r	$f_{osc.}$	x	y
1	1,400	1,575	64.617	0.0195865
2	1,000	1,175	126.650	0.0109110
3	600	775	351.806	0.0047424

Step 3. $A = 140.0126$; $B = 27.3531$.

$$X = -5.1103; \qquad Y = 0.1138 \times 10^{-2}.$$

Step 4. $K = 1.2863$.
Step 5. $u = 5.3097$; $v = 2.5700 \times 10^{-2}$; $m = 0.7574$.
Step 6. $C_1 = 20.17\ \mu\mu$f; $C_2 = 1451.5\ \mu\mu$f.
Step 7. $L_{osc.}$ = 151.50 microhenries.

30. Tapped Tank Circuits. In some cases the high impedance of a parallel LCR circuit at resonance is a disadvantage, *e.g.*, at the end of a low impedance transmission line where the correct termination is about 500 ohms. However, the low impedance can be obtained by tapping the LCR circuit in either the L or C branch as shown in Fig. 36. The result is a coupled circuit, that part of the reactance between B and C being the mutual impedance.

1. *Capacity Tapped*. In Fig. 36*a*, the impedance at B-C is

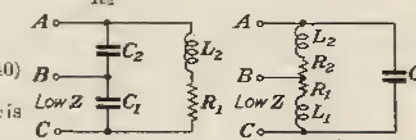
$$|Z_{BC}| = \frac{\sqrt{L_2^2 C_2^2\left(\frac{1}{C_1(C_1 + C_2)}\right)^2 + \frac{R_2^2 L_2 C_2}{C_1(C_1 + C_2)}}}{R_2} \tag{39}$$

If R_1 is small,

$$|Z_{BC}| \doteq \frac{L_2 C_2}{R_2 C_1(C_1 + C_2)} \tag{40}$$

and its ratio to the impedance Z_{AC} is

$$\frac{|Z_{BC}|}{|Z_{AC}|} = \frac{C_2^2}{(C_1 + C_2)^2} \tag{41}$$



(*a*) Capacity Tapped (*b*) Inductance Tapped

FIG. 36.—Tapped tank circuits.

The resonant frequency is

$$f_r = \frac{1}{2\pi\sqrt{L\frac{C_1 C_2}{C_1 + C_2}}} \tag{42}$$

and the impedances Z_{AC} and Z_{BC} are both purely resistive at resonance. The ratio of C_1 to C_2 for a given ratio between Z_{AC} and Z_{BC} is

$$\frac{C_1}{C_2} \doteq \left(\sqrt{\frac{Z_{AC}}{Z_{BC}}} - 1\right) \tag{43}$$

In terms of the resonant frequency, inductance, and the impedance ratio,

$$C_1 = \frac{1}{4\pi^2 f_r^2 L}\sqrt{\frac{Z_{AC}}{Z_{BC}}} \tag{44}$$

$$C_2 = \frac{1}{4\pi^2 f_r^2 L\left(1 - \sqrt{\frac{Z_{BC}}{Z_{AC}}}\right)} \quad (4$$

2. *Inductance Tapped.* In Fig. 36*b* the inductance is tapped, and the impedance at *B-C* is (assuming no mutual inductance between L_1 and L_2)

$$|Z_{BC}| = \frac{\sqrt{\left(R_1R_2 - \frac{L_1L_2}{(L_1 - L_2)C_2} + \frac{L_2}{C_2}\right)^2 + \left(\frac{R_2L_1}{\sqrt{(L_1+L_2)C_2}} + \frac{R_1L_2}{\sqrt{(L_1+L_2)C_2}} - \frac{R_1\sqrt{(L_1+L_2)C_2}}{C_2}\right)^2}}{R_1 + R_2} \quad (4$$

If R_1 and R_2 are small,

$$|Z_{BC}| = \frac{L_2}{C_2(R_1 + R_2)} \cdot \frac{L_2}{(L_1 + L_2)} \quad (4$$

and its ratio to the total impedance Z_{AC} is

$$\frac{|Z_{BC}|}{|Z_{AC}|} = \frac{L_2^2}{(L_1 + L_2)^2} \quad (4$$

The resonant frequency is

$$f_r = \frac{1}{2\pi\sqrt{(L_1 + L_2)C_2}} \quad (4$$

and the impedances Z_{AC} and Z_{BC} are both resistive at resonance.

The ratio of L_1 to L_2 for a given ratio between Z_{AC} and Z_{BC} is

$$\frac{L_1}{L_2} = \sqrt{\frac{Z_{AC}}{Z_{BC}}} - 1 \quad (5$$

In terms of the frequency, capacity, and the impedance ratio,

$$L_1 = \frac{1}{4\pi^2 f_r^2 C_2}\left(1 - \sqrt{\frac{Z_{BC}}{Z_{AC}}}\right) \quad (5$$

$$L_2 = \frac{1}{4\pi^2 f_r^2 C_2}\sqrt{\frac{Z_{BC}}{Z_{AC}}} \quad (5$$

31. Measurement of Parallel Resonance Impedance. A convenient method of experimentally determining the resonance impedance of a parallel circuit is shown in Fig. 37. *LC* is the circuit to be measured. This method is based on the fact that the circuit just commences to oscillate when the "negative resistance" of the tube characteristic is numerically equal to the impedance of the *LC* plate circuit. In practice, a type 22 or 24 tube is satisfactory, in which case *B* should be about 12 volts and *C* about 25 volts. The potentiometers *G* and *P* control the grid bias and plate voltages, respectively. The latter should be between 60 and 80 volts for the *B* voltage mentioned. A receiver or other indicating device is loosely coupled to *LC* to detect the point where oscillation starts. *G* and *P* are adjusted until the circuit is on the verge of oscillation. Then *LC* is short-circuited by closing the key *S*

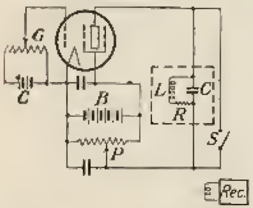


FIG. 37.—Circuit for measuring resonant impedance of parallel circuit.

and *P* is varied a few volts above and below the setting at which oscillation occurred and the values of plate current noted. The values of *G* and *B* are, of course, unchanged during this latter adjustment. The slope of the $e_p - i_p$ curve through the value of e_p where oscillation occurred is the negative resistance and is numerically equal to the impedance $|Z_0|$. If *L* and *C* are known, *R* can be computed from Eq. (28):

$$|Z_0| = \frac{L}{RC}$$

or

$$R = \frac{L}{|Z_0|C}$$

This also suggests the use of the above circuit for measuring r-f resistance, by inserting an unknown resistance in series in the *LC* circuit and measuring its impedance before and after the insertion is made. By a similar process, capacity or inductance may also be measured. The method as outlined is limited by tube characteristics to impedances of about 10,000 ohms and over.

COUPLED CIRCUITS

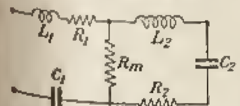
32. Coupling. If two circuits have one or more common impedances, they are said to be electrically *coupled.* A common impedance is any impedance so situated that it causes the current in one circuit to influence the current in the other. The impedance may be resistive, reactive, or both.

33. Coefficient of Coupling. The coefficient of coupling is

$$K = \frac{X_m}{\sqrt{X_1X_2}} \quad (53)$$

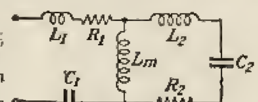
where X_m is any one component of the mutual impedance (resistance, capacitive reactance or inductive reactance) and X_1 and X_2 are the total impedance components of the same kind in the respective circuits. *K* varies in value between zero and 1; if it is nearly 1, the coupling is *close* or *tight;* if near zero, the coupling is *loose.*

34. Direct and Indirect Coupling. If the common impedance is a resistance, inductance, or capacitance connected directly between the



$$K = \frac{R_m}{\sqrt{R_1'R_2'}} \qquad R_1' = R_1 + R_m \qquad R_2' = R_2 + R_m$$

FIG. 38.—Direct resistive coupling.



$$K = \frac{L_m}{\sqrt{L_1'L_2'}} \qquad L_1' = L_1 + L_m \qquad L_2' = L_2 + L_m$$

FIG. 39.—Direct inductive coupling.

two circuits, the coupling is *direct.* Such circuits are shown in Figs. 38, 39, and 40. If the common impedance is a transformer, the coupling

Coupled Circuits: Indirect Inductive

Equivalent impedance:

$$Z_0 = j\left[\left(\omega L_1' - \frac{1}{\omega C_1}\right) - \frac{(\omega M)^2}{\left(\omega L_2' - \frac{1}{\omega C_2}\right)}\right]$$

Circuit Resonance Curve

Equivalent direct-coupled circuit: Indirect inductive coupling is equivalent to direct inductive coupling if

$$L_1 = L_1' - M$$
$$L_2 = L_2' - M$$
$$L_m = M$$

where L_1' and L_2' are the self-inductances of the coils.

$$f_1 = \sqrt{\frac{f_a^2 + f_b^2 - \sqrt{(f_a^2 - f_b^2)^2 + 4k^2 f_a^2 f_b^2}}{2(1 - k^2)}} \qquad f_a = \frac{1}{2\pi\sqrt{L_1'C_1}}$$

$$f_2 = \sqrt{\frac{f_a^2 + f_b^2 + \sqrt{(f_a^2 - f_b^2)^2 + 4k^2 f_a^2 f_b^2}}{2(1 - k^2)}} \qquad f_b = \frac{1}{2\pi\sqrt{L_2'C_2}}$$

$$k = \frac{M}{\sqrt{L_1'L_2'}}$$

Special cases:

a. Both circuits tuned to same frequency ($f_a = f_b$).

$$f_1 = \frac{f_a}{\sqrt{1 + k}} \qquad f_2 = \frac{f_a}{\sqrt{1 - k}}$$

b. Loose coupling ($f_a = f_b$; $M << L_1'$ and L_2'; $k \doteq 0$).

$$f_1 \doteq f_2 \doteq f_a \doteq \frac{1}{2\pi\sqrt{L_1'C_1}} \doteq \frac{1}{2\pi\sqrt{L_2'C_2}}$$

c. Close coupling ($f_a = f_b$; $M >> L_1'$ and L_2'; $k \doteq 1$).

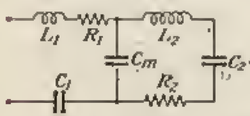
$$f_1 \doteq \frac{f_a}{\sqrt{2}} \doteq \frac{1}{2\pi\sqrt{2MC_1}}$$
$$f_2 \doteq \infty$$

d. Both circuits identical.

$$\begin{cases} f_a = f_b \\ L_1' = L_2' \\ C_1 = C_2 \end{cases}$$

$$f_1 = \frac{1}{2\pi\sqrt{(L_1' + M)C_1}}$$
$$f_2 = \frac{1}{2\pi\sqrt{(L_1' - M)C_1}}$$
$$k = \frac{M}{L_1'}$$

is indirect and is usually called merely *inductive coupling.* This type of coupling is illustrated in Fig. 41. Indirect capacitive coupling is illustrated in Fig. 42.



$$K = \sqrt{\frac{C_1C_2}{(C_1 + C_m)(C_2 + C_m)}}$$

FIG. 40.—Direct capacitive coupling.

FIG. 41.—Indirect inductive coupling.

From Figs. 38 to 40 it is apparent that direct-coupled circuits may be considered as networks of impedances in series and parallel, as in Fig. 43.

The notion of "equivalent impedance" (Art. 7) is a useful concept in the treatment of such circuits. In the present treatment of coupled circuits the equivalent impedance is determined by combining the various impedance elements of the circuits according to the laws of parallel and series combination as discussed in Eqs. (5), (6), and (7).

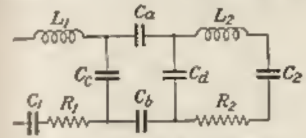


FIG. 42.—Indirect capacitive coupling.

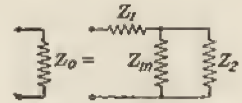


FIG. 43.—Equivalent impedance of direct-coupled circuits.

The equivalent impedance of the network of Fig. 43 is

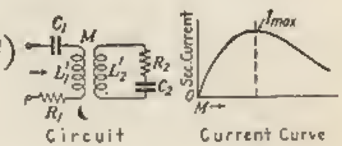
$$Z_0 = Z_1 + \frac{Z_m Z_2}{Z_m + Z_2}$$
$$= \frac{Z_1 Z_m + Z_1 Z_2 + Z_m Z_2}{Z_m + Z_2} \qquad (54)$$

Coupled Circuits: Inductive or Transformer with Resistance

Equivalent Impedance:

$$Z_0 = R_0 + jX_0 = \left(R_1 + \frac{\omega^2 M^2 R^2}{|Z_2^2|}\right) + j\left(X_1 - \frac{\omega^2 M^2 X_2}{|Z_2^2|}\right)$$

where
$$|Z_2^2| = R_2^2 + X_2^2$$
$$X_1 = \omega L_1' - \frac{1}{\omega C_1}$$
$$X_2 = \omega L_2' - \frac{1}{\omega C_2}$$



Circuit Current Curve

Special case:

If M is variable, and both circuits tuned to the same frequency, the current in the secondary varies with M as shown in the figure.

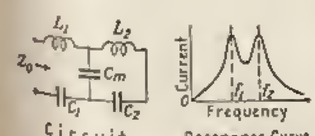
The maximum secondary current occurs at

$$\omega M = \sqrt{R_1 R_2}$$

Coupled Circuits: Direct Capacitive

Impedance:

$$Z_0 = j\frac{\frac{1}{\omega C_m}\left(\omega L_1 - \frac{1}{\omega C_1}\right) - \left(\omega L_1 - \frac{1}{\omega C_1}\right)\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \frac{1}{\omega C_m}\left(\omega L_2 - \frac{1}{\omega C_2}\right)}{\frac{1}{\omega C_m} - \omega L_2 + \frac{1}{\omega C_2}}$$



Circuit Resonance Curve

General case: L_1, L_2, C_1, C_2 and C_m unrestricted.

$$f_1 = \sqrt{\frac{f_a^2 + f_b^2 - \sqrt{(f_a^2 - f_b^2)^2 + 4k^2 f_a^2 f_b^2}}{2}}$$

$$f_2 = \sqrt{\frac{f_a^2 + f_b^2 + \sqrt{(f_a^2 - f_b^2)^2 + 4k^2 f_a^2 f_b^2}}{2}}$$

where $f_a = \dfrac{1}{2\pi\sqrt{L_1\dfrac{C_1C_m}{C_1 + C_m}}}$ $\qquad f_b = \dfrac{1}{2\pi\sqrt{L_2\dfrac{C_2C_m}{C_2 + C_m}}}$

Coefficient of coupling:

$$k = \sqrt{\frac{C_1C_2}{(C_1 + C_m)(C_2 + C_m)}}$$

Special cases:

a. Both circuits tuned to same frequency ($f_a = f_b$).

$$f_1 = f_a\sqrt{1 - k} \qquad f_2 = f_b\sqrt{1 + k}$$

b. Loose coupling ($f_a = f_b$ and $C_m >> C_1$ and C_2; $k \doteq 0$).

$$f_1 \doteq f_2 \doteq f_a \doteq \frac{1}{2\pi\sqrt{L_1C_1}} \doteq \frac{1}{2\pi\sqrt{L_2C_2}}$$

c. Close coupling ($f_a = f_b$ and $C_m << C_1$ and C_2; $k \doteq 1$).

$$f_1 \doteq 0 \text{ and } f_2 \doteq \sqrt{2}\, f_a \doteq \frac{\sqrt{2}}{2\pi\sqrt{L_1C_m}} \doteq \frac{\sqrt{2}}{2\pi\sqrt{L_2C_m}}$$

d. Both circuits identical.

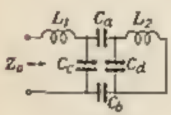
$$\begin{cases} f_a = f_b \\ L_1 = L_2 \\ C_1 = C_2 \end{cases}$$

$$f_1 = \frac{1}{2\pi\sqrt{L_1C_1}} \qquad f_2 = \frac{1}{2\pi\sqrt{L_1\dfrac{C_1C_m}{2C_1 + C_m}}}$$

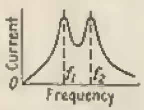
Coupled Circuits: Indirect Capacitive

Equivalent impedance:

$$Z_e = j\left[\omega L_1 - \frac{1}{\omega C_c}\,\frac{\left(\omega L_2 - \dfrac{1}{\omega C_d}\right)\dfrac{1}{\omega C'} + \dfrac{L_2}{C_d}}{\left(\omega L_2 - \dfrac{1}{\omega C_d}\right)\dfrac{1}{\omega C''} + \dfrac{L_2}{C_d}}\right]$$



Circuit



Resonance Curve

where $\qquad C' = \dfrac{C_aC_b}{C_a + C_b}$

$$C'' = \frac{C_aC_bC_c}{C_aC_b + C_aC_c + C_bC_c}$$

General case: L_1, L_2, C_a, C_b, C_c and C_d unrestricted.

$$f_1 = \sqrt{\frac{f_a^2 + f_b^2 - \sqrt{(f_a^2 - f_b^2)^2 + 4k^2f_a^2f_b^2}}{2}}$$

$$f_2 = \sqrt{\frac{f_a^2 + f_b^2 + \sqrt{(f_a^2 - f_b^2)^2 + 4k^2f_a^2f_b^2}}{2}}$$

where $\quad f_a = \dfrac{1}{2\pi\sqrt{L_1\left(C_c + \dfrac{C_dC'}{C_d + C'}\right)}} \qquad f_b = \dfrac{1}{2\pi\sqrt{L_2\left(C_d + \dfrac{C_cC'}{C_c + C'}\right)}}$

Coefficient of coupling:

$$k = \frac{C'}{\sqrt{(C_c + C')(C_d + C')}}$$

Special cases:

a. Both circuits tuned to same frequency ($f_a = f_b$).

$$f_1 = f_a\sqrt{1 - k} \qquad f_2 = f_a\sqrt{1 + k}$$

b. Loose coupling ($C_a + C_b$) $<<$ C_c and C_d; $k \doteq 0$; $f_a = f_b$.

$$f_1 \doteq f_2 \doteq f_a \doteq \frac{1}{2\pi\sqrt{L_1C_c}} \doteq \frac{1}{2\pi\sqrt{L_2C_d}}$$

c. Close coupling ($f_a = f_b$); ($C_a + C_b$) $>>$ C_c and C_d; $k \doteq 1$.

$$f_1 \doteq 0$$

$$f_2 \doteq \sqrt{2}\, f_a \doteq \frac{1}{\pi\sqrt{2L_1(C_c + C_d)}} \doteq \frac{1}{\pi\sqrt{2L_2(C_c + C_d)}}$$

d. Both circuits identical.

$$\begin{cases} f_a = f_b \\ L_1 = L_2 \\ C_1 = C_2 \end{cases}$$

$$f_1 \doteq \frac{1}{2\pi\sqrt{L_1(C_c + 2C')}}$$

$$f_2 \doteq \frac{1}{2\pi\sqrt{L_1C_c}}$$

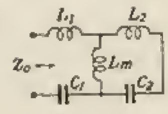
$$k = \frac{C'}{C_c + C'}$$

Coupled Circuits: Direct Inductive

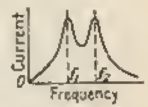
Equivalent impedance:

$$Z_e = j\frac{\omega L_m\left(\omega L_1 - \dfrac{1}{\omega C_1}\right) + \left(\omega L_1 - \dfrac{1}{\omega C_1}\right)\left(\omega L_2 - \dfrac{1}{\omega C_2}\right) + \omega L_m\left(\omega L_2 - \dfrac{1}{\omega C_2}\right)}{\omega L_m + \omega L_2 - \dfrac{1}{\omega C_2}}$$

General case: L_1, L_2, L_m, C_1 and C_2 unrestricted.



Circuit



Resonance Curve

$$f_1 = \sqrt{\frac{f_a^2 + f_b^2 - \sqrt{(f_a^2 - f_b^2)^2 + 4k^2f_a^2f_b^2}}{2(1 - k^2)}}$$

$$f_2 = \sqrt{\frac{f_a^2 + f_b^2 + \sqrt{(f_a^2 - f_b^2)^2 + 4k^2f_a^2f_b^2}}{2(1 - k^2)}}$$

where $\qquad f_a = \dfrac{1}{2\pi\sqrt{(L_1 + L_m)C_1}} \qquad f_b = \dfrac{1}{2\pi\sqrt{(L_2 + L_m)C_2}}$

Coefficient of coupling $\qquad k = \dfrac{L_m}{\sqrt{(L_1 + L_m)(L_2 + L_m)}}$

Special cases:

a. Both circuits tuned to the same frequency ($f_a = f_b$).

$$f_1 = \frac{f_a}{\sqrt{1 + k}} \qquad f_2 = \frac{f_a}{\sqrt{1 - k}}$$

b. Loose coupling ($f_a = f_b$; $L_m << L_1$ and L_2; $k \doteq 0$).

$$f_1 \doteq f_2 \doteq f_a \doteq \frac{1}{2\pi\sqrt{L_1C_1}} \doteq \frac{1}{2\pi\sqrt{L_2C_2}}$$

c. Close coupling ($f_a = f_b$; $L_m >> L_1$ and L_2; $k \doteq 1$).

$$f_1 \doteq \frac{f_a}{\sqrt{2}} \doteq \frac{1}{2\pi\sqrt{2L_mC_1}}$$

$$f_2 \doteq \infty$$

d. Both circuits identical.

$$\begin{cases} f_a = f_b \\ L_1 = L_2 \\ C_1 = C_2 \end{cases}$$

$$f_1 = \frac{1}{2\pi\sqrt{(L_1 + 2L_m)C_1}}$$

$$f_2 = \frac{1}{2\pi\sqrt{L_1C_1}}$$

$$k = \frac{L_m}{L_1 + L_m}$$

35. Use of Resistanceless Circuits in Calculations. Each impedan[ce] in Eq. (54) is in general of the form $R_0 + jX_0$, so that the expressio[n] becomes somewhat involved if an exact solution is made. In many actu[al] applications, however, coupled circuits are also sharply tuned, which [is] tantamount to saying that their resistances are small compared with th[e] reactances. For such cases computations are much simplified witho[ut] undue sacrifice of accuracy if the circuits are assumed to be resistanceles[s].

36. Stray Coupling. Because of the apparent increase in resistan[ce] of a circuit when another circuit is coupled to it, spurious and uninte[n]tional coupling due to stray fields and the proximity of other apparat[us] may appreciably affect the resistance of r-f circuits and introduce unne[c]essary losses unless precautions are taken to avoid it. Stray effects a[re] due principally to capacity coupling and stray inductive coupling. Th[e] former varies with the areas of conductors and a-c voltages involv[ed] and inversely with the distances between the conductors, while the latt[er] varies with ampere-turns, the diameter of the heavy current path in th[e] circuit, and inversely with the distance between the circuit and oth[er] conductors in which induced currents flow.

SOME SPECIAL APPLICATIONS OF LCR CIRCUITS

37. Band-pass R-f Circuits. If two identical tuned circuits ar[e] capacitatively or inductively coupled (upper part Fig. 44*a* and *b*), th[e] circuit acts as a band-pass filter with a band width approximately

$$f_s = f_1 - f_2 \doteq \frac{\sqrt{X_m^2 - R^2}}{2\pi L} \qquad (55)$$

The band width varies with the tuning, increasing with the frequenc[y] in the inductive case, and decreasing with the frequency in the capacit[a]tive case (lower parts Fig. 44*a* and *b*). These opposing effects may b[e] combined in the manner shown in Fig. 44*c*, so that the band width i[s] maintained substantially constant while the circuits are tuned over a wide range of frequency by adjustment of C_1 and C_2.

Uehling[1] has shown that this condition obtains when

$$X_{m_n} = \pm\sqrt{R_n^2 + 4\pi^2L^2f_s^2} \qquad (56)$$

where R_n is the resistance and L the total inductance of each branch and f_s is the band width. With X_m computed for the two boundary frequencie[s] f_a and f_b of the tuning range, the values of M and C_m required are given by

[1] *Electronics*, p. 279, September, 1930.

$$M = \frac{X_{m_a}f_b - X_{m_b}f_a}{2\pi(f_a^2 - f_b^2)} \qquad (57)$$

$$C_m = \frac{f_a^2 - f_b^2}{2\pi f_af_b(X_{m_a}f_b - X_{m_b}f_a)} \qquad (57a)$$

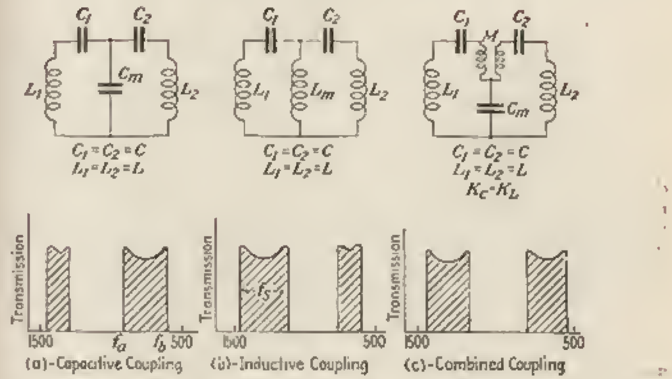


FIG. 44.—Coupled circuits as band-pass filters.

Representative values of M and C_m for f_a = 1,500 kc, f_b = 550 kc, R_n = 30 ohms, R_b = 10 ohms, $L = 200 \times 10^{-6}$ henry, and f_s = 10 kc, which are typical constants of broadcast circuits, are

$$M = 3.2 \times 10^{-6} \text{ henry}$$

and

$$C_m = 0.06\ \mu f$$

The inductive coupling M must be negative so that its effect will be additive to that of C_m. This may be obtained by winding the coils M (Fig. 44) of two wires side by side and by connecting the "start" ends of the coils to C_1 and C_2 and the "finish" ends to C_m.

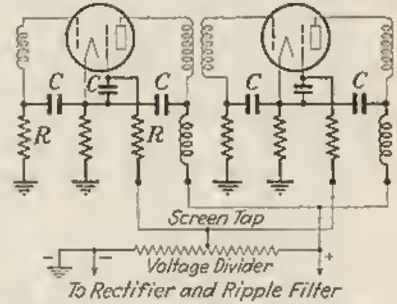


FIG. 45.—Resistance-capacity filter usage.

38. Decoupling Filters. When the plate current for several tubes of a high-gain amplifier is obtained from a single source, the internal resistance of the source is common to all the plate circuits and is likely to act as a coupling between stages. Similar couplings may exist through a bleeder circuit when screen voltage for two or more tubes is taken from a common tap or through a bias resistor common to the control-grid circuits of several tubes. To reduce such stray couplings to negligible amounts, decoupling filters are generally inserted in the circuits of each tube and separate bias resistors are used.

A typical application of decoupling filters is shown in Fig. 45, the filt elements being indicated by heavy lines. The condensers C furni low-impedance paths back to the cathodes for the signal currents flowi in the grid, screen-grid, an plate circuits, while the hig impedance resistors R a chokes in the leads to the vol age divider prevent any appr ciable flow of signal currents i that direction. The choice values for these resistors ar chokes depends principal upon the currents in the lea and the permissible d-c voltag

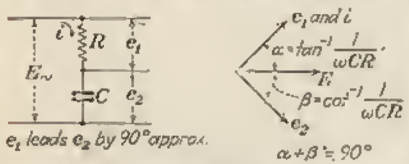


Fig. 46.—*CR* circuit for obtaining out-of-phase voltages.

drop in each filter. The impedance of each by-pass condenser should b not more than 10 per cent of that of the associated resistor or choke, at th frequency for which the amplifier is designed to operate. On the oth hand, the value of C should not be so large in any filter that "blocking" or motorboating occurs due to too high a time constant.

The impedance of a choke coil (neglecting its resistance) is

$$X_L = 6.28fL \text{ ohms,}$$

and that of a condenser is

$$X_C = \frac{10^6}{6.28fC} \text{ ohms}$$

where f = frequency in cycles per second
L = inductance in henrys
C = capacity in microfarads.

The value of each cathode resistor, when separate biasing resistors ar used, is equal to the bias required, divided by the total cathode d.c.

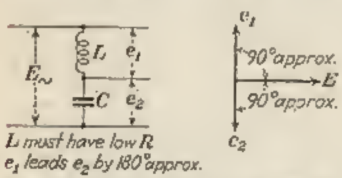


Fig. 47.—*LC* circuit for obtaining out-of-phase voltages.

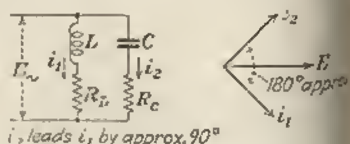


Fig. 48.—Circuit for obtaining cur rents out of phase by 90 deg.

that tube. The screen-grid filter resistors serve as voltage-droppi resistors as well as filters, and their values are determined by the IR drop required for correct screen voltages.

39. Circuits for Obtaining Out-of-phase Voltages and Currents. Tw circuits producing voltages 90 or 180 deg. out of phase are shown i Figs. 46 and 47 with their vector diagrams. These are often useful circuit designs and oscillograph measurements. To maintain the phase relations, high impedance circuits only should be connected acros e_1 and e_2.

A circuit for obtaining currents 90 deg. out of phase with each other is shown in Fig. 48.

To utilize these currents, non-reactive loads R_L and R_C are introduced, with values such that $R_LR_C = L/C$ and $R_L = R_C$.

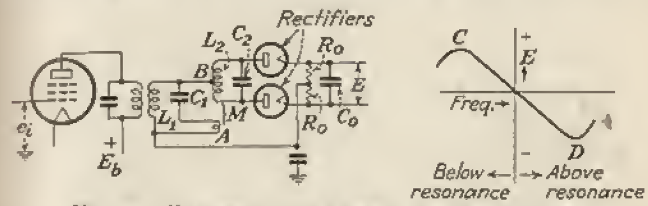


Fig. 49.—Frequency discriminator circuit and curve.

40. Frequency Discriminator Circuit. The frequency discriminator circuit shown in Fig. 49 is applied in automatic frequency control, frequency-modulation detection, and frequency-drift indicators, etc. L_1C_1 and L_2C_2 are tuned to the same frequency and doubly coupled: (1) directly at B and (2) inductively by M. After rectification, a bias E is obtained which, between limits C and D, is proportional to the difference between the frequency of the input voltage and the resonance frequency of LC. The time constant of R_0C_0 should be much less than the period of one cycle of the frequency variation in the input voltage.

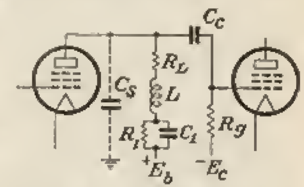


Fig. 50.—Compensated resistance-coupled amplifier.

41. Compensation in Resistance-coupled Amplifier. In a conventional resistance-coupled amplifier (Fig. 50) the amplification falls off at low frequencies because of increasing impedance of C_c and at high frequencies because of the shunting effect of stray capacitance C_s. In wide-band amplifiers, the compensating impedances L and R_1C_1 are added. For approximately constant gain between frequency limits f_1 (low) and f_2 (high),

$$\left.\begin{aligned} R_L &= \frac{1}{2\pi f_2 C_s} & L &= \frac{R_L}{4\pi f_2} \\ C_1 &= \frac{1}{2\pi f_1 R_L} & R_1 &= 2R_L \end{aligned}\right\} \qquad (58)$$

This type of compensation also tends to correct for phase shift near the limits f_1 and f_2.

RECURRENT NETWORKS

42. General Types. *Recurrent networks* are iterative combinations of L, C, and R, such as those shown in Fig. 51.

The transmission characteristics of such structures vary with frequency in a singular manner and introduce both useful and detrimental effects in r-f and a-f circuits. Examples of recurrent networks are transmission lines (actual and artificial) and wave filters.

43. Terminating Conditions for No Reflection and Maximum Powe Transfer. If a recurrent network is terminated at the nth section in a impedance equal to its image impedance, there is no reflection at th termination, and the network behaves as though it had an infinite numb of sections, in so far as its input terminals are concerned.

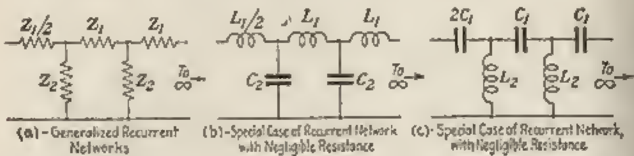


FIG. 51.—Types of infinitely long recurrent network structures

A long line so isolates its terminating impedances (the source and lo impedances) that the apparent value of each as measured from the oppo site end of the line is very nearly equal to the line impedance and pra tically independent of the terminations. Consequently, to obtain maximum transfer of power from source to line and from line to load, th source and load impedances must equal the characteristic impedance o the line, or be matched to the line b transformers whose turns ratios a equal to the square root of the ra of termination and line impedance A line terminated in its characterist impedance at both ends also has minimum reflection from its termi nals, and in general a line thus ope ated has the lowest total transmissio loss.

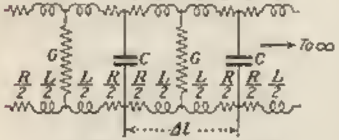


FIG. 52.—Line constants.

In a structure having lumped co stants, and terminated at one of its series elements, the series impedan in each end section is one-half the value of the series impedance in th internal sections (Fig. 51). If the termination is at a shunt elemen the shunt impedance at each end is made twice the shunt impedan in the internal sections.

44. Transmission Lines. Transmission lines are recurrent structur having continuously distributed impedances. Two wires in space hav besides their ohmic resistance, shunt capacity and series inductance a are thus equivalent to the recurrent structure of Fig. 52, where L, and R are the constants of a very short length (Δl) of the line and G is th conductance due to leakage between the wires in the same length.

45. General Properties of a Transmission Line. The *characteristic imp ance* is

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \text{ ohms} \qquad (5$$

or

$$|Z_0| = \sqrt[4]{\frac{(R^2 + \omega^2L^2)}{(G^2 + \omega^2C^2)}} \text{ ohms} \qquad (6$$

or

$$Z_0 = \sqrt{Z_{oc}Z_{sc}} \text{ ohms} \qquad (6$$

where Z_{oc} and Z_{sc} are the input impedances with the far end open- and short-circuited, respectively.

The *propagation constant* is

$$P = \sqrt{(R + j\omega L)(G + j\omega C)} = A + jB \qquad (62)$$

R, L, G, and C being the resistance, inductance, leakance, and capacitance per unit length of the line.

Attenuation Constant. The real part (A) of P is the attenuation constant and is

$$A = 6.141\sqrt{\sqrt{(R^2 + \omega^2L^2)(G^2 + \omega^2C^2)} + RG - \omega^2LC} \text{ db per unit length} \qquad (63)$$

Wave-length Constant. The quadrature part (B) of P is the wave-length constant and is

$$B = 0.707\sqrt{\sqrt{(R^2 + \omega^2L^2)(G^2 + \omega^2C^2)} - RG + \omega^2LC} \text{ radians per unit length} \qquad (64)$$

The *velocity of propagation* is

$$V = \frac{\omega}{B} = \frac{2\pi f}{B} \text{ unit lengths per second} \qquad (65)$$

The *wave length* is

$$\lambda = \frac{2\pi}{B} \text{ unit lengths} \qquad (66)$$

The *retardation time* is

$$t = \frac{B}{\omega} = \frac{B}{2\pi f} \text{ sec. per unit length} \qquad (67)$$

Input Impedance of a Line Terminated at Its Far End by an Impedance Z_e. Let Z_i = input impedance of the line
Z_0 = characteristic impedance of the line
Z_e = terminating impedance at the far end
θ = propagation factor.

The input impedance of a line so terminated is

$$Z_i = Z_0\left[\frac{Z_e \cosh\theta + Z_0 \sinh\theta}{Z_0 \cosh\theta + Z_e \sinh\theta}\right] \qquad (68)$$

The *propagation factor* is

$$\theta = lP \qquad (69)$$

where l = length
P = propagation constant per unit length.

In the communication field, transmission lines may be classified according to the frequencies they are used to transmit, as *audio-* or *radio-frequency* lines. Simplified forms of the general transmission line formulas result from the introduction of approximations appropriate to each case.

46. Audio-frequency Lines. In open-wire lines and large-gage cables, G is negligible, so that

$$A \doteq 6.14\sqrt{\omega C\sqrt{R^2 + \omega^2L^2} - \omega^2LC} \text{ db per unit length} \qquad (70)$$

and

$$B \doteq 0.707\sqrt{\omega C\sqrt{R^2 + \omega^2L^2} + \omega^2LC} \text{ radians per unit length} \qquad (71)$$

In small-gage cables, both L and G become negligibly small, and

$$A \doteq 15.39\sqrt{fRC} \text{ db per unit length} \tag{72}$$

and

$$B \doteq 1.772\sqrt{fRC} \text{ radians per unit length} \tag{73}$$

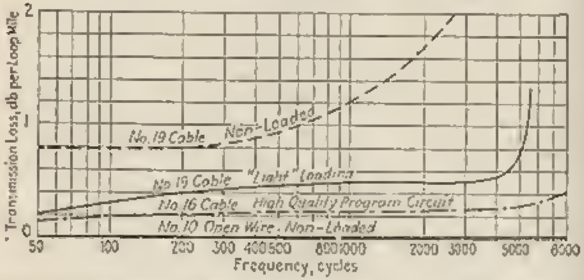


FIG. 53.—Transmission-loss characteristics of various audio-frequency circuits.

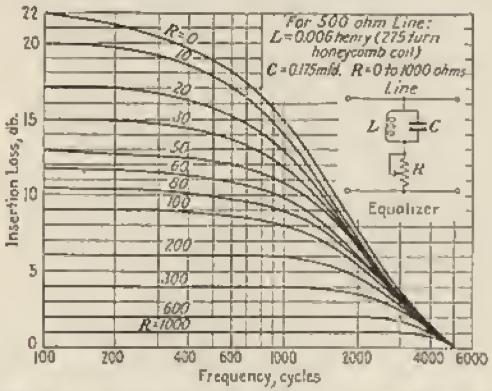


FIG. 54.—Attenuation-frequency characteristic of equalizer shunted across a 500-ohm circuit.

In both cases the attention is seen to vary with frequency. The transmission-loss frequency characteristics of various kinds of a-f circuits are shown in Fig. 53, and other characteristics of typical audio lines are shown in Table II.

47. Equalization of Transmission-loss Characteristic. From the curves in Fig. 53 it is evident that if a band of frequencies is transmitted over a line, the higher frequencies will suffer more attenuation than the low frequencies, resulting in distortion. The prevention of this condition necessitates the use of attenuation equalizers in high quality circuits. A typical 5,000-cycle equalizer for this purpose and its transmission-loss curves are illustrated in Fig. 54, and the curves for the bare line, equalizer alone, and the equalized line are shown in Fig. 55. The equalizer is

usually connected in shunt across the receiving end of the line, preceding other apparatus.

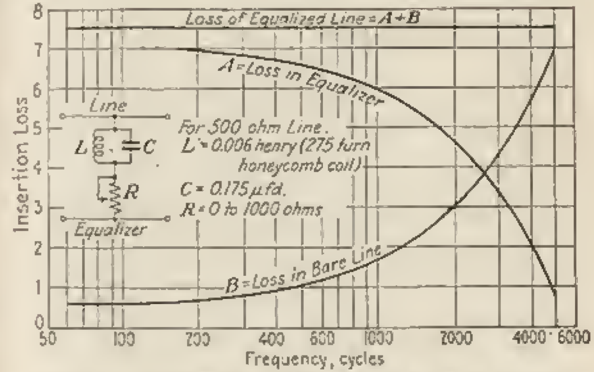


FIG. 55.—Attenuation equalizer for short cable circuits.

48. Artificial Lines. An artificial line is a compact network of lumped impedances to simulate the electrical characteristics of an actual line. Such a network having approximately the characteristics of an unloaded cable or open-wire circuit may be constructed as shown in Fig. 56 and is useful in laboratory measurements and investigations.

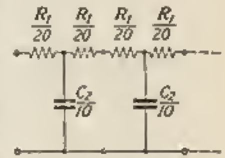


FIG. 56.—Artificial non-loaded cable.

The constants R_1 and C_2 are the loop resistance and capacity of the full length of the line to be represented. For standard cable $R_1 = 88$ ohms and $C_2 = 0.051\mu f$ per loop mile; values for various other lines are given in Table II. As the similarity between the artificial and the actual line increases with the number of sections in the former, it is preferable to use at least ten sections, and not more than 1 mile of cable or 10 miles of open wire should be represented by one section. The end sections should be "mid-series" terminated—*i.e.*, their series impedances should be one-half that of the internal sections.

49. RC Filter for Small Currents. An economical RC filter for small currents as suggested by Scott[1] is shown in Fig. 57. An especial feature is the shunting circuit through R_A to feed voltage 180 deg. out of phase to the point X. This can be adjusted to give a very high attenuation at one particular frequency which it is desired to eliminate. As shown, this filter is low pass; a similar high-pass structure can be made by transposing the R's and C's.

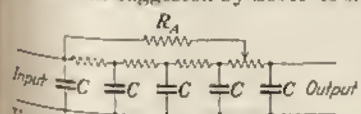


FIG. 57.—RC filter for small currents.

50. Resistance Pads. Resistance pads are artificial lines whose series and shunt elements are pure resistances and are used principally

[1] SCOTT, H. H., *Electronics*, August, 1939.

TABLE II.—CHARACTERISTICS OF NON-LOADED AUDIO-FREQUENCY CIRCUITS (Per loop mile at 1,000 cycles)

Type of circuit	R Ohms	L Henrys	G μ mhos	C Micro-farads	Z Ohms	λ Miles	V Miles per second	A db per mile	B Radians per second
No. 10 open-wire NL	10.4	0.00394	0.8	0.0078	739	177	176,600	0.65	0.0356
No. 16 cable NL	42.2	0.001	0.87	0.062	331	64.5	64,550	0.73	0.0975
No. 19 cable NL	83.2	0.001	0.87	0.062	462	47.5	47,500	1.065	0.1322
No. 22 cable NL	171	0.001	1.75	0.073	610	31.7	31,700	1.72	0.198

as attenuators in a-f circuits. The amount of loss caused by insertion of a pad in a circuit may be accurately computed and is independent of frequency if the terminating impedances are resistances.

Either π or T structures may be used as pads, as shown in Fig. 58*a*. Both are electrically equivalent, but for identical values of loss and impedance one type may require resistors of more convenient values than the other. A pad which is to be used in a circuit that is balanced to ground should be of the balanced π or T type; otherwise the unbalanced network is satisfactory and requires several less resistors to build.

51. Pad Design. To design a pad, three constants must be known: the input and output impedances and the loss in decibels. The input and output impedances of a pad are usually made equal to those of the circuit to be connected to it. The design procedure depends upon whether these are equal or are different from each other.

1. *Equal Input and Output Impedances.* In this case, the value of each element is found by multiplying the proper constants selected from Table III in connection with Fig. 58*a*, by the value of the input or output impedance Z in ohms.

Example: To design a 10-db, 500/500-ohm pad of the balanced T type: From Table III for 10-db attenuation, $a = 0.5195$ (hence $a/2 = 0.2597$) and $b = 0.7027$. Then the required resistances are $0.2597 \times 500 = 129.85$ for the series elements and $0.7027 \times 500 = 351.35$ ohms for the shunt element.

2. *Unequal Input and Output Impedances.* In this case, the design involves more computation. The value of each element is indicated by Fig. 58*b*, the constants of which are to be found in Table III. The ratio of input to output impedance (or vice versa) of a pad of given loss is limited by the fact that for large values of the impedance ratio certain of the pad resistors would have to be negative in value if the loss of the pad were to be below a certain minimum value. The maximum impedance ratio which a 10-db pad can have, for example, is 3.018. Stated in another way, this means that, if the impedance ratio of a pad is to be 3.018, its loss must be at least 10 db. The maximum impedance ratios for various values of pad losses are also given in Table III. These are the same for both π and T pads.

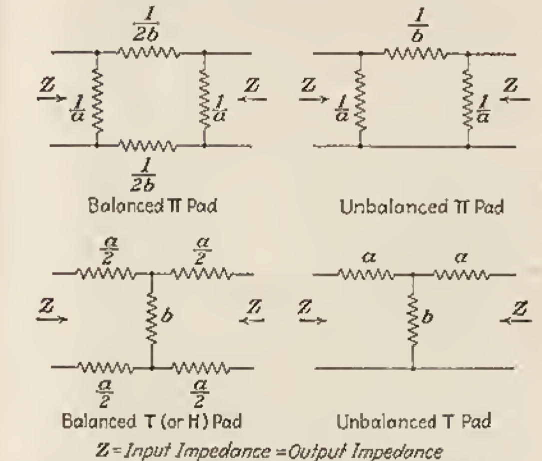


FIG. 58*a*.—Equivalent balanced and unbalanced pads.

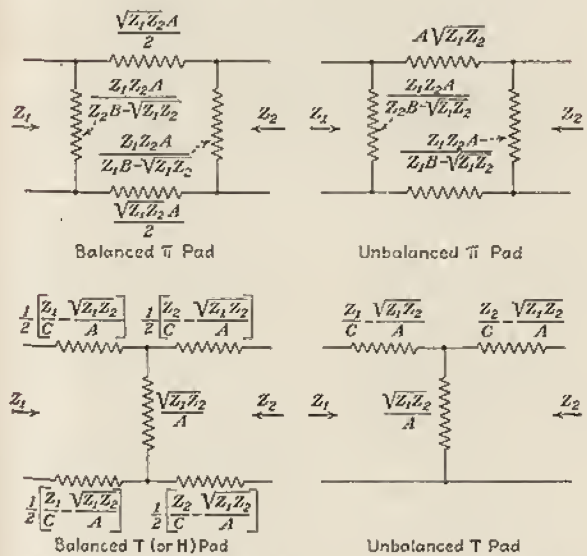


FIG. 58*b*.—Pads to be used between unequal impedances.

TABLE III.—CONSTANTS FOR PADS OF FIG. 58

Loss, decibels	A	B	C	a	b	$1/b$	$1/a$	$1/2b$	Maximum ratio Z_1/Z_2 or Z_2/Z_1
1	0.1154	1.007	0.1150	0.0575	8.664	0.1154	17.39	0.0577	1.01[illegible]
2	0.2323	1.027	0.2263	0.1146	4.305	0.2323	8.724	0.1161	1.05[illegible]
3	0.3523	1.060	0.3325	0.1710	2.838	0.3523	5.848	0.1761	1.12[illegible]
4	0.4770	1.108	0.4305	0.2263	2.097	0.4770	4.410	0.2385	1.22[illegible]
5	0.6084	1.170	0.5192	0.2801	1.645	0.6084	3.570	0.3042	1.30[illegible]
6	0.7472	1.248	0.5986	0.3323	1.330	0.7472	3.009	0.3736	1.35[illegible]
7	0.8960	1.343	0.6673	0.3825	1.116	0.8960	2.615	0.4480	1.50[illegible]
8	1.0570	1.455	0.7264	0.4305	0.9462	1.0570	2.323	0.5285	2.11[illegible]
9	1.2320	1.586	0.7763	0.4762	0.8118	1.2320	2.100	0.6160	2.[illegible]
10	1.4218	1.738	0.8181	0.5195	0.7027	1.4218	1.925	0.7109	3.[illegible]
11	1.6324	1.914	0.8527	0.5601	0.6127	1.6324	1.785	0.8162	3.[illegible]
12	1.8059	2.117	0.8814	0.5986	0.5359	1.8659	1.670	0.9329	4.[illegible]
13	2.1223	2.346	0.9046	0.6343	0.4712	2.1223	1.576	1.0611	5.[illegible]
14	2.4067	2.605	0.9235	0.6672	0.4155	2.4067	1.498	1.2033	6.[illegible]
15	2.7230	2.901	0.9387	0.6981	0.3672	2.7230	1.432	1.3615	8.41[illegible]
20	4.9522	5.052	0.9802	0.8182	0.2020	4.9522	1.222	2.4761	25.[illegible]
25	8.8612	8.918	0.9940	0.8932	0.1128	8.8612	1.119	4.4306	79.[illegible]
30	15.800	15.830	0.9980	0.9387	0.06331	15.800	1.065	7.900	250.[illegible]
35	28.004	28.112	0.9994	0.9649	0.03560	28.094	1.036	14.047	790.[illegible]
40	50.000	50.0094	0.9998	0.9802	0.020000	50.000	1.020	25.000	2,500
45	88.928	88.933	0.9999	0.9888	0.01124	88.928	1.011	44.464	7,909
50	158.1	158.102	1.0000	0.9937	0.006325	158.10	1.006	79.050	24,950
60	500	500	1.0000	0.9980	0.002000	500	1.002	250	
70	1,581	1,581	1.0000	0.9994	0.000632	1,581	1.001	790	
80	5,000	5,000	1.0000	0.9998	0.000200	5,000	1.000	2,500	
90	15,810	15,810	1.0000	0.9999	0.0000632	15,810	1.000	7,905	
100	50,000	50,000	1.0000	1.0000	0.0000200	50,000	1.000	25,000	

$$A = \sinh\theta \qquad a = \frac{1}{C} - \frac{1}{A}$$

$$B = \cosh\theta$$

$$C = \tanh\theta \qquad b = \frac{1}{A}$$

$$\theta = \frac{\text{loss in decibels}}{8.686} \qquad \text{Maximum ratio } \frac{Z_1}{Z_2} \text{ or } \frac{Z_2}{Z_1} = B^2$$

Example: To design a 20-db 500/200-ohm pad of the unbalanced π type:

$$Z_1 = 500 \text{ ohms}, \qquad Z_2 = 200 \text{ ohms}$$

From Table III, $A = 4.9522$ and $B = 5.0522$. Then,

$$\text{Input shunt element} = \frac{Z_1 Z_2 A}{Z_2 B - \sqrt{Z_1 Z_2}} = 713 \text{ ohms}$$

$$\text{Series element} = \sqrt{Z_1 Z_2}\,A = 1{,}567 \text{ ohms}$$

$$\text{Output shunt element} = \frac{Z_1 Z_2 A}{Z_1 B - \sqrt{Z_1 Z_2}} = 430 \text{ ohms}$$

52. Characteristic Impedance of R-f Line. At high frequencies R and G usually become negligible as compared with ωL and ωC, respectively. The characteristic impedance of a line at radio frequencies is then

$$Z_0 = \sqrt{\frac{L}{C}} \text{ ohms} \tag{74}$$

where L and C are in henrys and farads per unit length.

1. *Special Case: Line of Two Parallel Wires.* In terms of the dimensions of the line

$$Z_0 = 277 \log_{10} \frac{2s}{d} \text{ ohms} \tag{75}$$

for parallel wire, where s is the spacing from center to center of the wires and d the diameter, both being measured in the same units. Equation (74) is based on the assumption that s is at least ten times d and that the height of the line above the ground is at least ten times s. The characteristic impedances of open-wire r-f lines of commonly used dimensions are shown in Fig. 59.

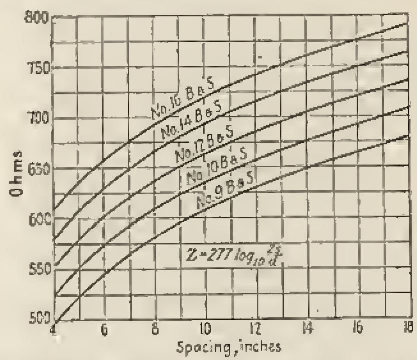


FIG. 59.—Characteristic impedance of open-wire r-f transmission line.

2. *Special Case: Line of Two Coaxial Conductors.* Radio-frequency lines are often constructed with one conductor in the form of a metal tube and the other a coaxially placed wire or tube of smaller diameter. The advantage of such construction lies principally in the effective shielding that can be obtained by grounding the outer tube.

The characteristic impedance of a line having such coaxial conductors is

$$Z_0 = 138.5 \log_{10} \frac{r_o}{r_i} \text{ ohms} \tag{76}$$

where r_o is the inside radius of the outer tube, and r_i is the outside radius of the inner conductor. For a line whose outer and inner conductors are respectively ¾ and ¼ in. in diameter, $Z_0 = 65$ ohms.

53. Other Properties of R-F Lines.

Velocity of propagation is

$$V \doteq \frac{1}{\sqrt{L_1 C_2}} \doteq 186{,}000 \text{ miles per second} \tag{77}$$

Wave-length constant is

$$B = \omega\sqrt{L_1C_2} \text{ radians per unit length} \quad (78)$$

$$= \frac{\omega}{186{,}000} \text{ radians per mile} \quad (79)$$

Wave length is

$$\lambda = \frac{2\pi}{\omega\sqrt{L_1C_2}} = \frac{1}{f\sqrt{L_1C_2}} \text{ unit lengths} \quad (80)$$

$$= \frac{186{,}000}{f} \text{ miles} \quad (81)$$

$$= \frac{300{,}000{,}000}{f} \text{ meters} \quad (82)$$

Retardation time is

$$t = \sqrt{L_1C_2} \text{ sec. per unit length} \quad (83)$$

$$= 5.39 \times 10^{-6} \text{ sec. per mile} \quad (84)$$

Attenuation constant is

$$A = 4.346R\sqrt{\frac{C}{L}} \text{ db per unit length} \quad (85)$$

For *parallel wires* this becomes

$$A = \frac{0.0157R}{\log_{10}\frac{2s}{d}} \text{ db per unit length} \quad (86)$$

where R = loop resistance per unit length
s = spacing of wires, center to center
d = diameter of each wire, s and d being measured in the same units

For *coaxial conductors*, the attenuation is

$$A = \frac{0.0314R}{\log_{10}\frac{r_o}{r_i}} \text{ db per unit length} \quad (87)$$

where R = loop resistance (sum of the resistance of the two conductors)
r_o = radius of outer tube
r_i = radius of inner conductor, r_0 and r_i being measured in the same units.

54. Input Impedance of Line Terminated in Impedance Z_a at Its Far End. *Special Cases for Radio Frequencies.* At high frequencies the attenuation constant A of a line approaches zero, and the propagation constant is nearly equal to the wave-length constant B,

$$P \doteq jB \doteq j\omega\sqrt{LC} \quad (88)$$

and from Eq. (69)

$$\theta = lP = jlB = j\omega l\sqrt{LC} \quad (89)$$

Then Eq. (68) becomes

$$Z_i = Z_0\left[\frac{Z_a \cos lB + jZ_0 \sin lB}{Z_0 \cos lB + jZ_a \sin lB}\right] \text{ ohms} \quad (90)$$

This input impedance has certain interesting and useful values when the length of the line is a multiple of a quarter- or half-wave length.

1. *Lines Quarter-wave Length Long.* In this case

$$l = \frac{\lambda}{4}, \; B = \frac{2\pi}{\lambda}, \text{ and } lB = \frac{\pi}{2}$$

Then (90) reduces to

$$Z_i = \frac{Z_0^2}{Z_a} \text{ ohms} \quad (91)$$

Owing to this property quarter-wave lines are made use of as impedance-matching transformers. If, for example, a line whose characteristic impedance is Z_1 is to be connected to an antenna system whose input impedance is Z_2, a quarter-wave line having characteristic impedance $Z_0 = \sqrt{Z_1Z_2}$ is inserted. Since $Z_2 = Z_a$, the impedance facing the line is $Z_i = Z_1Z_2/Z_2 = Z_1$ ohms, and the impedance facing the antenna is $Z_i' = Z_1Z_2/Z_1 = Z_2$ ohms, which results in a perfect impedance match at each junction.

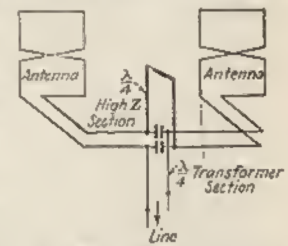


Fig. 60.—Use of quarter-wave short-circuited line to by-pass low-frequency currents for sleet melting without disturbing the r-f impedance of the system.

Quarter-wave Line Short-circuited at Far End. In this case, $Z_a = 0$, and $Z_i = \infty$. Such a line is thus antiresonant at the radio frequency corresponding to four times its length and is often used in antenna systems to by-pass low-frequency current around large r-f impedances, for melting sleet. Such a use is illustrated in Fig. 60.

Quarter-wave Line Open-circuited at the Far End. In this case $Z_a = \infty$, and $Z_i = 0$. Such a line thus has practically no impedance at the r.f. which corresponds to four times its length.

Half-wave Line Terminated in Impedance Z at Far End. Here, $l = \lambda/2$ and $lB = \pi$. Consequently Eq. (90) becomes

$$Z_i = Z_a \quad (92)$$

Thus the input impedance of a half-wave line is equal to the termination impedance at its far end and is independent of the characteristic impedance of the line.

Lines Whose Lengths Are Integral Multiples of Quarter- or Half-wave Lines. Such lines can be shown to have the same properties as quarter- or half-wave lines, due to the periodicity of the sine and cosine functions in Eq. (90).

55. Termination Impedances at Radio Frequencies. At r.f. proper termination of lines is even more important than at a.f., since reflection resulting from mismatched impedances at the junctions produces standing waves which in turn cause radiation along the line and a decrease in efficiency. Impedance irregularities in a line also tend to set up reflections, and bends in the line should therefore be gradual, with a minimum radius of about one-fourth wave length. For the same reason the line should be kept free (at least one-fourth wave length) from large masses of conducting or dielectric materials.

56. Efficiency of Lines at Radio Frequencies. In a properly constructed and terminated line the power losses are practically all due to the inherent ohmic resistance of the line, and the efficiency may be fairly high. For ordinary designs, the efficiency is approximately

$$(100 - 2l) \text{ per cent} \tag{93}$$

where l is the length of the line in wave lengths.

57. Tapered Lines as Impedance Transformers. A gradual smooth change with length in the inductance and capacity of a line causes the characteristic impedance to vary along the line and can be shown to introduce no reflections. Consequently a section of line with variable spacing or diameter of the wires is, like the quarter-wave-length line, a useful impedance matching transformer, the dimensions being so chosen that the end impedances of the line equal their respective terminating impedances.

WAVE FILTERS

58. Wave filters are forms of artificial lines, such as those of Fig. 51b and c, purposely designed to transmit efficiently current in a desired band of frequencies and more or less completely to suppress all other frequencies. The boundary frequencies between transmission bands and attenuation bands are called *cutoff frequencies.*

The following brief discussion of wave-filter design is intended to serve as a guide to the design of simple filters for use where the requirements are not very severe. For complete information concerning the design of filters to meet more exacting specifications, the references listed in the bibliography at the end of this section should be consulted.

Filters are divided into four classes, according to the frequency bands which they are intended to transmit, *viz., low pass, high pass, band pass,* and *band elimination.*

59. Losses in Filters, and Effects of Dissipation. The elements of ideal wave filters are always pure reactances; practically, however, some dissipation must always be tolerated owing to the resistance of coils and condensers, but this is made as small as possible by employing high-Q elements.

The terminating impedances of a filter are usually resistances equal in value to the image impedances of the filter. Then the loss within the transmitted bands (except near the cutoff frequency) is mainly due to dissipation in the elements and is usually small. In the vicinity of cutoff and the point of maximum attenuation, the total insertion loss of a filter involves the reflection and interaction losses as well as the attenuation. The loss elsewhere in the attenuated bands is very nearly the sum of the attenuation constants of the various sections, minus a gain of approximately 6 db which is due to reflections resulting from impedance mismatches occurring in these regions. Methods for the exact calculation of filter losses are beyond the scope of this handbook but are available in the published works of Zobel, Johnson, and Shea.

60. The Basic Filter Section. The basis of filter design is the full L section, consisting of a series element Z_1 and a shunt element Z_2 as shown at L in Fig. 61. The relation of such a section to an infinite line is also indicated. In a wave filter, where the number of sections is finite and small instead of infinite, symmetrical sections are used. These are either

T or π networks as shown at b and c in Fig. 61. The T section may be considered as being cut from the infinite line (Fig. 61a) at the mid-points (a-a) of two consecutive series elements Z_1 and is said to be "mid-series terminated." The π section may be considered as being cut at the mid-points (b-b) of two consecutive shunt elements and is said to be "mid-

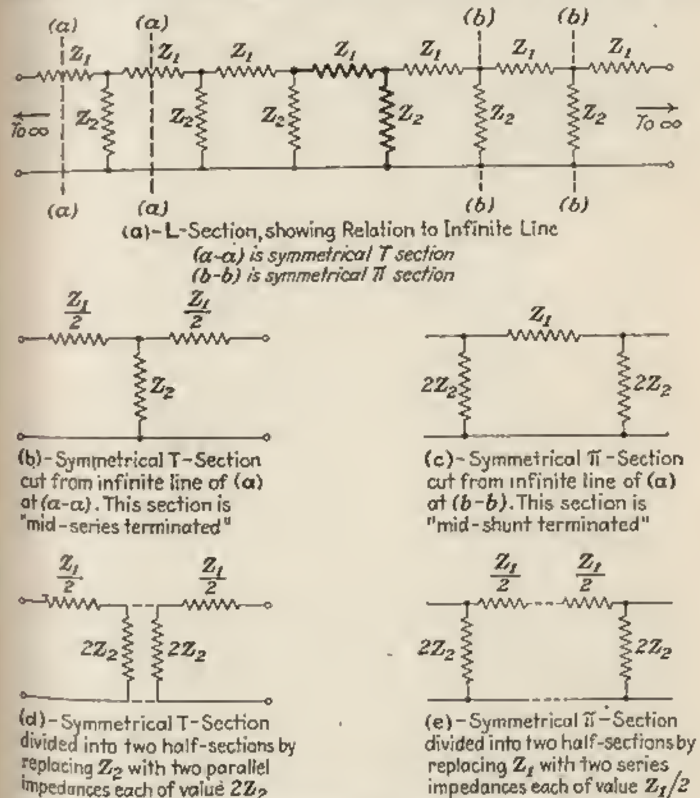


FIG. 61.—Equivalence of T and π networks.

shunt terminated." (To form a mid-shunt termination, each full-shunt element is replaced by an equivalent two impedances in parallel, each of value $2Z_2$.) Either a T or π section may be divided into pairs of equivalent half sections as shown at d and e in Fig. 61.

61. Types of Sections. 1. *Constant-K Sections.* The simplest and most common type of filter section is that in which the impedances Z_1 and Z_2 are so related that their product is a constant

$$Z_1 \times Z_2 = K^2$$

at all frequencies. From this it derives its name "constant-K" section. The configuration and circuit constants of the four classes of constant-K sections are shown in the filter-design formulas in Art. 65. The image impedances of mid-series and mid-shunt terminated constant-K sections within the transmission bands are functions of frequency, but each approaches the value K at some frequency within the band. The value K is therefore taken as the nominal resistance of the constant-K section for design purposes. If a constant-K section is used with one or both of its terminals connected to a pure resistance of value $R = K$, the impedances will be mismatched for all frequencies within the transmitted band except one, and the actual insertion or transmission loss of the filter will be increased by reflection losses at the terminations. This causes an even more gradual cutoff for the constant-K section than its attenuation curve would indicate.

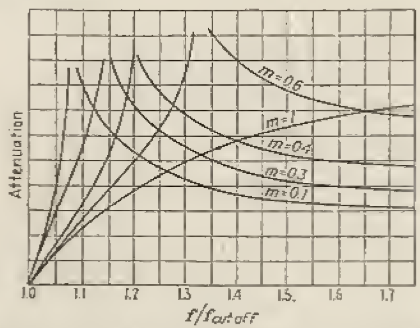


FIG. 62.—Effect of m upon sharpness of cutoff in a low-pass filter structure.

2. *m-Derived Sections.* In many filters, a sharper cutoff than that given by a constant-K type of structure is required. Such a characteristic may be realized in the so-called *m-derived section*, which is due to Otto J. Zobel.[1] This type of section is derived from the constant-K section as a prototype but is made to have sharper cutoff than the prototype by the addition of impedance elements in either the shunt or series arms so that infinite attenuation occurs at some frequency beyond cutoff. Each impedance of the m-type section is related to those of the constant-K section by a factor which is a function of a constant m. The latter is in turn a function of the ratio between the frequency of infinite attenuation and the cutoff frequency and may have any value between 0 and plus 1. The sharpness of cutoff increases as m approaches 0. This effect is illustrated in Fig. 62 for various values of m. It will be noted that, when m is equal to 1, the structure is identical with the constant-K structure. Also, from Fig. 62, it appears that from the viewpoint of obtaining a uniform degree of attenuation throughout the attenuated band the combination of a constant-K section ($m = 1$) (having gradual cutoff but large attenuation remote from cutoff) with one having a small value of m and sharp cutoff ($m = 0.3$, for example)

[1] *Bell System Tech. Jour.*, January, 1923.

would be desirable. This principle is valuable in the design of composite filters.

3. *Shunt-derived and Series-derived m Sections.* Two forms of m-derived sections exist; if the extra impedance is added to the shunt arm, the section is called *series derived*, while, if it is added to the series arm, the section is called *shunt derived*. (See illustrations of derived sections under Filter-design Formulas, Art. 65.)

62. Assembly of Sections into Filters. A filter may consist of any number of sections from a single one-half section to five or six full sections, depending on the amount of attenuation of unwanted frequencies required. The *amount* of attenuation in the rejected band depends upon

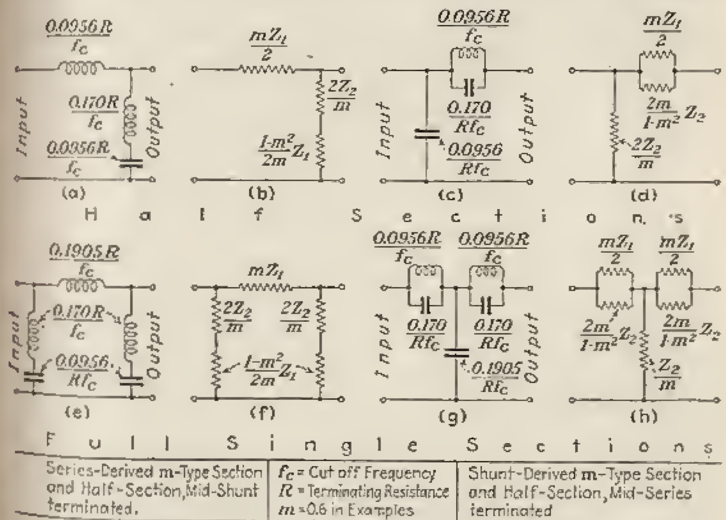


FIG. 63.—Half-section compared with full-section structures.

the number of filter sections used, while the *shape* of the transmission curve depends upon the types of sections employed.

63. One-half- and One-section Filters. If a half section or one full section is used alone as a filter and the requirements regarding the cutoff are not too sharp, an m-derived section is usually preferable, with $m = 0.6$. This will provide the best impedance match with resistance terminations. Either of the structures shown in Fig. 63 is suitable for use with terminations of resistance R.

64. Multi-section Filters. Filters having more than one section are of two types:

A *uniform* filter is one in which all sections are identical with the exception of the end sections. The latter are ordinarily half sections suitable for connecting the filter to its terminating resistances.

A *composite* filter is one made up of two or more sections having different characteristics, each of which is designed to contribute some especial property to the characteristic of the filter as a whole. For example, one section which has sharp cutoff but a diminishing attenuation beyond cutoff may be combined with another section having a gradual cutoff and increasing attenuation beyond as shown at I and II in Fig. 64. The resulting composite structure will then have both sharp cutoff and high attenuation beyond, as shown at III. In general, constant-K sections have gradual cutoffs with increasing attenuation beyond, while m-sections with small values of m have the sharpest cutoff characteristics. Still other types of sections may be added to match impedances at the junctions of the filter and its terminating resistances, or to further alter the transmission characteristics.

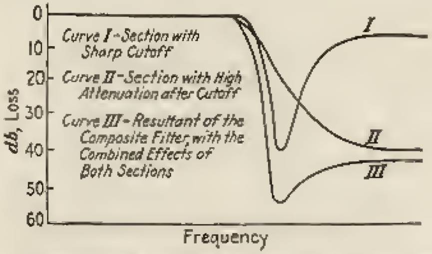


Fig. 64.—Transmission curves for composite low-pass filter.

In a composite filter it is essential that the image impedances be matched at each junction of the component sections, to avoid reflection losses which would impair the transmission curve of the filter. Likewise, the end terminations of the filter should as nearly as possible match the terminating resistances. One of the principal advantages of the m-type structure is that its image impedances can be made identical with other m-type sections or with constant-K sections; or they can be made to approximate resistances over the transmission band for terminating purposes. A complete analysis of the impedance conditions within a wave filter is not possible in the limited space available here but may be found in the References listed at the end of this section. The following will suffice as working rules in designing simple filters for ordinary requirements:

End Terminations. *Resistance.* A mid-shunt termination of a series-derived m-type section or half section, or a mid-series termination of a shunt-derived section or half section, with $m = 0.6$ in either case.

For Parallel or Series Connection with Other Filters. An 0.8-series constant-K section or half section (*i.e.*, one terminated in a series arm equal to 0.8 of a full series arm, Z_1).

Here, as well as in the two preceding paragraphs, the image impedance of the internal section next to the end section in either case must match the image impedance at the inner terminals of the latter, in accordance with the following.

Internal Junctions. The following terminations of the types of filter sections for which formulas are given in Art. 65 may be joined together without impedance mismatches at the junction points:

Mid-series termination of constant-K type to mid-series termination of series-derived m type.

Mid-shunt termination of constant-K type to mid-shunt termination of shunt-derived m type.

Mid-series termination of constant-K, series-derived m type or shunt-derived m type, to mid-series termination of another section of the same type.

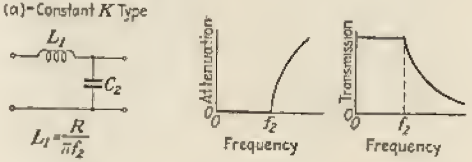
Mid-shunt termination of constant-K, series-derived m type or shunt-derived m type, to mid-shunt termination of another section of the same type.

(Note. In the latter two cases, the values of m in the two sections to be joined, if they are of the m type, may be, and frequently are, different. Both sections must be of the same type and termination, however.)

65. Filter-design Formulas. Formulas for calculating the capacities and inductances of constant-K, series-derived m-type and shunt-derived m-type basic sections are given in the following pages. These are expressed in terms of R, the terminating resistances, the factor m, and the values of f_c, the cutoff frequency, and other critical frequencies. These factors must be predetermined on the basis of the filter requirements and the considerations outlined above.

I. LOW PASS FILTERS

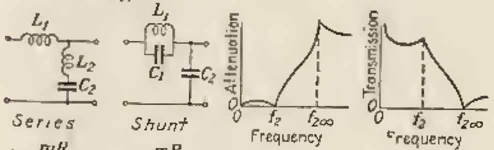
(a)—Constant K Type



$$L_1 = \frac{R}{\pi f_2}$$

$$C_2 = \frac{1}{\pi f_2 R}$$

(b) m—Derived Type



Series:

$$L_1 = \frac{mR}{\pi f_2} \qquad L_2 = \frac{(1-m^2)R}{4m\pi f_2} \qquad C_2 = \frac{m}{\pi f_2 R}$$

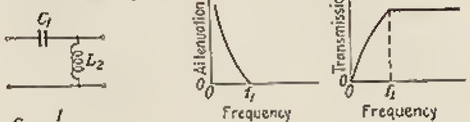
Shunt:

$$L_1 = \frac{mR}{\pi f_2} \qquad C_1 = \frac{(1-m^2)}{4m\pi f_2 R} \qquad C_2 = \frac{m}{\pi f_2 R}$$

$$m = \sqrt{1 - \frac{f_2^2}{f_{2\infty}^2}}$$

II - HIGH PASS FILTERS

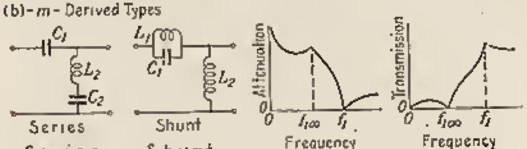
(a)-Constant *K* Type



$C_1 = \frac{1}{4\pi f_1 R}$

$L_2 = \frac{R}{4\pi f_1}$

(b)-*m*-Derived Types

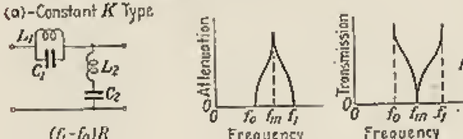


Series	Shunt
$C_1 = \frac{1}{4\pi f_1 m R}$	$L_1 = \frac{mR}{(1-m^2)\pi f_1}$
$L_2 = \frac{R}{4\pi f_1 m}$	$C_1 = \frac{1}{4\pi f_1 m R}$
$C_2 = \frac{m}{(1-m^2)\pi f_1 R}$	$L_2 = \frac{R}{4\pi f_1 m}$

$$m = \sqrt{1 - \frac{f_{1\infty}^2}{f_1^2}}$$

III - BAND ELIMINATION FILTERS

(a)-Constant *K* Type



$f_m = \sqrt{f_0 f_1}$

$L_1 = \frac{(f_1 - f_0)R}{\pi f_0 f_1}$

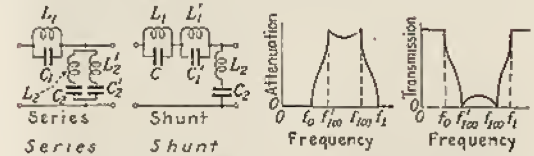
$C_1 = \frac{1}{4\pi(f_1 - f_0)R}$

$L_2 = \frac{R}{4\pi(f_0 - f_1)}$

$C_2 = \frac{f_1 - f_0}{\pi R f_0 f_1}$

III - BAND ELIMINATION FILTERS (continued)

(b)-*m*-Derived Types



Series	Shunt
$L_1 = \frac{mR(f_1 - f_0)}{\pi f_0 f_1}$	$L_1 = \frac{(f_1 - f_0)R}{\pi f_0 f_1 b}$
$C_1 = \frac{1}{4\pi(f_1 - f_0)mR}$	$C_1 = \frac{a}{4\pi(f_1 - f_0)R}$
$L_2 = \frac{aR}{4\pi(f_1 - f_0)}$	$L_1' = \frac{(f_1 - f_0)R}{\pi f_0 f_1 a}$
$C_2 = \frac{(f_1 - f_0)}{\pi f_0 f_1 b R}$	$C_1' = \frac{b}{4\pi(f_1 - f_0)R}$
$L_2' = \frac{bR}{4\pi(f_1 - f_0)}$	$L_2 = \frac{R}{4\pi(f_1 - f_0)m}$
$C_2' = \frac{(f_1 - f_0)}{\pi f_0 f_1 a R}$	$C_2 = \frac{m(f_1 - f_0)}{\pi f_0 f_1 R}$

$$m = \sqrt{\frac{\left(1 - \frac{f_0^2}{f_{1\infty}^2}\right)\left(1 - \frac{f_{1\infty}^2}{f_1^2}\right)}{1 - \frac{f_0}{f_1}}}$$

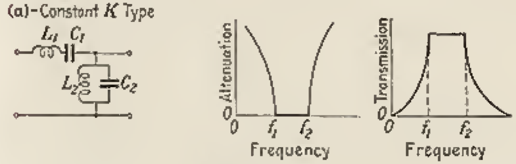
$$a = \frac{1}{m}\left(1 + \frac{f_0 f_1}{f_{1\infty}^2}\right)$$

$$b = \frac{1}{m}\left(1 + \frac{f_{1\infty}^2}{f_0 f_1}\right)$$

$$f_{1\infty}' = \frac{f_0 f_1}{f_{1\infty}}$$

IV. BAND PASS FILTERS

(a)-Constant *K* Type

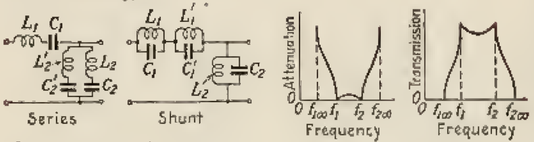


$L_1 = \frac{R}{\pi(f_2 - f_1)}$ $\qquad L_2 = \frac{(f_2 - f_1)R}{4\pi f_2 f_1}$

$C_1 = \frac{(f_2 - f_1)}{4\pi f_2 f_1 R}$ $\qquad C_2 = \frac{1}{\pi(f_2 - f_1)R}$

IV. BAND PASS FILTERS (continued)

(b)-m-Derived Types



Series	Shunt	
$L_1 = \dfrac{mR}{\pi(f_2-f_1)}$	$L_1 = \dfrac{(f_2-f_1)R}{4\pi f_1 f_2 b}$	$m = \dfrac{h}{1-\dfrac{f_1 f_2}{f_{2\infty}^2}}$
$C_1 = \dfrac{(f_2-f_1)}{4\pi f_1 f_2 m R}$	$C_1 = \dfrac{a}{\pi(f_2-f_1)R}$	$h = \sqrt{\left(1-\dfrac{f_1^2}{f_{1\infty}^2}\right)\left(1-\dfrac{f_2^2}{f_{2\infty}^2}\right)}$
$L_2 = \dfrac{aR}{\pi(f_2-f_1)}$	$L_1' = \dfrac{(f_2-f_1)R}{4\pi f_1 f_2 a}$	$a = \dfrac{(1-m^2)f_1 f_2}{4h f_{1\infty}^2}\left(1-\dfrac{f_{1\infty}^2}{f_{2\infty}^2}\right)$
$C_2 = \dfrac{(f_2-f_1)}{4\pi f_1 f_2 b R}$	$C_1' = \dfrac{b}{\pi(f_2-f_1)R}$	$b = \dfrac{(1-m^2)}{4h}\left(1-\dfrac{f_{1\infty}^2}{f_{2\infty}^2}\right)$
$L_2' = \dfrac{bR}{\pi(f_2-f_1)}$	$L_2 = \dfrac{(f_2-f_1)R}{4\pi f_1 f_2 m}$	$f_{1\infty} = \dfrac{f_1 f_2}{f_{2\infty}}$
$C_2' = \dfrac{(f_2-f_1)}{4\pi f_1 f_2 a R}$	$C_2 = \dfrac{m}{\pi(f_2-f_1)R}$	

Examples of Filter Design: 1. *Single-section Filter.* Required: High-pass single-section filter to be connected between resistance terminations of R = 1,000 ohms, with a cutoff frequency of 1,000 cycles and maximum attenuation occurring at 800 cycles.

To secure the attenuation peak at 800 cycles, an m-type filter section is required. Either the shunt- or series-derived type may be used. Choosing the latter, we have from the filter formulas II (b), Art. 65, in which f_1 = 1,000 cycles, $f_{1\infty}$ = 800 cycles, R = 1,000 ohms, and m = 0.6,

$$C_1 = 0.1325 \times 10^{-6} \text{ farad}$$
$$L_2 = 0.1325 \text{ henry}$$
$$C_2 = 0.298 \times 10^{-6} \text{ farad.}$$

From the considerations involving impedance matching at the end terminals a mid-shunt termination facing each resistance termination is seen to be desirable for a series-derived section. Hence the structure of Fig. 63f is indicated. One full-series element (C_1) will be required, with a double-impedance shunt

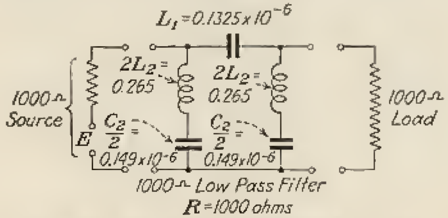


Fig. 65.—Example of single-section filter.

arm ($2L_2 + C_2/2$) at each end. The completed filter will then be as shown in Fig. 65.

2. *Multi-section Composite Filter.* Required: Low-pass filter to be connected between resistance terminations of R = 600 ohms, with sharp cutoff at 1,000 cycles and high attenuation beyond.

There is no unique solution or "best" filter design for this problem. A large number of filters might be designed to meet these requirements, each of which would serve as well as any of the others. The relative merits of different designs will depend upon their economy of coils and condensers in accomplishing the required results. One suitable design is shown here:

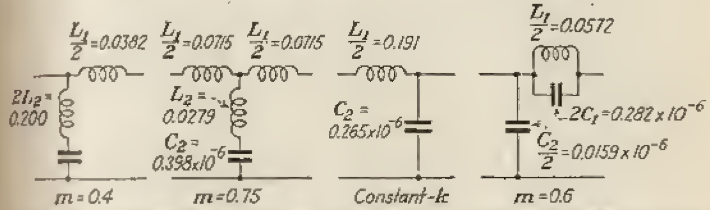


Fig. 66.—Low-pass filter for use between 600 ohms with sharp cutoff at 1,000 cycles.

Let the input-end section be a half-section mid-series-derived m type, with its mid-shunt termination facing the input to match impedances at that point. Let m = 0.4 for this half section to give a sharp cutoff.

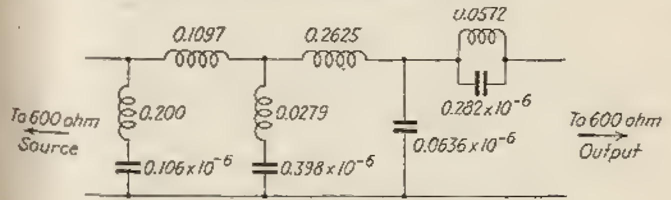


Fig. 67.—Final filter as designed by Fig. 66.

This will be followed by a symmetrical full section of the series-derived m type, mid-series terminated, with m = 0.75. Then a half section of the constant-K type with mid-series termination facing the full section and mid-shunt termination facing the end-terminating half section, which will be shunt-derived m type, with m = 0.6. The latter will have a mid-shunt termination facing the constant-K half section and a mid-series termination facing the output termination.

References

Circuit Theory:
Bryant, Correll and Johnson: "Alternating-current Circuits," McGraw-Hill Book Company, Inc.
Colebrook: "Alternating Currents and Transients," McGraw-Hill Book Company, Inc.
Everitt: "Communication Engineering," McGraw-Hill Book Company, Inc.
Johnson: "Transmission Circuits for Telephonic Communication," D. Van Nostrand Company, Inc.
Kennelly: "Electric Lines and Nets," McGraw-Hill Book Company, Inc.
Morecroft: "Principles of Radio Communication," John Wiley & Sons, Inc.

SHEA: "Transmission Networks and Wave Filters," D. Van Nostrand Company, I
TERMAN: "Radio Engineering," McGraw-Hill Book Company, Inc.

Wave Filters.

CAMPBELL: Physical Theory of Electric Wave Filters, *Bell System Tech. Jour.*, 1, No. 2, 1.
———: Cisoidal Oscillations, *Trans. A.I.E.E.*, April, 1911.
CARSON and ZOBEL: Transient Oscillations in Electric Wave Filters, *Bell System Te Jour.*, July, 1923.
DELLENBAUGH: Electric Filters, *QST*, July and August, 1923.
JOHNSON: "Transmission Circuits for Telephonic Communication," Chaps. XVI a XVII, D. Van Nostrand Company, Inc.
JOHNSON and SHEA: Mutual Inductance in Wave Filters, with an Introduction Filter Design, *Bell System Tech. Jour.*, 4, No. 1, 52.
PIERCE: "Electric Oscillations and Electric Waves," Chap. XVI, McGraw-H Book Company, Inc.
SHEA: "Transmission Networks and Wave Filters," D. Van Nostrand Company, I
ZOBEL: Theory and Design of Uniform and Composite Wave Filters, *Bell System Te Jour.*, 2, No. 1, 1.
———: Transmission Characteristics of Electric Wave Filters, *Bell System Tech. Jou* 3, No. 4, 567.

Transmission Lines:

EVERITT: "Communication Engineering," McGraw-Hill Book Company, Inc.
FELDMAN and STERBA: Transmission Lines for Short-Wave Radio Systems, *Be System Tech. Jour.*, 9, No. 3, 411; and *Proc. I.R.E.*, 20, No. 7.
FRANKLIN and TERMAN: "Transmission Line Theory," Franklin and Charles.
TERMAN: "Radio Engineering," McGraw-Hill Book Company, Inc.

SECTION 7

ELECTRICAL MEASUREMENTS

BY R. F. FIELD[1] AND JOHN H. MILLER[2]

True basic measurements of electrical quantities are rarely made except in standardizing laboratories, owing to the inherent difficulties in the procedure. Ordinary measurements are made by comparison devices of one form or another. Direct-reading instruments, having an electrical torque-producing means functioning against a spring, are calibrated against accurate standards which are in turn calibrated against basic measuring devices. Such torque-producing instruments are used for measuring current, voltage, power, and resistance. Instruments for measuring phase relations, frequencies, and other factors may have two torque-producing systems, each torque varying with the position of the moving element and bearing different functional relations to the quantity measured. The result is for the moving system carrying the pointer to take up a position where the torques balance, this being different for each different value of the quantity in question, and the scale may be marked accordingly.

STANDARDS

1. Current. Current is measured, absolutely, in terms of the force of attraction or repulsion between two coils connected in series and carrying that current, and the various dimensions of the coils. This current is then used to deposit silver in the *silver voltammeter* to determine the electrochemical equivalent of silver. The silver voltammeter is thus the standard of current. One ampere of continuous unvarying current will deposit 0.001118 g of silver per second when following the standard procedure. The use of this standard is tedious and time consuming, and it is generally used only for the exact calibration of a standard cell and a known resistance.

2. Resistance. Resistance is measured absolutely by a number of methods in terms of a speed of revolution of a disk or coil and its various dimensions. The resistance is then compared with a mercury column of uniform cross section by a suitable bridge method. Such a column of mercury, having a mass of 14.4521 g, a uniform cross section (practically equivalent to 1 sq mm) of a length of 106.3 cm, and at the temperature of melting ice, has a resistance of 1 ohm. Practical secondary standards are coils of manganin wire immersed in oil and sealed in metal containers. Such sealed standards built by Leeds & Northrup Company to the specifications of the U. S. Bureau of Standards are adjusted to an accuracy of 0.01 per cent and may be relied upon to

[1] General Radio Company, Inc., Cambridge, Mass.
[2] Weston Electrical Instrument Company, Newark, N. J.

hold their calibration to 1 part in 100,000 for considerable periods time. The sealing of the containers is important to prevent the absor tion, by the oil, of moisture from the atmosphere, for such moisture wou deposit upon the shellac or other insulating material on the wire whi in turn, will cause mechanical strains to distort the values beyond norm expectancy.

3. Voltage. Voltage measurements cannot be measured absolut with an accuracy sufficient to make the measurement desirable, account of the smallness of the electrostatic forces involved. The secon ary standard of voltage is the saturated cadmium or Weston cell.

These cells, as built by Weston and by the Eppley Laboratory, a correct to 0.01 per cent. They may be depended upon to hold th voltage to 1 part in 100,000 when proper correction for temperatu is made. The unsaturated cadmium cell must be compared with t saturated type for its initial calibration. Its temperature coeffici is negligible. Its voltage is constant to 1 part in 10,000.

As stated above, the cell is calibrated basically in terms of the standa mercury ohm and the standard ampere as obtained by the *silver-vol ammeter* method.

4. Reactance. The self and mutual inductance of single-layer air-co coils and the capacitance of two-plate air condensers having gu rings may be calculated from their dimensions, with an accuracy better than 2 parts in 100,000.

5. Frequency. The absolute standard of frequency is the m solar day as measured by astronomical observations. The mechani vibrations of piezoelectric quartz crystals or of tuning forks made fr carefully stabilized metals provide standards of frequency when perm nently connected into suitable vacuum-tube circuits and allowed oscillate continuously at constant temperature. Over long periods time their frequency is constant to better than 1 part in 1,000,00 recent advances indicate a stability of 2 parts in 10,000,000 is obtainab The frequency of the crystal with which such accuracy may be attain is restricted to the neighborhood of 100 kc. Tuning-fork standard usually operate at 1,000 cycles. By means of suitable frequency mult pliers and dividers all other frequencies from 1 cycle to 100 Mc may b obtained with the same accuracy.

Quartz crystals whose frequencies remain constant to 5 parts in 1,000, 000 may be made for the frequency range 20 kc to 10 Mc. Metal such as nickel and certain iron alloys, having the property of magnet striction, may be used as oscillators in suitable vacuum-tube circui Their frequency range extends from 5 to 100 kc. Their stability about 2 parts in 100,000. For the lower frequencies tuning forks an metal bars are used. Their frequency range is 25 to 1,000 cycles.

CURRENT-MEASURING INSTRUMENTS

6. Moving-coil permanent-magnet instruments of the pointer typ or reflecting galvanometers, consist of a coil, usually wound on a met frame for damping purposes, which can rotate in an intense uniform magnetic field produced by a permanent magnet.

The current I flowing through the turns N of the coil reacts with th magnetic field H in the air gap to produce a force F acting on each conduct proportional to the product IHl of the current, magnetic field, and length o conductor in the field. If the coil is pivoted at its center, a torque will b

exerted, tending to rotate the coil about an axis parallel to the sides of the coil and perpendicular to the magnetic field. Some kind of restoring torque is provided which is proportional to the angle θ through which the coil rotates. Expressing the sensitivity S of the instrument as the angular deflection per unit current, it is given by

$$S = \frac{\theta}{I} = \frac{HNlb}{\tau} \tag{1}$$

where b is the diameter of the coil and τ is the restoring torque per unit angular displacement. For maximum sensitivity as a galvanometer, the permanent magnet should be very strong and the restoring force very weak. However, for pointer-type indicating instruments swung on pivots between sapphire V jewels, there is a minimum torque which may be used for a given moving element weight in order that frictional effects will be unobservable. For instruments mounted on a switchboard and having a horizontal axis, the ratio of the full-scale torque in milligram-centimeters with the weight in grams should not be less than 40 for small instruments, 60 for larger instruments of 1 per cent accuracy, and still greater if greater accuracy is required.

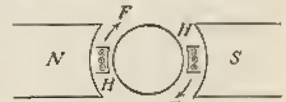


FIG. 1.—Moving-coil galvanometer.

For portable instruments having a vertical axis, it has been found that heavy elements, over 1 g, show greater friction than given by the above relation, and lighter elements show less friction. Hence for such vertical axis instruments for portable service the torque/weight$^{3/2}$ ratio is used and this ratio should be over 40 for small instruments and over 60 to 100 for large instruments for unobservable friction. Ratios much lower than this may be satisfactory for highly sensitive laboratory instruments used with care and not subject to vibration or handling.

The magnetic field obtained from the permanent magnet must be constant so that the electrical characteristics of the instrument may remain unchanged. The constancy of a magnetic system is determined by the ratio K, which is equal to the product of the effective length of the magnet times the effective cross section of one of the air gaps, divided by the product of the cross section of the magnet and the total air-gap length. This constant should be over 100 for chrome and tungsten magnet steels and over 30 for high cobalt steels. For the various nickel-aluminum or MK steels the constant will vary, but 10 may be taken as a median value. Tungsten and chrome steels are most generally used; high cobalt steels will cost two to three times as much but can be made somewhat smaller and will give increased flux, which may be very valuable for aircraft instruments and where the utmost in sensitivity is required. Nickel-aluminum steels require such radical redesigns for efficient use that at this time their use is rather limited; the use of these steels in future designs may be expected to increase materially. Fabrication cost is high and over-all cost is probably governed more by the method of use than by the material itself.

The flux density in the air gap is between 500 and 2,500 gauss. The structure of a pole piece and a core is used to decrease the length of the air gap and to make the magnetic flux uniform and radial. Where distorted d-c scales are required to balance other factors such as decibel relations, the pole tips may be cut away to produce a markedly distorted field resulting in a more uniform scale for the quantity measured.

The deflection of any sensitive galvanometer is indicated by the angu rotation of a beam of light, the so-called *optical lever*, which is reflec from a mirror, either plane or convex, mounted above the moving c The older form of telescope and scale is now being replaced by a spot light containing cross hairs which moves along a scale. The use of spot of light is much less fatiguing than observation through a telescop and a wider range of view is obtained. The usual scale length is 50 c with zero in the center. The standard distance from mirror to scale 1 meter. The maximum angular deflection is about 14 deg. Practica all pivot instruments use pointers. Full-scale deflection corresponds approximately 90 deg. This is increased to 120 deg. in some centr station meters by careful shaping of the pole pieces. It may be increas to 270 deg. by a radical change in design.

The moving element of every deflection instrument provided w a restoring torque proportional to the angular deflection is in effe a torsional pendulum. As such it has a moment of inertia P, a peri T, and a damping factor. If the damping factor is low, the instrum will oscillate several times about its position of rest, each oscillati being less than the preceding one in accordance with the decrement of t system. For most rapid indication it is desirable that the instrume be not quite aperiodic or deadbeat but rather that it overswing fr 3 to 5 per cent. (For a complete discussion of this see Drysdale an Jolley, "Electrical Measuring Instruments," Vol. 1, Chap. 3, Conditio for Rapid Indication.)

Normal ammeters and voltmeters may be expected to have a peri of the order of 1 to 2 sec. The smaller instruments, if equipped wi magnets for very high gap densities and extremely light moving elemen may have a period as short as 0.2 sec. (Weston high-speed power-le indicators.) Instruments of ultrahigh sensitivity, where very litt energy is available, may have a period as high as 5 sec. Sensiti suspension galvanometers may have a period as long as 12 sec.

The period of an instrument is important because the time necess for any deflection instrument to attain a new position when its deflecti force is altered cannot be less than its period. High-speed indicat in indicating instruments is very desirable, particularly when the ph nomena being observed are rapidly changing, as in the monitoring voice-frequency circuits; instruments with a long period will integra the energy while high-speed instruments will give indications of peaks.

The friction of the suspension and the surrounding air is not suffici to prevent the moving coil oscillating back and forth about its equilibriu position when a deflecting force is applied. The amount of damping measured by the rate at which the amplitude of the oscillations decreas The ratio of any two successive swings is constant. The Napierian hyperbolic logarithm of this ratio is called the *logarithmic decrement* of th instrument. The smallest amount of damping which will cause the coil t come to rest with no oscillation whatever is called the *critical damping* and the coil is said to be critically damped. Increasing the dampin beyond this point increases the time necessary for the coil to come to re and produces overdamping. The shortest time in which the coil ca come within a given small distance of its position of rest occurs when th coil is slightly underdamped. It has a value of about 1.5 times th period of the coil. The extra damping necessary to critically damp

coil is usually obtained magnetically from the motion of the coil in the field of the permanent magnet, which sets up counter electromotive forces. The amount of damping produced by the current in the coil depends upon the total resistance of the coil and connected circuit. That resistance which produces critical damping is called the *critical damping resistance*. A galvanometer is usually so designed that its critical damping resistance is at least five times its coil resistance so that it may be shunted for critical damping without losing much sensitivity. All but the most sensitive pivot instruments are critically damped on open circuit by the current set up in the metal winding form, and resistance of the connected circuit has little effect on the damping.

The *current sensitivity* of any galvanometer varies directly as the number of turns on its moving coil and as the square of its period. For a given winding space on the coil, its resistance varies as the square of the number of turns, assuming that the portion of the winding space occupied by insulation remains constant. The deflection is proportional to the current and to the square root of the resistance, *i.e.*, to the square root of the power dissipated in the coil.

TABLE I.—CHARACTERISTICS OF D-C GALVANOMETERS

Make	Type	E, μv	I, μa	T,	R coil, Ω	Rc.D., Ω	W, μμw
		Suspended-coil type with mirror					
L. & N.	2285a	0.032	0.0027	7.5	12	37	0.00009
	2285b	0.046	0.0038	5	12	52	0.00017
	2285f	0.032	0.00004	20	800	71,000	0.0000013
	2290	0.008	0.00001	40	800	101,000	0.00000008
	2500b	0.25	0.0005	6	500	10,500	0.00012
	2500e	1.5	0.003	3	500	2,500	0.0045
	2500f	0.05	0.0001	14	500	14,500	0.000005
	2239a	1.7	0.014	8	115	10,000	0.022
	2239b	1.0	0.001	14	1,000	10,000	0.001
	2239f	1.6	0.0002	18	8,000	54,000	0.00032
		Suspended-iron type with mirror					
L. & N.	2270	0.008	0.0002	5	40	......	0.0000016
		Suspended-coil type with self-contained scale					
L. & N.	2400c	10	0.01	3	1,000	16,000	0.1
	2420c	25	0.025	3	1,000	16,000	0.62
	2310d	125	0.125	3.5	1,000	11,000	15.6
		Double-pivot type with pointer and scale					
Weston	440	37.5	0.25	2.7	150	1,150	9.4
	440	200	0.05	2.7	4,000	60,000	10.0

Values of voltage E, current I, and power W are for a scale deflection of 1 mm at a scale distance of 1 m for the galvanometers having mirrors; for those having self-contained scales the values given are for a deflection of the smallest division, usually 1 mm. The voltage drop in the external critical damping resistance is not included in the voltage given.

In the selection of galvanometers it should be noted that in general those of high sensitivity will also be slow in action, and in general the natural period and critical damping resistance for a galvanometer as listed by the several makers should be considered as carefully as the sensitivity. Further, galvanometers of highest sensitivity will require great care in leveling; they will be responsive to minor vibrations and in many installations may require special supports.

Where vibration in a building is a factor, the Julius suspension m be used, a somewhat complex system of weights supported by sprin with oil-damping vessels. A simpler method although not so perfe is to rest a 200-lb. block (of concrete) on an air cushion; this will abso all vibration usually encountered in factories, at least for galvanomete of moderate sensitivity. Galvanometers with a single suspension hav the greatest sensitivity, those with a taut suspension less, and those wi double pivots least. For the most sensitive type of galvanomete increasing the period from 5 to 40 sec. allows the power to be decreas from 11 to 0.005 $\mu\mu$w. The minimum current sensitivity is 10^{-11} am per millimeter. The smallest current sensitivity for a taut suspensio is 10^{-8} amp. per millimeter, and for a double-pivot pointer instrume 5×10^{-8} amp. per scale division.

Galvanometers of the suspended type are used mainly as null ind cators for d-c bridges and potentiometers and as deflection instrume in comparison methods. In the latter case a *differential galvanomete* sometimes used. This is a galvanometer having two separate insula windings on the suspended coil. They have equal numbers of tur and are so connected that, when equ currents flow through the two coils, deflection is produced.

FIG. 2.—Ayrton-Mather universal shunt.

The sensitivity of a galvanometer most easily reduced by shunting, an since it is desirable to keep the galvano eter critically damped, the Ayrt Mather universal shunt shown in F 2 is most convenient. This arrangeme is also used in multiple-range ammeters and milliammeters and frequently known as a "series shunt." The total resistance of the sh is made approximately equal to the critical damping resistance of t galvanometer or indicating instrument with which it is used.

Pointer-type instruments of the pivot type are used as ammete and voltmeters of all ranges and as the indicating portions of therm couple, rectifier, and various vacuum-tube instruments. The minimu range of the ammeters extends from 5 μa to an upper limit determin only by the size of shunt desired, commercial shunts having been made 50,000 amp. Above 15 to 30 ma the movements are shunted, in wh case the copper or aluminum winding of the moving coil must ha sufficient manganin swamping resistance in series with it to give a go temperature coefficient when shunted by the manganin resistan Voltmeters may be made with a full-scale range from 1 mv to as high series resistance can be arranged to care for the requirements. Instr ments are made with self-contained series resistance up to a few hundr volts; higher ranges usually require an external resistor with the instr ment placed in the grounded or low-potential side of the circuit for t sake of safety and to reduce electrostatic effects on the moving system.

Voltmeter sensitivity, at the present time, is almost exclusively 1 m for full-scale deflection, this decreased current being almost a requis to limit power requirements at high voltages. While series resistor for low-range voltmeters are of conventional spool type, for ranges over 1,000 volts tubular-type units are widely used, having resista spools of special design, electrostatically shielded in sections contai in insulation tubes and filled with inert wax. Such units are complet

moistureproof and mechanically well protected and are almost universally used for the measurement of plate potentials.

In general, pointer-type indicating instruments can be made to give full-scale deflection on as little as 0.1 μw, although for a rugged instrument from 1 to 5 μw is required. Moving-element resistances may be made from about 1 ohm to 10,000 ohms. Low-resistance elements are limited by the spring or suspension resistance which becomes a very appreciable part of the total, reducing the energy available for torque; high-resistance elements are limited by the available wire, and many are now being wound of enameled copper wire 0.001 in. in diameter.

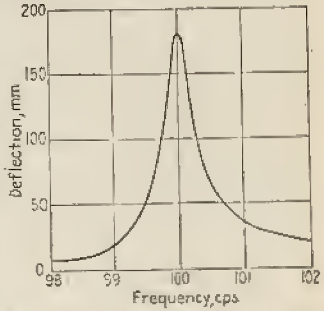


FIG. 3.—Resonance curve of vibration galvanometer.

As in the output circuits of vacuum-tube amplifiers, the resistance of the instrument or galvanometer should be matched to the circuit in which it is placed for maximum energy transfer, and this is particularly important where the energy is limited. On the other hand, this will frequently result in overdamping galvanometers of ultrahigh sensitivity, and a compromise must usually be made between speed of response and sensitivity requirements. It should be noted, however, that this matching is not of vital importance since the loss by a very approximate match in error by as much as 20 per cent is very small.

7. Moving-coil Vibration Galvanometers. When an alternating voltage is applied to the coil of a permanent magnet galvanometer, the coil will follow the alternations of the current if the frequency is of the same order as that defined by its period. Maximum amplitude of vibration will occur at the natural frequency of the coil. The relation between amplitude and frequency is similar to the resonance curve of an electrical circuit. The ratio of the maximum amplitude at its natural frequency to the amplitude for an equal d-c voltage is between 25 and 150. The period of the ordinary d-c galvanometer is never less than 1 sec., while the frequencies at which measurements are made are rarely less than 30 cycles. The upper limit for a taut single suspension is around 300 cycles. This limit may be raised to 1,000 by the use of a taut bifilar suspension. Electrical characteristics of commercial vibration galvanometers are given in Table II. At 60 cycles their sensitivity is equal to that of a good d-c galvanometer. A resonance curve when tuned to a frequency of 100 cycles is shown in Fig. 3.

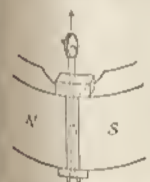


FIG. 4.—Bifilar suspension.

The natural frequency may be raised still further by eliminating the coil entirely and using the single-turn loop formed by the bifilar suspension. The mirror is then placed at the center of the taut wires. The general method of construction is shown in Fig. 4. By this means a natural frequency of 12 kc may be obtained. The sensitivity decreases inversely as the first power of the frequency. On this account it is as sensitive at 10 kc as the bifilar-coil galvanometer was at 1 kc. In com-parison with other null detectors at these frequencies, its sensitivity is low that it is not much used in this form.

8. The Einthoven string galvanometer uses the simplest possi[ble] moving system for a galvanometer. A single conducting string moves [in] the narrow air gap of the magnetic system, which may be a perman[ent] magnet or an electromagnet depending on the sensitivity desired. [Its] motion is observed through a microscope or by its shadow thrown on [a] screen from a point light source. Electrical characteristics of th[e] Einthoven string galvanometer built by the Cambridge Instrum[ent] Company are given in Table II, using a silvered glass string and [a] magnification of 600 times. The string galvanometer may also b[e] used as an oscillograph. The shadow of the string is observed on [a] translucent screen as reflected from a revolving mirror. The moti[on] of the string may also be photographed on film or bromide paper. Th[e] usual paper speed is 10 in. per second, but this may be increased to [a] maximum of 100 in. per second. At this latter speed, phenomena lasti[ng] a millisecond appear 0.1 in. long.

TABLE II.—CHARACTERISTICS OF A-C GALVANOMETERS

Make	Type	f, cycles	E, μv	I, μa	R, Ω	W, $\mu\mu w$
		Vibrating-coil type				
Cambridge	Campbell bifilar	50	8.5	0.017	500	0.1[illegible]
		100	17.5	0.05	350	0.5[illegible]
		350	53	0.33	160	17
		750	104	2.0	52	200
		1,000	175	5.0	35	800
	Campbell unifilar	30	1.5	0.05	30	0.0[illegible]
		50	1.2	0.04	50	0.0[illegible]
		100	3.0	0.025	120	0.0[illegible]
		200	7.0	0.10	70	0.7[illegible]
L. & N.	2350a	60	17.5	0.025	700	0.4[illegible]
Cambridge	Duddell oscillograph	100	5.0	0.02	250	0.1[illegible]
		1,000	50	0.2	250	10
		2,000	100	0.4	250	40
		Vibrating-string type				
Cambridge	Einthoven	100	100	0.025	4,000	2.5
		300	800	0.2	4,000	160
		Suspended-coil type with electromagnet				
L. & N.	2570	60	0.06	0.005	12	0.0[illegible]
		Electromagnet type				
L. & N.	2440	60	16	0.05	325	800,000
		Vibrating-diaphragm type (telephone)				
W. E. Co.	....	800	400	0.02	6,000	2.4

Values of voltage E, current I, and power W are for a scale deflection of 1 mm at scale distance of 1 m for all galvanometers except the telephone, for which the thresh[old] of audibility is used. The moving system is tuned to the frequencies given for a[ll] instruments except the suspended-coil galvanometer with electromagnet.

9. Moving-coil A-c Instruments. If a steady deflection is desir[ed] with a.c., the magnetic field must change in direction with the curre[nt] in the coil and must have the same phase. This requires that the fie[ld] be an electromagnetic one. In the case of galvanometers and partic[u-]

larly null indicators, a field of laminated iron may be used, excited at the same frequency as the moving coil. When used as a null indicator in a bridge network, the field is connected across the same supply as the bridge, while the moving coil is connected to the detector terminals. Since the current through the field and the flux produced will be nearly 90 deg. out of phase with the voltage applied to the bridge, the galvanometer will be most sensitive to the reactance balance and will be little affected by the resistance balance. These conditions may be equalized or reversed by the introduction of resistance in series with the field, or reactance in series with the bridge, to make the field current and bridge current differ in phase by 45 deg. or be in phase. The phase selectivity of the a-c galvanometer may be of advantage in certain special cases, but in general it is a considerable disadvantage. The electrostatic field of the main field winding exerts a considerable force on the moving coil so that it must be carefully shielded. Its sensitivity is very high, and it compares favorably with the best d-c galvanometers.

10. Electrodynamometer. When the iron core is omitted from the field winding, the moving coil and field coil may be connected in series. The deflection is then proportional to the square of the current flowing in the windings, and the instrument is called an *electrodynamometer.* Instruments of this type read the same on both a.c. and d.c. and are suitable as transfer instruments, provided certain precautions are taken. Protection from external magnetic fields is most important. This is usually accomplished in pivot-type instruments by shielding with soft iron. It may also be effected by making the instrument *astatic.* When a.c. is used, an error is introduced if the distribution of current in the coils is affected by eddy currents in the conductors themselves—the so-called *skin effect*—or by capacitance between the windings. The former effect is minimized by the use of conductors with insulated strands—so-called *litzendraht*—the latter by careful spacing and by electrostatic shielding.

Electrodynamometers may be used as galvanometers, ammeters, voltmeters, and wattmeters. Their sensitivity as galvanometers is so low compared with vibration galvanometers and other meters that they are now rarely used. As ammeters, voltmeters, and wattmeters, they are the standard instruments for use at commercial frequencies. In general the sensitivity of a-c instruments is of the order of 1/1,000 of that of d-c instruments, this being due to the difference in field intensity of the electromagnetic field as compared with that which can be obtained from a permanent magnet. Electrodynamometer instruments of the highest precision will take from 1 to 3 watts full scale, the total energy varying with the square of the deflection. Suspension-type electrodynamometers may have sensitivities 100 times as great.

Electrodynamometer ammeters have their fields and moving coils in series up to several hundred milliamperes above which the moving element is shunted across a resistor in series with the fixed coils. Above 50 amp., or so, current transformers are used, and these are now available with special alloy cores which will give accuracies of the order of 1/10 of 1 per cent. Electrodynamometer instruments are ordinarily made to function up to 125 cycles without correction but may be used on frequencies up to several thousand cycles if especially designed or if corrections are made. Note that low-range voltmeters have very low resistance in order to get the required energy; dynamometer voltmeters with full-scale values of 2 volts may draw as much as 0.5 amp. Hig voltages above 1,000 volts are measured with potential transformers.

Electrodynamometer instruments are also used as wat meters when the field is excited in series with the load and the moving coil is acro the load in series with suitable resistance, the readings being proportion to $EI \cos \theta$. For polyphase circuits a multiplicity of similar element may be arranged on a single shaft, the most usual variety being th two-element instrument or three-phase circuits. Such an instrumen gives true power wi hout relation to phase angle.

TABLE III.—CHARACTERISTICS OF A-C AMMETERS

Make	Type	E, v	I, amp.	R, Ω	W, w
		Electrodynamometer type			
Weston	326	2.6	1.0	2.6	2.6
	341	1.0	0.5	2.0	0.5
	370	21	0.015	1,400	0.31
		Moving-iron type			
Weston	155	31	0.02	1,540	0.62
	433	14	0.03	460	0.41
	476	30	0.015	2,000	0.45
	517	30	0.015	2,000	0.45
	528	30	0.015	2,000	0.45
		Thermocouple type			
G. R. Co.	493	0.8	0.008	100	0.00[illegible]
		0.2	0.10	2	0.02[illegible]
Cambridge	....	0.24	0.008	30	0.00[illegible]
		0.12	0.12	1	0.014
		0.08	0.70	0.12	0.059
Weston	412	0.25	0.01	25	0.0025
	425	0.13	0.10	1.35	0.0135
		0.62	0.12	5.2	0.075
		0.59	0.50	1.18	0.295
Cambridge	Duddell	1.5	0.01	150	0.015
		1.5	0.10	1.5	0.015
		Rectifier type			
Weston	301	1	0.001	1,000	0.001
G. R. Co.	488	3	0.00075	4,000	0.00[illegible]
		2	0.0005	4,000	0.001
		2	0.00025	8,000	0.00[illegible]
		2	0.0001	20,000	0.00[illegible]
Westinghouse	PY-4	13.4	0.010	1,270	0.13
	NA	13.6	0.010	1,300	0.13
	NA	13.6	0.010	1,300	0.13

Values of voltage E, current I, and power W are for full-scale deflection.

11. Moving-iron Instruments. Galvanometers may be constructed with a stationary coil and a moving-iron vane or magnet. The moving system consists of small permanent magnets placed at the center of the coil at right angles to the axis of suspension. To avoid the effect of outside magnetic fields, the system is duplicated with the magnets pointing in the opposite direction to make it astatic, and the whole galvanometer is surrounded by multiple soft-iron shields. Its sensitivity (see Table I) is nearly equaled by the best moving-coil galvanometer so that it is very little used.

Soft iron may also be used in the moving element, either alone or in conjunction with a fixed piece of soft iron, both of which are magnetized by the fixed coil.

Soft-iron meters are much used as a-c ammeters and voltmeters in a wide variety of ranges and sizes. They may also be used on d.c. Electrical characteristics are given in Table III. The range of the ammeters is from 20 ma to 500 amp. The upper limit is ten times that of dynamometer-type meters, because the current coil is fixed. Currents up to 5,000 amp. are measured by the use of current transformers. Frequencies to 500 cycles may be used. The range of the voltmeters is from 1 to 750 volts. Their resistances are such as to give from 3 to 200 ohms per volt, the values increasing with the voltage. Higher voltages are measured by the use of either multipliers or potential transformers. Frequencies up to 500 cycles may be used, the normal limit being 125.

In general the sensitivity of pointer-type indicating instruments using the moving-iron principle will be from 0.1 to 1 watt full scale. Instruments using short vanes, usually of the arcuate type, take about 1 watt full scale. Instruments with long radial vanes are more sensitive with a minimum of 0.1 watt full scale but in general are more sensitive to external fields and must be well shielded and kept away from strong external fields. Moving-iron instruments in general are less satisfactory on badly distorted wave forms as the hysteresis loop of the iron is represented in the measurement. They are, however, widely used on power circuits and are generally available in all sizes from the small 2-in. instruments up to the larger switchboard types.

HIGH-FREQUENCY CURRENT METERS

12. For the measurement of currents of high frequency, the only satisfactory means is through the heat developed in a resistor, which heat may be measured by the expansion of a wire, by measuring the thermoelectric voltage developed by a thermocouple adjacent to the resistor wire, by bolometer methods, and by other heat-measuring systems.

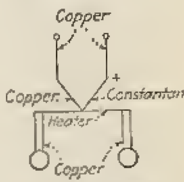


FIG. 5.—Thermocouple meter.

13. The hot-wire expansion type of instrument is today practically obsolete. Its defects of varying in indication with ambient temperature, the lack of perfect resiliency in the heated expansion wire, and its low overload capacity together with the advent of the thermocouple instrument have practically made this type obsolete.

14. A thermocouple meter consists of a heater member, a thermocouple adjacent to it, and a d-c galvanometer or millivoltmeter. Figure 5 shows the basic diagram of the device. Such a simple assembly, however, does not compensate for variations in temperature of the terminals or for ambient temperature variations.

The Weston thermal ammeter as developed by W. N. Goodwin, Jr., is as shown in Fig. 6. The heater is a wire or tube of platinum alloy of very short length whereby most of the heat is conducted to the terminals, thus wiping out largely the effect of convection currents of air. The temperature of the heated member may be represented as a parabola in its gradient from center to each terminal lug, and it is this temperature difference or gradient from the center of the heater to its end which is measured by the thermocouple. The couple proper consists of a pair of wires, usually constantan and a platinum alloy, permanently welded to the center of the heater at the junction end, with the effective cold ends soldered to a pair of copper strips which are thermally connected to, but electrically insulated from, the terminal lugs. Their heat capacity is such that the difference in temperature between the center of the heated member and the center of the two copper compensating strips is always the same as from the center of the heated member to the terminal lug, regardless of ambient temperature changes or general rise in temperature of the surroundings due to heating of the lugs themselves or temperature rise due to the total heat generated. The thermo-electric voltage is, therefore, strictly proportional to the temperature difference between the center and ends of the heated member which in turn is proportional to the square of the current causing this temperature rise and a d-c instrument connected to the couple may be calibrated in terms of this current.

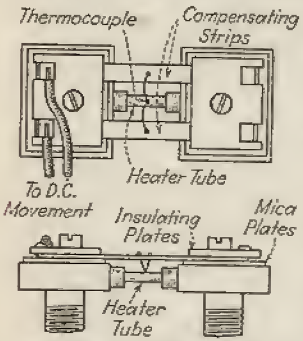


Fig. 6.—Compensated high-frequency thermocouple and heating element.

Couples may be designed to give suitable indication on instruments of commercial types from 200 ma up to whatever may be required. Solid round wires may be used for the heated member up to about 2 amp., but for higher currents and at the higher frequencies skin-effect phenomena cause the readings to be too high. For higher ranges, therefore, the heated member should preferably take the form of thin-walled tube of such dimensions that at the frequency being measured the ratio of h-f resistance to l-f resistance is not over 1.02. This limits the frequency error, on a square-law scale, to less than 1 per cent.

While standard instruments have a square-law scale as the result of the I^2R production of heat, instruments are available in which the upper four-fifths of the scale is approximately linear through the use of special d-c indicating mechanisms having non-linear air gaps whereby the d-c sensitivity is progressively lower as the pointer moves up the scale. By a proper combination of such specially shaped pole pieces a nearly linear scale may be produced. (See Figs. 14*c* and 14*d*.)

Instruments having the linear expanded scale are useful in small broadcast transmitters licensed for a lower power at night than during the day; sufficiently accurate readings of the high and low values of antenna current may be had on the same instrument to be satisfactory, and instruments of this type are listed as complying with FCC rule No. 143.

For low ranges so-called *bridge-type* couples are used, as shown in Fig. 7, whereby a number of couples are arranged in series-parallel to give a higher thermal e.m.f. The impedance of these couples is higher than for a single couple, and for the common current-squared galvanometer the effective resistance is 4.5 ohms. The indicating instrument for the standard single couples has a sensitivity of 12 mv and a resistance of about 5 ohms.

For still higher sensitivities the couple may be placed *in vacuo*. Such couples show no increase in sensitivity until the vacuum is better than

0.01 mm of mercury; but above this point a great increase in sensitivity is obtained up to as much as twenty-five times that obtained in air for certain extremely fine wire couples. The heaters for such couples may be carbon or graphitized wire. Commercial vacuum couples are intended to function with a 12-ohm 200-μa d-c instrument and may be obtained in ranges down to as low as 2 ma in the heater circuit for full-scale deflection on the instrument with a heater resistance of from 700 to 1,000 ohms. Vacuum couples are rarely used for currents higher than a few hundred milliamperes, and the air couples are quite satisfactory for these higher ranges.

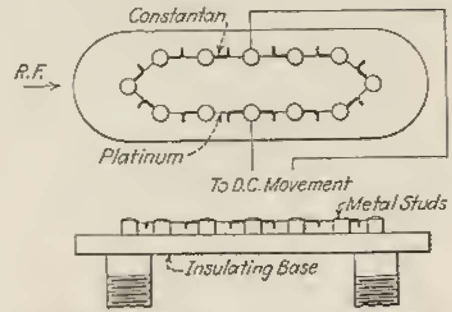


Fig. 7.—Galvanometer or bridge-type couple.

Thermocouple instruments in general are calibrated on commercial frequency a.c., and, if used on d.c., the mean of reversed readings should be taken to make certain that any d-c drop in the heater picked up by the couple is canceled out.

Thermocouple instruments may be obtained with separate couples for use in indicating at a distance as where a couple is placed in the antenna of a transmitting station and the leads brought back to an instrument in the transmitting building. The couple should be placed in the h-f circuit at a point close to ground potential to reduce circulating r-f currents in the leads to the instrument. If this cannot be done, the thermocouple, of low range, around ½ amp., is connected to a loop of wire that is inductively coupled to a loop in the main antenna circuit. The thermocouple circuit may then be grounded. The instrument is scaled to read the total antenna current, and the final calibration is made by adjusting the inductive coupling between the two loops until the remote reading instrument indicates the same value as an instrument placed directly in the antenna itself. Note that FCC rules require an instrument in the main antenna circuit which may be used for this purpose but which under normal operating conditions is kept short-circuited to prevent damage due to lightning. The switch is opened when the instrument is read for logging purposes, and the remote indicator, usually located on the transmitter panel, is used for normal operation.

The ratio of the power available to operate the indicating meter to that put into the heater is about 1 to 2,000 for the most efficient couples; hence a very sensitive d-c instrument is required for low r-f energies.

Thermocouple voltmeters are constructed by using one of the h... sensitive couples with sufficient series resistance to give the desi... voltage range. Their range is from 0.3 to 150 volts with resistance... 125 ohms per volt above 1 volt, and 500 ohms per volt above 10 vo... if desired. Their frequency range is determined by that of the seri... resistance. The small resistance spools which must be used in inst... with self-contained resistors change their resistance rapidly with f... quency so that their frequency limit is 3 kc. Frequencies of 1 Mc m... be attained with an error of 1 per cent with special h-f resistors.

Since the e.m.f. produced by the thermocouple is proportional to ... power input and hence to the square of the current, this meter will r... correctly on both d.c. and a.c. and may therefore be used as a trans... instrument. It is necessary, however, to take the average of the readin... for both directions when using d.c.

RECTIFIER METERS

15. An a.c. may be changed to a pulsating current having a stead... component by the process of rectification. If the current-voltage chara... teristic is as shown in Fig. 8*a* the effect is called *half-wave rectification*. The negative half cycles are eliminated and the positive half cycles repr... duced undistorted. The value of the steady component is half th... average value of a half sine wave. The ratio of the d.c. to the effecti...

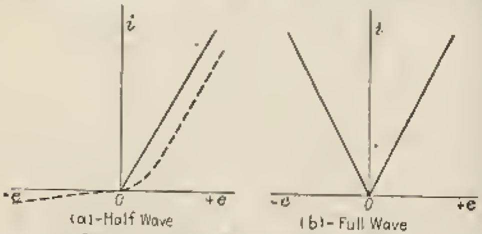


Fig. 8.—Rectifier characteristics.

value of an a-c current having a sine wave form which would flow if th... rectifier were replaced by a pure resistance of the same value as that o... the rectifier is $\sqrt{2/\pi}$, or 0.450. By a combination of rectifiers it is possi... ble to obtain the characteristic shown in Fig. 8*b*, which gives full-wav... rectification. The d.c. is then 0.900 of the a.c. Actual rectifiers have ... curved characteristic as shown by the dotted line in Fig. 8*a*. For nega... tive voltages the resistance is not infinite. The ratio of the positive an... negative half-cycle resistances is sometimes as low as 8. Because of th... curvature of the characteristic, the ratio of d.c. to a.c. is a function both of the magnitude of the current and of wave form.

The crystal rectifiers used with early radio receivers may be used with a sensitive d-c meter for rectifying an a.c. Carborundum, galena, silicon, and many other crystals may be used. The crystal is cast in a low melting-point alloy and the top contact made with a fine copper wire. Rectification occurs at the points of contact of copper and crystal.

16. Commercial rectifier instruments contain a full-wave rectifier consisting of four copper oxide rectifier disks connected in bridge relation

as shown in Fig. 9. The rectification is by virtue of the oxide film formed on the copper disk. Current flows readily from the oxide to the copper and much less readily in the reverse direction. For instrument use the rectifier consists of four small plates arranged in a stack with suitable terminals between adjacent disks for connection to the instrument and the external circuit. The disks may be as large as $\frac{3}{16}$ in. square or round, which size is rated at about 1 volt and 5 ma maximum. This rating is somewhat less than a maximum rating for power purposes since in an instrument some overload capacity is required and stability rather than maximum power is the main requirement. Somewhat smaller disks are used in low-range instruments and for those designed for special characteristics in order to maintain a relatively high current density at lower currents, thus reducing frequency errors. Contact with the oxide is made in a variety of ways through the use of lead washers, graphite, or various metals applied to the surface. The main requirement here is permanence of contact over an extended period.

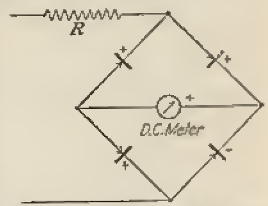


Fig. 9.—Copper oxide rectifier bridge.

The sensitivity[1] of the device depends upon the resistance and full-scale current of the d-c instrument. The d-c instrument measures the average

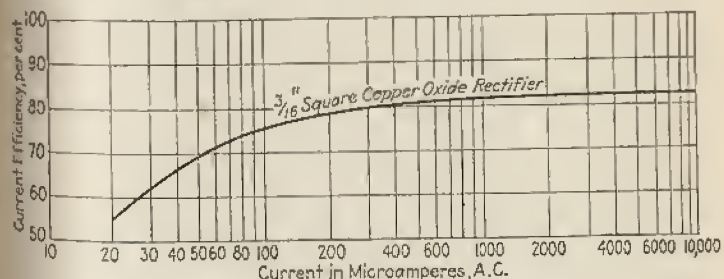


Fig. 10.—Current-efficiency characteristic.

value of a rectified wave, while a.c. is usually measured by methods which give the r-m-s value of the wave. It is customary to calibrate rectifier instruments in terms of the r-m-s value, of a stated wave form, usually a sine wave. If a rectifier instrument is used on a wave form differing widely from the wave for which it is calibrated, an error proportional to the form factor will result. Calibration also corrects an error due to imperfect rectification, which varies with current, temperature, and frequency.

The performance of rectifier instruments can be best expressed by considering the d-c instrument and the rectifier as a unit according to Fig. 9. The current efficiency, $F = \frac{\text{average d-c current}}{\text{r-m-s a-c current}}$, is 80 to 89 per cent for a sinusoidal a-c current in the order of 0.001 amp. It is therefore impossible to use an a-c rectifier instrument for d.c. without first making a suitable

[1] The following several paragraphs and Tables IV and V have been contributed by F. S. Stickney of the Westinghouse Electric & Mfg. Co.

change in circuit or calibration. Figure 10 shows the effect of current o current efficiency for a sinusoidal wave. This variation is corrected calibrating.

The 60-cycle impedance of a 20-ma rectifier instrument is shown in Fig. 11. Other ranges using different rectifiers will have different values, but in gener the slope of the characteristic as plotted in logarithmic coordinates will entirely similar.

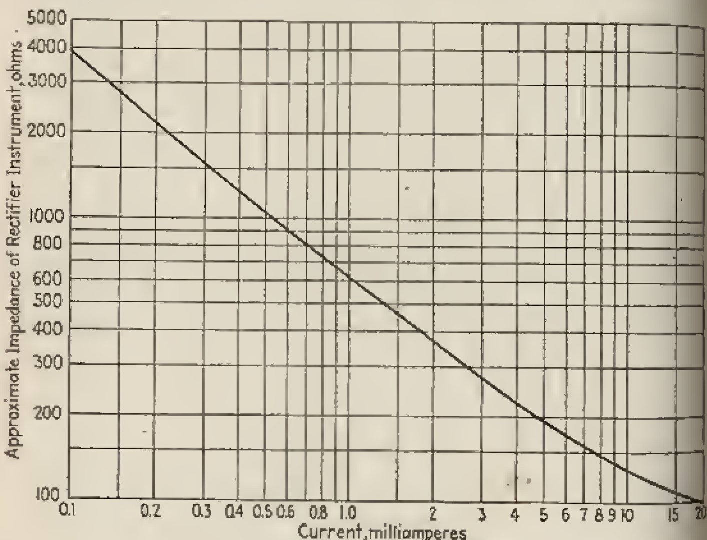


FIG. 11.—Impedance characteristic.

Temperature variations have considerable effect on both the impedance and accuracy of rectifier instruments. Figure 12 shows temperature-voltage variations for a specific group of milliammeters from which impedance can be determined. Figure 13 shows temperature-efficiency relations of this group at various current values. The point must be stressed, however, that the curvature of these characteristics varies with the several parameters of rectifier-disk size, current density, processing time, and the resistance of the instrument, and it is quite possible to modify these curves materially for special requirements. Standard instruments, by the same token, can hardly be represented by any particular group of curves. It might be stated that rectifier instruments have been materially improved in recent years as to the flattening of the curves and that design possibilities have broadened to the point where materially improved instruments can be made for particular requirements.

Higher temperatures adversely affect the rectifying film, and rectifier instruments may become erratic at temperatures in excess of 45°C. High-temperature locations should be avoided in application; where the instrument becomes unduly warm, instruments with external rectifiers are sometimes used with the rectifier placed in a relatively cool location.

Frequency errors are the result of capacity between disks. Since the disk resistance is lower at higher currents and since capacity is a function of rectifier size, the smallest rectifier is preferred for good frequency characteris-

tics. This in turn means a high current density with which good accuracy is obtainable somewhat above audio frequencies. With low-current density, errors may be as large as 1 per cent per 1,000 cycles.

In general, low-range voltmeters are more subject to temperature and frequency errors than high-range voltmeters. Low-range voltmeters have scales which are compressed at

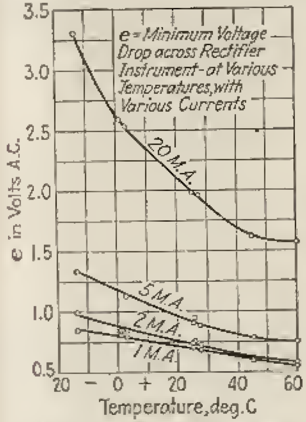


FIG. 12.—Effect of ambient temperature on the voltage drop across a rectifier instrument at various currents.

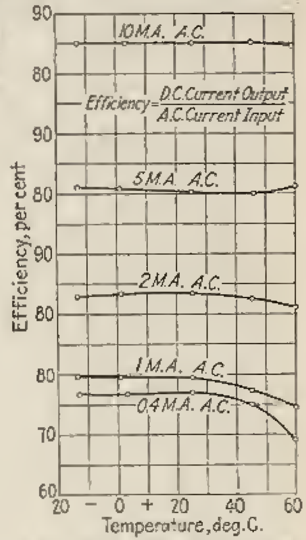


FIG. 13.—Ambient temperature-efficiency relation.

the lower end due to variations of impedance with current. High-range voltmeters and milliammeters have nearly uniform scale distribution.

Tables IV and V give approximate constants of commercial rectifier instruments.

TABLE IV.—MILLIAMMETERS AND MICROAMMETERS

Full Scale, Milliamperes	Approximate 60-Cycle Impedance at Full Scale[1]
15	100
10	130
5	190
2	370
1	600
0.5	1,140
0.2	1,950
0.1	4,200
0.05	6,300
0.02	10,000

[1] Individual copper oxide rectifiers vary considerably from the average in characteristics. Impedance values given may vary ±15 per cent, and efficiency values vary ±3 per cent for the product of *one* manufacturer. Much greater variations may be expected between the products of *different* manufacturers.

17. Power-level instruments used in the monitoring of voice-frequency circuits are usually voltmeters with scales calibrated to read power on the basis of a fixed-resistance load. The indications of power are usually in decibels above or below a specified zero power level. Prior to 1939 considerable confusion existed in this field of measurement owing to the fact that zero levels of 0.001, 0.006, and 0.0125 watt were used into loads of 500 or 600 ohms. The instruments themselves, fundamentally volt-

TABLE V.—VOLTMETERS

Full scale, volts	Full scale, approximate ohms per volt	Approximate fixed resistance, ohms	Approximate 60-cps impedance of rectifier and d-c instrument at full scale, ohms
150	1,000	149,400	600
50	1,000	49,400	600
10	1,000	9,400	600
4	1,000	3,400	600
3	2,000	4,860	1,140
2	2,000	2,860	1,140
1.5	2,000	1,860	1,140
1.	5,000	3,050	1,950
0.5	5,000	550	1,950

meters of the rectifier type, have been quite satisfactory. The usual impedance has been 5,000 ohms and higher to avoid too great a loss due to the addition of the power-level indicator and also to avoid adding harmonics to the line due to the non-linear shunt resistance of the instrument-rectifier network.

This situation has been largely cleared due to the work of Messrs. Chinn, Gannett, and Morris[1] in the development of the so-called VU meter. This is fundamentally a rectifier voltmeter having very definitely specified electrical and ballistic characteristics and a new scale. To this standard of reference the majority of organizations using such instruments have agreed.

Two instrument scales have been standardized, as shown in Fig. 14. The upper scale, known as the *A scale*, emphasizes the VU markings and has an inconspicuous voltage scale. The lower, known as the *type B scale*, emphasizes the per cent voltage and has a relatively inconspicuous VU scale. This latter scale is largely used in broadcast monitoring since the voltage scale indicates in a rather direct fashion the per cent utilization of the facilities. The scales are printed on buff paper to reduce eyestrain; the narrow arc and the figures above it are in black with the heavy arc to the right, the markings above it as well as the markings below the arc in red.

The instrument mechanism, which is identical for both scales, has very definite ballistic characteristics which may be completely defined by the fact that, if a voice-frequency voltage of such amplitude as to give a steady reading of 100 on the voltage scale is suddenly applied, the pointer should reach 99 on this scale in 0.3 sec. and should then overswing the 100 point by between 1 and 1.5 per cent.

Zero level was agreed upon as 1 mw in 600 ohms. Since a voice-frequency channel may contain many components of different frequencies and since they may affect different instruments in a different manner, the ballistic standards above listed are a very necessary part of the new standard. The instrument is standardized on sine-wave voltage and is adjusted to read the 100 mark on the voltage scale with 1.225 volts applied, this representing

[1] CHINN, GANNETT, and MORRIS, *Proc. I.R.E.*, January, 1940; A New Standard Volume Indicator and Reference Level, *Bell System Tech. Jour.*, January, 1940.

4 db above 1 mw in 600 ohms and is applied to the standard instrument as furnished, plus a 3,600-ohm external series resistance.

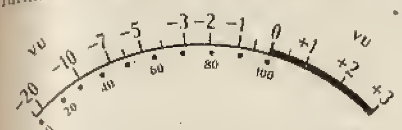


FIG. 14a.—"A" scale for VU meter.

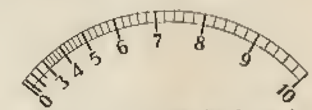


FIG. 14c.—Standard scale using a conventional d-c movement.

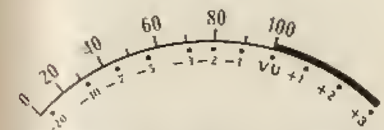


FIG. 14b.—"B" scale for VU meter.

FIG. 14d.—Linear expanded scale using the mechanism with specially shaped pole pieces.

With such an instrument, the readings obtained from it when voice-frequency currents are applied may then be stated as so many VU, taking into account that 4 VU must be added to the scale reading plus the number of VU lost in the attenuator placed in the network.

The required network is shown in Fig. 14f. The fundamental total resistance of the instrument is 7,500 ohms. To this are added 300 ohms representing a 600-ohm source, and load in parallel, making a total of 7,800 ohms. To simplify the use of an attenuator, this is split in the center to give 3,900 ohms each side, which will allow for a simple T-pad attenuator to be inserted at this point. The instrument proper, therefore, has an internal resistance of 3,900 ohms and must be used with the separate 3,600-ohm resistor. Since the normal instrument level is +4 VU, the attenuator dial is marked 4 VU at zero attenuation, and for other true attenuation values 4 VU are added. Table VI, shows values for such attenuators.

FIG. 14e.—Scale of db meter.

This instrument is available commercially and because of its deliberate action is found most readable. The standardization of the instrument by

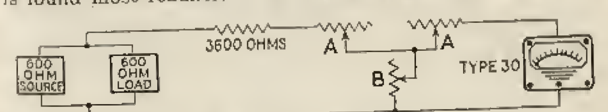


FIG. 14f.—Network for use with VU meter.

the majority of those concerned is of fundamental importance, particularly where levels along a transmission line are to be read, forwarded over an order wire to a common point, and compared. While instruments of several sizes are available, the one in most common use is approximately 4 in. square and available either with or without internal illumination.

The advent of this new level indicator has very largely superseded the use of db meters as such, although the previously available high-speed instruments still find some utility, particularly in the cutting of records where instantaneous indication and control of high level is necessary to prevent overcutting.

Table VII is a useful tabulation of power levels, ratios, and voltages, all in terms of the zero level of 1 mw in 600 ohms, and, when interpreted on an instrument of the characteristics described, the values of db above and below this level will also represent VU.

TABLE VI.—ATTENUATORS FOR VU METER

Attenuator loss, db	Level, VU	Arm A, ohms	Arm B, ohms	Attenuator loss, db	Level, VU	Arm A, ohms	Arm B, ohms
0	+ 4	0	Open	24	+28	3,437	494.1
1	+ 5	224.3	33,801	25	+29	3,485	440.0
2	+ 6	447.1	16,788	26	+30	3,528	391.9
3	+ 7	666.9	11,070	27	+31	3,566	349.1
4	+ 8	882.5	8,177	28	+32	3,601	311.0
5	+ 9	1,093	6,415	29	+33	3,633	277.1
6	+10	1,296	5,221	30	+34	3,661	246.9
7	+11	1,492	4,352	31	+35	3,686	220.0
8	+12	1,679	3,690	32	+36	3,708	196.1
9	+13	1,857	3,166	33	+37	3,729	174.7
10	+14	2,026	2,741	34	+38	3,747	155.7
11	+15	2,185	2,388	35	+39	3,764	138.7
12	+16	2,334	2,091	36	+40	3,778	123.7
13	+17	2,473	1,838	37	+41	3,791	110.2
14	+18	2,603	1,621	38	+42	3,803	98.21
15	+19	2,722	1,432	39	+43	3,813	87.53
16	+20	2,833	1,268	40	+44	3,823	78.0[illegible]
17	+21	2,935	1,124	41	+45	3,831	69.52
18	+22	3,028	997.8	42	+46	3,839	61.9[illegible]
19	+23	3,113	886.3	43	+47	3,845	55.22
20	+24	3,191	787.8	44	+48	3,851	49.21
21	+25	3,262	700.8	45	+49	3,857	43.8[illegible]
22	+26	3,326	623.5	46	+50	3,861	39.0[illegible]
23	+27	3,384	555.0				

MEASUREMENTS OF PULSATING CURRENTS AND POTENTIALS

In making measurements of current and voltage which are neither true a.c. nor d.c., care must be taken to make the measurement with the correct type of instrument in order that a measurement be had of the actual value required.

18. Rectified current, which may or may not be filtered, should in general be measured with a moving-coil permanent-magnet type of d-c instrument. This gives the average value. It is the value of current or voltage of interest when charging a battery and in general is the value of interest in vacuum-tube technique. Iron-vane and electrodynamometer instruments indicate the r-m-s value which is used for determining the heating effect.

Direct-current instruments, particularly voltmeters, have a sufficiently large heat-overload capacity so that they may ordinarily be used on pulsating currents without danger.

To measure the a-c component of voltage, a condenser may be placed in series with an a-c voltmeter of suitable range; the d-c component is blocked and the a-c value only is measured. The impedance of the

TABLE VII.—USEFUL TECHNICAL DB DATA (WESTON)

Power level, db	Power ratio to 0 db. Also power, milliwatts, when 0 level = 1 mw	Voltage ratio to 0 db	Voltage—based on 1 mw in 600 ohms = zero level	Power level, db	Power ratio to 0 db. Also power, milliwatts, when 0 level = 1 mw	Voltage ratio to 0 db	Voltage—based on 1 mw in 600 ohms = zero level
−10	0.1000	0.31623	0.24495	20	100.00	10.0000	7.7461
− 9	0.1259	0.35481	0.27483	21	125.89	11.220	8.6912
− 8	0.1585	0.39811	0.30839	22	158.49	12.589	9.7514
− 7	0.1995	0.44668	0.34599	23	199.53	14.125	10.941
− 6	0.2512	0.50119	0.38820	24	251.19	15.849	12.276
− 5	0.3162	0.56234	0.43560	25	316.23	17.783	13.775
− 4	0.3981	0.63096	0.48875	26	398.11	19.953	15.459
− 3	0.5012	0.70795	0.54840	27	501.19	22.387	17.341
− 2	0.6310	0.79433	0.61527	28	630.96	25.119	19.457
− 1	0.7943	0.89125	0.69035	29	794.33	28.184	21.831
0	1.0000	1.00000	0.77461	30	1,000.00	31.623	24.495
+ 1	1.2589	1.1220	0.86912	31	1,258.9	35.481	27.484
+ 2	1.5849	1.2589	0.97514	32	1,584.9	39.811	30.837
+ 3	1.9953	1.4125	1.0941	33	1,995.3	44.668	34.600
+ 4	2.5119	1.5849	1.2276	34	2,511.9	50.119	38.822
+ 5	3.1623	1.7783	1.3775	35	3,162.3	56.234	43.560
+ 6	3.9811	1.9953	1.5459	36	3,981.1	63.096	48.875
+ 7	5.0119	2.2387	1.7341	37	5,011.9	70.795	54.840
+ 8	6.3096	2.5119	1.9457	38	6,309.6	79.433	61.527
+ 9	7.9433	2.8184	2.1831	39	7,943.3	89.125	69.035
+10	10.0000	3.1623	2.4495	40	10,000.0	100.000	77.461
+11	12.589	3.5481	2.7484	41	12,589.2	112.20	86.912
+12	15.849	3.9811	3.0837	42	15,848.9	125.89	96.698
+13	19.953	4.4668	3.4600	43	19,952.6	141.25	109.41
+14	25.119	5.0119	3.8822	44	25,118.9	158.49	122.76
+15	31.623	5.6234	4.3560	45	31,622.8	177.83	137.75
+16	39.811	6.3096	4.8875	46	39,810.7	199.53	154.59
+17	50.119	7.0795	5.4840	47	50,118.7	223.87	173.41
+18	63.096	7.9433	6.1527	48	63,095.7	251.19	194.57
+19	79.433	8.9125	6.9035	49	79,432.7	281.84	218.31

condenser at the frequency used (120 cycles for a full-wave rectifier system) should not be greater than 10 per cent of the instrument resistance; the impedances being in quadrature, the resulting error will be under 1 per cent. This is the simplest method of measuring hum in a rectified plate supply. Because of its high resistance, the rectifier voltmeter described previously is most satisfactory for this purpose.

Peak voltages and currents are best measured through the use of a vacuum-tube voltmeter with a large capacity shunted by an extremely high resistance d-c voltmeter (see Art. 29). A cathode-ray oscillograph is also useful in such studies.

VOLTAGE-MEASURING INSTRUMENTS

19. Use of Current Meters to Indicate Voltage. All current-measuring instruments having a sensitivity in milliamperes may, with the addition of suitable series resistance, be used to indicate potential. The current drain of the instrument must be sufficiently low to abstract negligible energy from the circuit, as otherwise corrections must be made. With modern instruments of high sensitivity this requirement can usually be met.

20. Direct measurements of voltage are obtainable through electrostatic means, but the instruments are of limited utility because of their low torque and because the minimum ranges are rarely under several hundred volts. They are essentially instruments for the research laboratory.

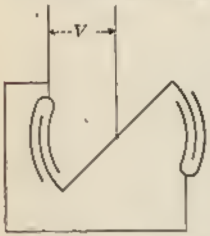


Fig. 15.—Suspended-vane meter.

Electrostatic voltmeters depend on the attractive force which exists between two conducting plates between which a difference of potential exists. In their simplest form, the force of attraction between a stationary and a movable disk is balanced by a calibrated spring. The *Kelvin absolute electrometer* is constructed in this manner. The force of attraction is proportional to the square of the difference of potential between the plates. Such meters give the same indication on steady and alternating voltages and have neither wave form nor frequency error.

One type of construction, used in suspended-vane meters, is shown in Fig. 15. The stationary plates are sections of two concentric cylinders into which the cylindrical rotor turns. With the opposite poles of a magnet placed outside the stator plates, satisfactory damping is obtained from the currents induced in the loop. This type of construction is that used in the *Ayrton-Mather electrostatic voltmeter* built by the Cambridge Instrument Company.

Electrostatic voltmeters are very useful because of their high resistance and low power consumption at low frequencies. They cannot be used on high voltage at frequencies much above a megacycle, because of the rapid increase of the power loss in the necessary insulation. This loss increases directly as the first power of the frequency and the square of the voltage. A hard-rubber insulator with a power factor of 0.004 and capacitance of 10 $\mu\mu f$ will have, at a frequency of 10 megacycles and voltage of 2.5 kv, a charging current of 1.5 amp. and a power loss of 15 watts, both of which values are excessive.

MEASUREMENT OF RESISTANCE

21. While bridge measurements of resistance give greatest accuracy (Art. 36*ff.*) direct-reading instruments are much used because there is no requirement for the manipulation of the controls, and they are widely used in production testing of resistance units as well as in general laboratory practice where the highest accuracy is not essential.

The simplest *direct-reading ohmmeter* consists of an ammeter and battery as shown in Fig. 16. Two readings are made, one with the terminals shorted, the other with the unknown resistance R connected

The fixed resistance S limits the current to about full-scale reading of the ammeter. The deflection is made exactly full scale by adjustment of the ammeter shunt B. The range of this type of meter is usually taken as that resistance which gives a deflection which is 5 per cent of full scale. On this basis the usual ranges are 1,000, 10,000, and 100,000 ohms.

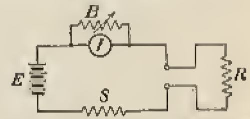


Fig. 16.—Direct-reading ohmmeter circuit.

Through the use of more complex networks, instruments with still wider ranges of capacity can be made available. The upper limit of resistance measurements by this means depends upon the instrument sensitivity and battery voltage; a 50-μa instrument at 15 volts gives an excellent deflection on several megohms. The lower limit, since a minimum battery voltage of 1.5 volts must be used, is dependent only on the current capacity of the battery and the resistance of the leads. In general, for accurate work, the effective battery resistance must be calculated into the circuit as a part of the total series resistance.

Note that in all series-type ohmmeters the center- or half-scale resistance value is exactly equal to the total effective ohmmeter resistance at its terminals.

The readings of an ohmmeter may be made independent of the applied voltage by dispensing with the controlling springs and obtaining the controlling torque from a separate coil connected across the supply voltage. Figure 17 shows the circuit used by Evershed and Vignole in their ohmmeters of this type.

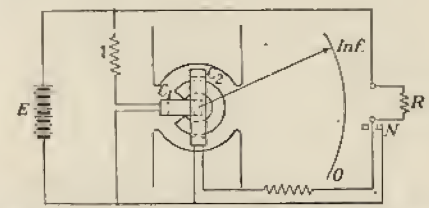


Fig. 17.—Ohmmeter of Evershed and Vignole.

This construction was first used by Evershed for an ohmmeter designed to measure high resistances up to 100 megohms. The source of voltage was a self-contained high-voltage magneto generator, giving voltages up to 500 volts. It was called a *megger*. The same principle has now been applied to ohmmeters of lower range using battery voltages. The resistance range extends from 1 ohm to 5,000 megohms.

22. Measurement of Impedance. When the voltmeter-ammeter method is used with a source of alternating voltage, the ratio of voltage to current gives the impedance of the load

$$Z = \frac{E}{I} \tag{2}$$

With the usual a-c instruments the corrections for the instruments are larger and more difficult to make because of their reactance. The high-resistance rectifier voltmeter and vacuum-tube voltmeter eliminat this difficulty.

The separation of impedance into its components requires the use of a wattmeter. The connections of Fig. 18*a* are usually used when no correcti for instrument errors is to be made, while those of Fig. 18*b* allow the correcti to be made quite easily. For this distinction the current coil of the wattmet is grouped with the ammeter and its potential coil with the voltmeter. A

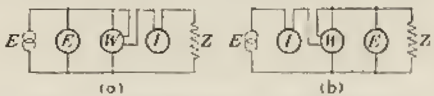


Fig. 18.—Measurement of impedance.

before, the impedance of the load is given by Eq. (2). Its power factor the ratio of the wattmeter readings to the product of voltage and current.

$$\text{P.f.} = \cos\theta = \frac{W}{EI} \tag{3}$$

where θ is the phase angle between voltage and current. The resistance of the load is

$$R = \frac{W}{I^2} \tag{4}$$

and the reactance

$$X = \sqrt{W^2 - R^2} \tag{5}$$

With the knowledge as to whether the load is inductive or capacitive, the inductance or capacitance may be calculated from

$$X = \omega L = -\frac{1}{\omega C} \tag{6}$$

where $\omega = 2\pi f$.

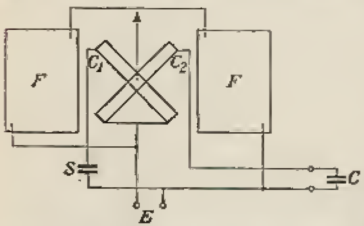


Fig. 19.—High-frequency microfarad meter. (*Weston.*)

23. Measurement of Capacitance. Since the power factor of the usual condenser is small, its reactance is approximately equal to its impedance. This may be measured directly by the voltmeter-ammeter method and the capacitance calculated from Eq. (6). At a given voltage and frequency, a single ammeter reading is sufficient, and the ammeter may be calibrated to read capacitance directly.

Capacitance may also be measured on a single indicating meter whose readings are independent of the applied voltage. The moving element consists of two coils set at right angles to each other. There are no controlling springs. The connections used in the high-frequency Weston microfarad meter are shown in Fig. 19.

Coils C_1 and C_2 are connected across the supply voltage, one in series with a fixed capacitance S, the other in series with the unknown C. The stationary field coils F are directly connected across the line voltage. With no condenser

connected in circuit with coil C_2, the coil C_1 sets itself in the plane of the field coils F and determines the zero of the scale. The introduction of C allows current to flow in the coil C_1 and provides an opposing torque which is proportional to the capacitance added. The resulting deflection is, of course, just as dependent on frequency as on capacitance, so that any particular instrument must be used on the exact frequency for which it was calibrated. The low-frequency Weston microfarad meter has the moving coils connected in series instead of in parallel with the field coils.

The capacitance range of the Weston microfarad meters extends from 0.05 to 10 μf at 60 cycles, 0.001 to 0.05 μf at 500 cycles, and 0.0005 μf at 1,000 cycles. The applied voltage must be large enough to provide sufficient torque to give a definite reading.

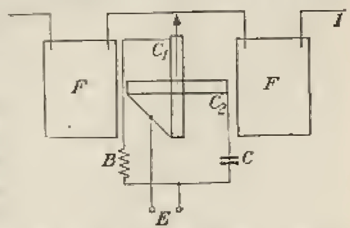


Fig. 20.—Power-factor meter. (*Weston.*)

24. Measurement of Power Factor. Instruments for measuring power factor are very similar to the moving-coil capacitance meters described above. The connections used in the Weston power-factor meter are shown in Fig. 20.

25. Measurement of Frequency. Frequency may be measured with an indicating instrument similar to the capacitance meter shown in Fig. 19, in which the capacitance C is fixed and the capacitance S is replaced by a resistance. The scale is, of course, calibrated in terms of frequency.

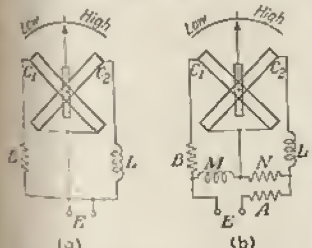


Fig. 21.—Frequency meter. (*Weston.*)

The functions of the moving and fixed coils may be transposed, the stationary part now consisting of two coils set at right angles to each other. The moving part is simply a vane of soft iron, since its sole function is to indicate the direction of the resultant magnetic field set up by the two stationary coils. The connections of such a frequency meter are shown in Fig. 21*a*. The tendency of the vane toward rotation is overcome in the Weston frequency meter by decreasing the phase difference between the currents in the two coils as shown in Fig. 21*b*. The rotation of the magnetic field is no longer uniform. The vane, being long and narrow, takes up a definite position, its inertia preventing it from following the irregular rotation of the magnetic field. The frequency range of the instrument is about 30 per cent of the mid-scale reading. These meters are usually built for the commercial frequencies 25 and 60 cycles. The General Electric Company has built them for higher frequencies, up to 2,000 cycles.

Frequency meters that make use of *vibrating reeds* are also constructed. A series of reeds, whose natural frequencies of vibration differ by regular intervals, are arranged in a line or in a circular arc in the order of ascending frequency. They are mounted on a suitably shaped electromagnet, whose winding is connected across the supply voltage of unknown frequency. That reed, having a natural frequency nearest to the supply frequency, will vibrate with an easily visible amplitude, and the frequency intervals between adjacent reeds are sufficiently small, compared to the damping, so that at least one will always vibrate.

MOVING-DIAPHRAGM METERS

26. The telephone is a very sensitive galvanometer, in which the indication of motion is acoustic. It is essentially a moving-iron vibration galvanometer, polarized with a permanent magnet. Its construction is shown in Fig. 22. The amplitude of vibration is proportional to the product of the steady flux in the air gap produced by the permanent magnet and the alternating flux produced by the coils carrying the a.c. The latter flux is much increased by placing the coils on laminated soft iron pole pieces. The reluctance of the hardened steel magnet to the alternating flux is so great that most of the a-c flux passes across the gap at the base of the pole pieces. This gap is made the proper length to make the product of the two fluxes at the diaphragm air gap a maximum. The diaphragm is a thin steel disk clamped at its outer edge. Its natural frequency of vibration is determined by its mass and stiffness. For silicon steel 0.01 in. in thickness, this frequency is about 900 cycles. By plugging the orifice in the earpiece, the natural frequency may be increased by as much as 50 per cent. The damping of the diaphragm is very small, being mainly due to the eddy-current losses in the iron. The variation of amplitude with frequency is a sharp resonance curve. Figure 23 shows such a curve for a Western Electric telephone. The damping is little affected by changes in stiffness and natural frequency. The impedance of a telephone winding increases with frequency in a regular way, except around the resonance frequencies. The resistance and reactance are generally of the same order of magnitude, so that its lag angle is about 45 deg. At a frequency of 1,000 cycles they are about ten times the d-c resistance of the winding. Near resonance the motion of the diaphragm introduces a counter e.m.f. into the circuit which is usually interpreted as additional resistance and reactance. These terms are referred to as *motional values*. In telephones of low damping, they may be as much as 70 per cent of the normal values. The actual numerical value of the resistance and reactance depends on the number of turns with which the magnets are wound. The d-c resistance varies from 30 to 1,000 ohms. The sensitivity of telephones is somewhat indefinite because it depends on the acuteness of hearing of the observer. It is usual to express it as the current necessary to produce a just audible response. Because of the existence of a threshold of hearing, this minimum current is reasonably definite and reproducible, at least for any one person. Values of this minimum current, together with the correspond-

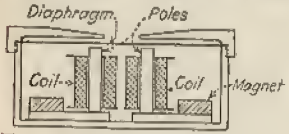


Fig. 22.—Construction of telephone.

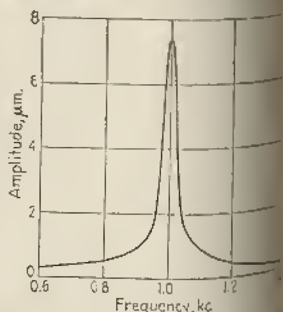


Fig. 23.—Resonance curve of Western Electric telephone.

ing voltage, resistance, and power are given in Table II for a Western Electric receiver. It is much more sensitive than any vibration galvanometer and at its resonant frequency is not far behind a good d-c galvanometer.

27. Other Types of Telephones. It is possible to use non-magnetic materials for the diaphragm by providing a separate steel armature so shaped and clamped that its natural frequency is higher than that of the diaphragm, to which it is attached by a stiff rod. When mica is used for the diaphragm, both the sensitivity and the selectivity are greater than for steel. On the other hand, the resonance curve can be broadened by using a corrugated diaphragm of suitable material.

The steel armature can be replaced by a coil carrying the a.c., which then may vibrate in the field as a **moving-coil** galvanometer. A light paper cone attached to the coil acts as a diaphragm. There is no single natural frequency, so that over a wide frequency range the sensitivity is essentially constant.

The **piezoelectric** effect exhibited in certain crystals is also used as the basis for a telephone. Rochelle salt crystals are used rather than quartz because of their greater piezoelectric effect. The construction is the same as is used in crystal microphones. The frequency characteristic of a telephone made in this manner is remarkably constant over the whole a-f range extending from 100 cycles to 5 kc. Its impedance decreases with frequency because it is essentially a condenser with crystal dielectric. In this respect it behaves in just the opposite manner from a permanent magnet telephone.

28. Thermophones. When a fine wire is heated by the passage of a.c., sound waves are produced in the surrounding air if the heat capacity of the wire is so small that the temperature of the surface of the wire follows the cyclic variations of the current. Instruments of this sort have been constructed, using gold foil as the heater. They are called *thermophones*. Their sensitivity in terms of sound energy is low. But they can be made small enough to be placed in the ear, so that their over-all sensitivity is quite satisfactory. Their response decreases slowly as the frequency is increased. The theory of this instrument has been studied in considerable detail because of its use as a standard in the production of sound.

ELECTRON TUBE METERS

29. Vacuum-tube Voltmeters. The simplest type of vacuum-tube voltmeter makes use of a three-electrode tube and a d-c galvanometer. Its connections are shown in Fig. 24. The grid bias E_c is so chosen that maximum plate rectification occurs, the relation between plate current and grid voltage being as shown in Fig. 25. When an alternating voltage e is applied between grid and filament, the average plate current increases from I_P to I_P'. This change in plate current is the quantity in terms of which the instrument is calibrated. The upper limit of applied voltage e is that for which the peak voltage equals the grid bias.

The zero of the plate-current meter may be suppressed mechanically so that the zero of the voltage scale may coincide with its electrical zero. This suppression may also be attained electrically as shown in Fig. 26. The single battery E_b supplies both grid and plate voltages through the drop wire composed of the three resistors R_c, R_b, R_m. The voltage drop in R_m is made equal to that in the adjustable resistor R_p caused by the plate current.

The galvanometer resistance should be small compared with R_p so that major part of the change in plate current will pass through the galvanometer.

The grid bias for the voltmeters shown in Figs. 24 and 26 may also obtained by connecting the grid return to a resistance R_b in the pl circuit as shown in Fig. 27. This method of obtaining the grid bias causes the bias to increase with the applied voltage. The relation resulting between meter deflection and signal voltage, while approximately a square-law relation for

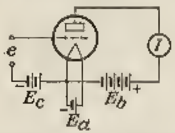


Fig. 24.—Vacuum-tube voltmeter.

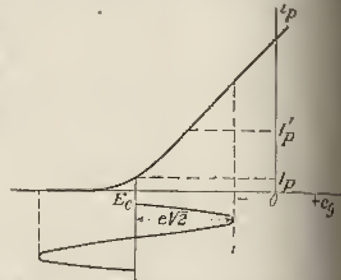


Fig. 25.—Vacuum-tube voltmet characteristic.

small voltages, becomes nearly linear for large voltages of from 20 to 10 volts. For a large grid bias, plate current flows only during the positi peak; hence the error due to wave form may become serious. The vo meter then becomes a peak or cr voltmeter. The voltage range of t meter can be changed by means of t bias resistor R_b. Each range mu however, have a separate calibrati

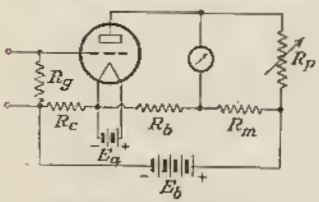


Fig. 26.—Single battery for plate and grid voltages.

The *sensitivity* obtainable with vacuum-tube voltmeter depends main ly upon that of the indicating meter The detection coefficients of the vari ous tubes available are not widely different and are not much affected by the value of plate voltage. A full scale reading of 3 volts is usual with d-c meter showing full-scale deflection on 200 μa. A 20-μa meter woul show a full-scale deflection on 1 volt. Wall galvanometers may be us to obtain increased sensitivity but the difficulty in maintaining the zero setting increases greatly.

The *input resistance* of a vacuum-tube voltmeter is high, being either the insulation resistance of the input terminals or the resistance R_g of Fig. 26 shunted between grid and filament to maintain the grid bias. This may be as high as 10 megohms. The plate load of the tube is sufficiently low so that it does not affect the input resistance. The input capacitance is essentially that of the terminals, socket, and grid-filament capacitance. By careful design this may be made as low as 5 $\mu\mu$f.

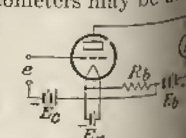


Fig. 27.—Gr bias from plate c cuit.

The *calibration* of a vacuum-tube voltmeter is usually independent of frequency over a wide range. At low frequencies an error appear

when the reactance of the plate by-pass condenser, connected between plate and filament to provide a low-impedance path for the alternating component of the plate current, becomes comparable with the plate load. If this condenser is omitted, in order that the meter may be calibrated and used at commercial frequencies, errors may appear at frequencies below 100 kc due to natural frequencies in the meter and resistances of the plate circuit. Finally, natural frequencies in the grid circuit, either in the resistance R_g of Fig. 26 or in the combination of R_g and the grid-filament capacitance of the tube, set an upper limit around 10 Mc.

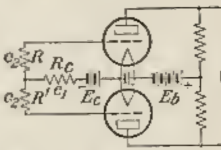


Fig. 28.—Balanced vacuum-tube voltmeter.

The sensitivity of the triode vacuum-tube voltmeter may be increased by the method suggested by Turner[1] in which two voltages are impressed on two balanced tubes connected as shown in Fig. 28. Equal voltages e_2 are applied to the two grids in opposite phase across resistances R and a separate voltage e_1 of the same frequency and the same phase as either is introduced into the common grid lead across the resistance R_c. With the grid bias adjusted for plate rectification, the differential current through the meter connected between the two plates is proportional to the product e_1e_2 of the two voltages. The voltage e_2 applied to

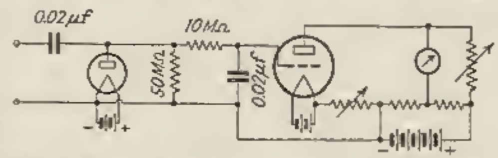


Fig. 29.—Two-electrode vacuum-tube voltmeter.

each grid is usually the small voltage to be measured and voltage e_1 is a high voltage which gives increased sensitivity. A special phase shifting network is generally necessary for the adjustment of voltage e_1. An effective amplification of 100 may be obtained.

If the two voltages are not in phase, the current through the ammeter is proportional to $e_1e_2 \cos\theta$, where θ is the phase angle between e_1 and e_2. This is the form for the expression for power in an a-c circuit. Hence, if e_1 is proportional to the voltage across any load, and e_2 is proportional to the current through that load, obtained as the fall of potential due to the flow of this current through resistances R, the ammeter deflection is proportional to the power dissipated in the load. Full-scale deflection may be obtained with powers as small as 20 μw. The frequency limits are those of the regular vacuum-tube voltmeter.

The use of a **two-electrode tube** in a vacuum-tube voltmeter allows the frequency range to be raised above 50 Mc. Since the rectified current is at the most only a few microamperes, it must be amplified by means of a triode in whose plate circuit the indicating meter is placed. The connections for such a voltmeter are shown in Fig. 29. The current rectified by the diode charges first the condenser in the input lead to the peak value of the applied voltage and then the 0.02-μf condenser which supplies part of the grid bias of the triode. The use of two condensers is required because the cathodes of both the diode and triode must be kept at essentially ground potential. A full-scale reading of 1.5 volts can be obtained with a 200-μa d-c meter. Other voltage ranges up to a maximum of 150 volts can be obtained by shunting the meter. Each range must, however, have a separate scale.

[1] Turner and McNamara, *Proc. I. R. E.*, **18**, No. 10, 1743–1747, October, 1930.

30. Electron-stream Meters. A stream of moving electrons is used in the *cathode-ray tube* to indicate and measure an electric or magnetic field. Electrons emitted from a hot cathode C are accelerated by a positive potential applied to the anode A as shown in Fig. 30. Most of the electrons strike the anode and form the anode or plate current. The remainder pass through a small hole in the center of the anode and continue at constant velocity to a fluorescent screen S of willemite or zinc sulfide, which is usually the enlarged end of the glass tube in which the various parts are mounted. The beam is naturally divergent because of the mutual repulsion of the individual electrons comprising it and must be focused on the screen in some manner in order to obtain a small sharp spot. In the earlier tubes this was accomplished by leaving enough residual gas in the tube to give a pressure of about 0.001 mm of mercury. The positive ions produced by the electron stream exert a

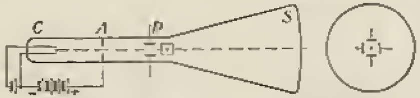


FIG. 30.—Electron-stream meter.

repulsive force on the electrons and prevent their divergence. Satisfactory focusing by this means demands a constant gas pressure which is difficult to maintain throughout the life of a tube. There is also an upper limit of perhaps 100 kc to the frequency for which sharp focusing can be obtained because of the relative slowness of the ionization process.

The beam may also be focused by a longitudinal magnetic field or a radial electric field, the latter being the more convenient. For this type of focusing, the gas pressure is reduced to the minimum necessary to prevent an accumulation of negative charge on the screen. Between the anode A and screen S there is placed a second anode having a positive potential between four and five times that of the first anode. In some designs the enlarged conical end of the tube is lined with a conducting layer and serves as this second anode. In others the second anode is a short cylinder or ring of larger diameter than the first anode. The cathode is usually of the oxide-coated type with a separate heater which, aside from its high efficiency in producing electrons, operates at a temperature sufficiently low so that light from it does not illuminate the screen. It is surrounded by a focusing cylinder with a partially closed outer end, which is connected directly to the cathode when the second anode is used. In tubes with residual gas the exact focusing of the beam is attained by varying the negative voltage applied to this cylinder.

The electron stream may be deflected by a transverse magnetic or electric field, applied beyond the first anode in the region where the electrons have a constant velocity. The losses inherent in the coils necessary to produce a transverse magnetic field limit their use to special cases. The transverse electric field is applied through four deflecting plates symmetrically disposed around the tube axis. When a difference of potential is applied to either pair of opposite plates, the stream of electrons is deflected toward the positive plate through an angle proportional to the strength of the electric field. The bright spot on the fluorescent screen, which marks where the electrons strike the screen, then moves proportionally. A voltage applied between the other pair of plates

produces a deflection of the spot in a direction at right angles to the first deflection. The deflection at the screen is inversely proportional to the higher anode voltage. It is of the order of 2 in. per 100 volts for an anode voltage of 1,000 volts.

When an alternating voltage is applied to a pair of plates, the electric field set up between the plates is continually varying in magnitude and direction. The stream of electrons is deflected back and forth between the plates, and the spot of light is drawn out into a line symmetrically disposed about the undeflected spot, provided the pair of plates is grounded at a point midway in potential between them. An alternating voltage applied to the other pair of plates will produce a line at right angles to the first. If the two voltages are applied to the two pairs of plates simultaneously, the electron stream follows the instantaneous resultant force exerted by both fields and traces on the screen a pattern which is closed, and therefore appears stationary, when the frequencies used bear a simple relation to one another. These patterns are called *Lissajous*[1] *figures*. For two equal frequencies the pattern is an ellipse of varying eccentricity which at the extremes becomes a straight line or a circle. The exact figure is determined by the phase difference of the two voltages. For other ratios of the two frequencies the patterns become reentrant. For the general case the ratio of the number of loops formed on adjacent sides of the pattern is that of the two frequencies.

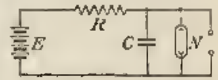


FIG. 31.—Timing circuit for cathode-ray tube.

31. Timing Axis. Since the electron stream can follow accurately all variations in applied voltage, it is only necessary to spread out the line of light which it produces on the screen into a two-dimensional picture to make visible its exact wave form. The second voltage of the same frequency giving the elliptical pattern just described does this but in such a manner that the whole pattern must be redrawn to be easily interpreted. The time axis, which the second voltage must provide, should be linear, not sinusoidal, and its return to zero value should be instantaneous.

A very convenient circuit for this purpose employs a neon tube as shown in Fig. 31. The potential across the condenser C builds up according to an exponential law determined by the time constant CR of the circuit, which over the first part of its range is nearly linear. At some potential between 100 and 300 volts, dependent on the shape of the electrodes and the pressure of the gas, the neon tube breaks down, and the condenser discharges very rapidly. At some lower voltage the neon tube goes out, and the charging process is resumed. If the resistance R is replaced by a two-electrode vacuum tube, the curvature of the exponential law of charging may be partially compensated for by the changing resistance of the vacuum tube as the voltage across it is varied. The frequency at which the condenser charges and discharges depends on the time constant CR of the charging circuit and is controlled by varying these quantities. Frequencies covering the range from 1 to 20,000 cycles are attainable. The wave form thus spread out on the screen will drift along the time axis unless the two frequencies are exactly equal or are simple multiples. It is very convenient to have the pattern stationary. The two frequencies may be synchronized by using a thyratron or three-electrode gas-filled tube in place of the two-electrode neon tube. Some voltage from the source of the wave form under observation is applied to the grid of the thyratron. When the control circuit is adjusted to produce approximately the correct frequency, this added voltage is sufficient to trigger off the discharge and maintain exact synchronism.

[1] BARTON, "Textbook on Sound," pp. 555–557.

A time axis may also be obtained by viewing the screen on a *revol mirror*. The pattern will be stationary when the speed of revolution of t mirror is an exact multiple of the frequency of the given wave.

Transient phenomena may be studied by photographing the sin trace of the electron stream as spread out by any of the method obtaining a time axis just described. The time axis may also be obtain by moving the photographic film itself. In this case, and also for t revolving-mirror method, the screen must be of the type in which t fluorescence does not persist, else the trace on the film will be blurr Screens with persistence times as short as 25 microsec. and as long 50 millisec. are available. The latter are useful in viewing very phenomena and in television, where it is helpful in reducing flicker.

COMPARISON MEASUREMENTS

32. Comparison of Voltages. A steady voltage may be compar with the difference of potential across a resistance-carrying current b the use of the simple potentiome shown in Fig. 32*a*.

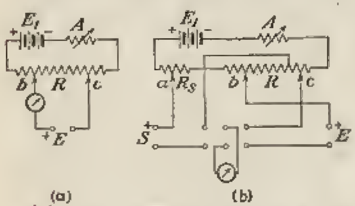


Fig. 32.—Potentiometer types: (*a*) simple; (*b*) with standard cell resistance.

A battery E_1 causes a current I flow in a resistance R. The unkno voltage E is connected to this res ance through a galvanometer, the resistance is adjusted to give deflection of the galvanometer. T voltage E is then equal to the pote drop IR. A second voltage E' m then be made equal to a different tential drop IR'. The two curre in the two cases are the same bec at balance no current flows in the g vanometer circuit. The two voltages are thus proportional to the two res ances. The potentiometer may be made direct-reading in voltage by using standard cell for one of the comparison voltages and connecting it across such portion of the resistance that the current must be adjusted to a predetermi decimal value in order to obtain balance. The unknown voltage is th connected through the galvanometer and balance is restored by adjustmen of resistance R, which may now be calibrated directly in volts. Connectio for this type of measurement are shown in Fig. 32*b*.

Two alternating voltages may be compared by the potentiomet principle only when they have the same frequency and the same phas They must at every instant be equal and opposite in order that the gal nometer in series with them shall show no deflection. Hence the poten ometer current must be taken from the same source as the voltage to measured, and some form of phase-shifting device must be provided which the output current is independent of its phase.

Drysdale used a two-phase induction regulator, feeding one phase throu a resistance and the other through a capacitance in order to obtain the t currents in quadrature. Such a device P is shown in Fig. 33 connected d-c potentiometer. The galvanometer G_A is an a-c galvanometer having sensitivity comparable to that of the d-c galvanometer G_D. Since there no standard of a-c voltage, a standard cell is used to adjust the potentiomet current to its proper value. This value is read on a transfer ammeter which may be either of the electrodynamometer or insulated heater therm

couple type. Its zero may be suppressed mechanically to give the effect of a longer scale and hence a greater accuracy of reading. Switches K and K_1 are then thrown to connect the potentiometer to the a-c voltages and the a-c current adjusted to produce the same deflection in ammeter I. Vacuum-tube voltmeters and rectifier voltmeters whose resistances are large compared with the resistance of the potentiometer may be calibrated directly without using the phase shifter, by connecting them directly to the terminals E. The voltage applied to them may be calculated from the settings of the contacts b and c.

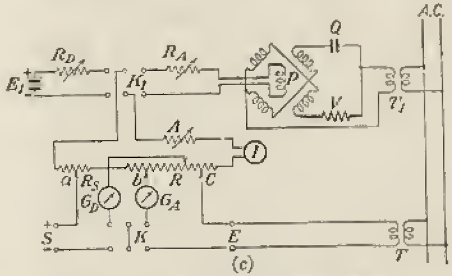


Fig. 33.—Drysdale potentiometer.

33. Comparison of Frequencies. Two nearly equal frequencies may be compared by measuring in a suitable manner their difference in frequency. When the two frequencies are in the audible range, this difference will appear as an audible beat—a waxing and waning in intensity—which may be counted if it is less than 10 beats per second. If the beats are faster than this or if the beating frequencies are above audibility, the beat must be rectified and a beat frequency produced. This beat frequency may then be measured by a suitable frequency meter. The accuracy of the comparison depends both on the accuracy of measurement of the beat frequency and on the ratio of this frequency to the original frequencies. The beat frequency is usually kept in the audible range.

If the two frequencies to be compared are not nearly equal, so that their frequency difference is large and above audibility, audible beats may usually be obtained between some of their harmonics. For a beat frequency b between the mth harmonic of a known frequency f and the nth harmonic of an unknown frequency f', the expression giving f' is

$$f' = \frac{mf \pm b}{n} \qquad (7)$$

the sign of b being determined by considering which harmonic, mf or nf' is the larger. Sufficient harmonics are usually present in most frequency sources for the purpose of this comparison, especially when emphasized and isolated by the use of tuned circuits. They can always be produced by the use of a rectifier tube.

In the most precise measurements the known frequency is a multiple or submultiple of a standard crystal frequency, obtained from the various multivibrators driven by the standard. For less precise work a variable standard may be used. The beat frequency is then made zero. Such a variable frequency oscillator, called a *heterodyne oscillator*, will have a limited frequency range, even though provided with multiple coils. Properly chosen for range, it may be used to measure a super-audio beat frequency, such as might be obtained when comparing two very high frequencies.

Frequency is measured in terms of inductance and capacitance by means of a tuned-circuit frequency meter consisting of a variable capacitance and a set of fixed inductances. The frequency range allotted to each coil determines the accuracy of setting, which ranges from 0.1 per cent to 0.001 per cent. Resonance is indicated in a variety of ways: thermocouple ammeter, heterodyne zero beat, or reaction on an oscillator, these being arranged in the order of their accuracy. In the third method the frequency meter is coupled closely enough to the oscillator whose frequency is being measured so that either the amplitude of its oscillation is affected or its frequency is altered. The frequency alteration is the more precise method but demands for greatest accuracy a second oscillator set at zero beat with the first. When the frequency meter is in exact resonance, the zero beat note of the two oscillators will be unaffected. In the second method a vacuum-tube oscillator is connected to the wavemeter so that it really becomes a heterodyne oscillator. A screen-grid tube, operating as a dynatron oscillator, may be connected to a frequency meter without the addition of extra coils or taps and converts it into a heterodyne-frequency meter.

34. Comparison of Impedances. An unknown resistance may be compared with a known resistance in a number of different ways. When the known resistance is variable, a *substitution method* may be employed.

The unknown resistance X is connected in series with a battery and shunted galvanometer g, the shunt resistance M having been adjusted to allow a full-scale deflection. The known variable resistance S is then substituted for X and the same current allowed to flow. Its value as thus determined is that of the resistance S. When the known resistance is not continuously variable, the value of the unknown resistance may be interpolated from the two readings of the meter. This method is frequently used for the measurement of very high resistances, such as insulation resistances from a megohm up. The known resistance is rarely larger than 1 megohm; hence under these conditions different values of the shunt M are used for the two measurements. The method is not applicable to measurements with a.c. because the phase angles of the source and load are indeterminate.

Two resistances may be compared by connecting them in series and measuring the voltage drops across them by means of a high-resistance voltmeter. Since the same current flows in both resistances, the value of the unknown resistance is

$$R = S\frac{E_R}{E_S}$$

where E_R and E_S are the voltages across the unknown and known resistances respectively. Except for the case of equal resistances, the resistance of the galvanometer must be either very large compared with the resistances being measured or a correction must be made for the current taken by the galvanometer. This method may be used with a.c. to compare all kinds of impedances. Either a vacuum-tube voltmeter or a high-resistance rectifier voltmeter must be used, since correction for the current taken by the voltmeter is difficult. The polarity of the voltmeter should be maintained as in d-c measurements in order to eliminate the errors of these voltmeters due to even harmonics. The upper limit for frequency is that imposed by the frequency characteristics of the known standard and by the capacitances to ground of the voltmeter in its two positions.

The power factor of an unknown impedance may be determined by the *three-voltmeter method*, in which the voltages across the unknown and known impedances and that applied to the two in series are read. The same precautions concerning polarity and capacitances to ground apply as in the two-voltmeter method. The vectorial relations between the three voltmeter readings together with the voltage components of the unknown impedance are shown in Fig. 34.

The expressions giving the unknown impedance Z, its resistance R, reactance X, and power factor $\cos\theta$ are

$$Z = S\frac{E_Z}{E_S}$$

$$R = S\frac{E^2 - E_Z^2 - E_S^2}{2E_S^2}$$

$$X = \sqrt{Z^2 - R^2}$$

$$\cos\theta = \frac{R}{Z} = \frac{E^2 - E_Z^2 - E_S^2}{2E_Z E_S} \tag{9}$$

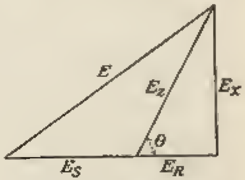


FIG. 34.—Vectorial relations in three-voltmeter circuit.

35. Variation Methods. The total resistance of a circuit may be measured by the *resistance-variation method*. Since with a constant applied voltage the current flowing in the circuit is inversely proportional to the total resistance, the circuit resistance is given by

$$R = S\frac{I'}{I - I'} \tag{10}$$

where I is the initial current and I' the current which flows when the resistance S is added. A plot of the reciprocal of the current flowing for different values of the added resistance against that resistance gives a straight line whose negative intercept on the resistance axis is the circuit resistance. The added resistance necessary to halve the current is also the circuit resistance. This method is sometimes used to measure the resistance of a sensitive galvanometer.

The resistance-variation method may be used with a.c. provided the circuit is tuned to resonance. The necessary connections are shown in Fig. 35. By reducing the reactance of the circuit to zero, the same equations and procedure may be used as for d.c. The ammeter used is usually of the thermocouple type. Halving the current on such a meter quarters the deflection; hence this type of measurement is sometimes called the *quarter-deflection method*. The ammeter may be replaced by a vacuum-tube voltmeter connected across the condenser. This arrangement is much more sensitive than the thermocouple ammeter and simplifies the grounding of the circuit by eliminating one series element. The upper limit for frequency is set by the frequency characteristic of the known resistance and the capacitances to ground of the different parts of the circuit. This method is the one usually adopted for the measurement of the resistance of inductors at high frequencies.

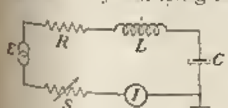


FIG. 35.—Added-resistance method.

Two reactances may be compared in a tuned circuit by a substitution method. The circuit is tuned to resonance both when the unknown reactance is connected in circuit and when it is disconnected. The change in reactance of the variable standard, with which the circuit is tuned, equal to the unknown reactance. When the unknown and known reactances are both inductive or both capacitive, the value of the unknown inductance or capacitance is obtained directly, independent of frequency, the two reactances being connected in series if inductive, and in parallel if capacitive. For these pairs of measurements it is unnecessary that the currents be kept of the same value.

Air condensers are much better standards at high frequencies than inductors, and it is therefore usual to measure an unknown inductance in terms of a variable condenser. Small inductances are connected in series and large inductances in parallel.

The resistance of the unknown reactance may be determined by noting the current at resonance when it is connected in circuit and then adjusting the current to this same value by adding sufficient resistance when it is disconnected. This added resistance, corrected for the change in resistance of the standard reactance with setting, is the resistance of the unknown reactance. The resistance of variable reactors must in general be measured by the added-resistance method described above or by one of the bridge methods. The resistance of a variable air condenser follows a definite law, and this fact may be used in this type of resistance measurements.[1]

The total resistance of the tuned circuit may also be measured by detuning the circuit. This method is called the *reactance-variation method*.[2] The change in reactance necessary to halve the squared current (deflection of a thermocouple meter) or to reduce the reading of a vacuum-tube voltmeter in the ratio of 1 to $\sqrt{2}$ (0.707) is equal to the resistance of the circuit. The resistance of an unknown reactance may be found by again measuring the total resistance of the circuit when the unknown is added. The difference in circuit resistance with the unknown in and out is the unknown resistance. The circuit resistance for the one case can also be found from the other by multiplying the known circuit resistance by the ratio of the voltmeter readings at resonance.

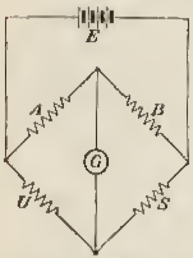


Fig. 36.—Wheatstone bridge.

D-C BRIDGE MEASUREMENTS

36. Whenever two resistances or impedances are compared by matching or comparing the deflections of any deflecting instrument, the accuracy of the measurement is determined by the accuracy of reading of the deflections themselves. This accuracy may be greatly increased by adopting a null method, in which a certain relation of the resistances being compared is indicated by a zero deflection. As this condition is approached, the sensitivity of the indicating instrument may be increased.

37. Four-resistance Network. The simple four-resistance network invented by Christie in 1833 and exploited by Wheatstone ten years later is shown in Fig. 36.

Two paths are provided for the current, one through the ratio arms A and B, the other through the unknown and known resistances U and S. The

[1] See Art. 44.
[2] Sinclair, D. B., *Proc. I.R.E.*, **26**, No. 12, 1466–1497.

galvanometer G is connected between the junctions of these pairs of resistances. The condition for a null deflection of the galvanometer is that these two junctions are at the same potential. Equating the voltage drops

$$AI_A = UI_U \quad \text{and} \quad BI_B = SI_S \tag{11}$$

or, since no current flows in the galvanometer,

$$\frac{A}{B} = \frac{U}{S} \quad \text{or} \quad U = \frac{A}{B}S \tag{12}$$

The ratio arms are usually only variable in steps of 10 so that the bridge is balanced by varying the known resistance S.

In commercial bridges the accuracy ranges from 0.1 to 0.02 per cent. Switching is accomplished by sliding contact-decade switches or taper plugs, and the ratio arms are reversible. There are four to six decades in the known resistance, hundredths to hundred thousands, and up to nine ratios, 0.0001 to 10,000. Comparisons of resistances on the best bridges using sealed standards, flat mercury contacts, and a temperature-controlled oil bath may be made to 1 part in 1,000,000, which is beyond the accuracy with which the primary standard of resistance is known.

38. The sensitivity of the null detector necessary to attain a given accuracy of bridge balance is determined by the relative magnitude of the resistances of the bridge arms and the voltage applied to the bridge. The ratio of the output voltage e to the input voltage E is given by

$$\frac{e}{E} = \frac{G/B}{1 + \frac{A}{B}} \cdot \frac{A/B}{\frac{A}{B}\left(1 + \frac{S}{B}\right) + \frac{G}{B}\left(1 + \frac{A}{B}\right)} d \tag{13}$$

where G is the resistance of the null detector and d is the fractional accuracy of balance demanded. For an equal-arm bridge

$$\frac{e}{E} = \frac{1}{4} \frac{G/B}{1 + \frac{G}{B}} d \tag{14}$$

This ratio lies between $\frac{1}{8}d$ and $\frac{1}{4}d$ for ratios of detector and bridge-arm resistances between one and infinity. In general, its value decreases rapidly when the bridge arms are made unequal and when the detector resistance is low compared to them. On this account resistances above a megohm cannot be accurately measured when a d-c galvanometer is used as a null detector.

For a very high resistance detector, Eq. (13) becomes

$$\frac{e}{E} = \frac{A/B}{\left(1 + \frac{A}{B}\right)^2} d \tag{15}$$

which is independent of the ratio S/B. This condition may be realized by the use of a vacuum-tube voltmeter as described in Art. 29. Thus for greatest sensitivity the detector should be connected from the junction of the highest resistances to the junction of the lowest. The battery, on the other hand, should be connected across the higher and lower resistance pairs, so that the amount of power drawn by the bridge is a maximum.

39. Slide-wire Bridges. When the known resistance is fixed, the bridge must be balanced by varying one or both of the ratio arms. In the slide-wire bridge shown in Fig. 37a the ratio arms A and B are p of a single uniform resistance along which the contact of the lead t galvanometer may slide. The position of the contact is read as a dista measured from one end, the whole length of the scale being L divisi The value of the unknown resistance in terms of these distances is

$$U = \frac{l}{L - l} S \quad (1$$

When the known and unknown resistances are nearly equal, the accura of measurement may be increased by placing extension coils in seri with the slide wire as shown in Fig. 37b. The slide wire may be calibrat

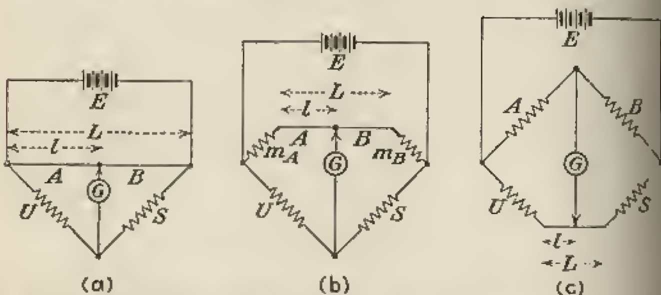


Fig. 37.—(a) Slide-wire bridge; (b) bridge with extension arms; (c) Car Foster bridge.

to read directly the percentage error of the unknown resistance U i terms of the standard resistance S.

Two nearly equal resistances may also be compared by means of th Carey Foster bridge shown in Fig. 37c. This is a slide-wire bridge i which the slide wire is placed between the two resistances being com pared. Two settings of the slide wire l and l' are made with the resist ances U and S as shown in Fig. 38 and transpose

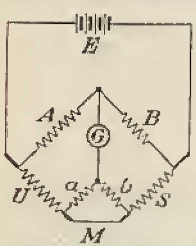


Fig. 38.—Kelvin double bridge.

The value of the unknown resistance is

$$U = S - (l - l')\rho \quad (1$$

where ρ is the resistance per unit length of the slid wire.

40. Kelvin Bridge. In the measurement of 0. ohm or less, the variation in contact resistance its terminals and the consequent variation in th lines of current flow near the terminals may pr duce appreciable errors. To overcome this diff culty, low-resistance standards are always built a four-terminal resistances. All ammeter shunts a so constructed. The two potential terminals are placed between the cur rent terminals and the resistance proper. The value of the resistance i that between the potential terminals.

Such four-terminal resistances cannot be compared on the ordinary Wheatstone bridge. They may be measured on the Kelvin double bridge shown in Fig. 38. The two four-terminal conductors U and S are connected in series, leaving an unknown resistance M between their adjacent potential terminals. The bridge is balanced by adjustment of the standard resistance S. The value of the unknown resistance U is given by

$$U = \frac{A}{B} S \quad (18)$$

when the double ratio arms are proportional, satisfying the condition $A/B = a/b$.

A-C BRIDGE MEASUREMENTS

41. Four-impedance Network. When an alternating voltage is applied to the simple Wheatstone bridge of Fig. 36, the conditions for balance of the bridge involve the impedances of the four arms, as shown in Fig. 39.

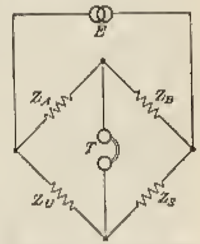


Fig. 39.—A-c bridge.

For a null deflection of the a-c galvanometer or telephones the two junctions, across which it is connected, must be at the same potential at all instants of the a-c cycle. Equating the voltage drops along the two parallel paths offered to the flow of the a.c.

$$Z_A I_A = Z_U I_U \text{ and } Z_B I_B = Z_S I_B \quad (19)$$

where Z_A, Z_B, etc., replace A, B, etc., in Fig. 37.

The four impedances are vectors of the form

$$Z = R + jX \quad (20)$$

Hence, since no current flows in the galvanometer,

$$\frac{Z_A}{Z_B} = \frac{Z_U}{Z_S} \quad (21)$$

Expanding these vectors into their rectangular components the two conditions of balance are

$$\frac{A}{B} = \frac{U}{S} + \frac{X_A X_S - X_B X_U}{BS} = \frac{X_U}{X_S} + \frac{U X_B - S X_A}{B X_S} \quad (22)$$

where the resistance components of the four arms are represented by the four letters A, B, U, S without subscripts. If the ratio arms have no reactance, so that $X_A = X_B = 0$, these conditions reduce to

$$\frac{A}{B} = \frac{U}{S} = \frac{X_U}{X_S} \quad (23)$$

The two reactances must have the same ratio as their resistances and as the ratio arms. Considering the reactances as both inductive or both capacitive, Eq. (23) becomes

$$\frac{A}{B} = \frac{U}{S} = \frac{L_U}{L_S} \quad \text{and} \quad \frac{A}{B} = \frac{U}{S} = \frac{C_S}{C_U} \quad (24)$$

respectively. These equations cover all the types of bridge measurements in which similar impedances are compared.

42. Power Supply and Null Detector. The power source at audio and radio frequencies is usually a vacuum-tube oscillator, capable of supplying several hundred milliwatts of power at varying potentials up to 100 volts. At the low audio frequencies, a-c generators with rotating parts may be used, as well as the commercial power supply at 60 and 25 cycles. The null detector most frequently used in the a-f range from 400 to 5,000 cycles is the head telephone. Vibration galvanometers and a-c moving-coil galvanometers are used at power frequencies. Rectifier voltmeters are used for all frequencies up to 50 kc, cathode-ray and "magic-eye" tubes up to 1 Mc, and vacuum-tube voltmeters at all frequencies. At super-audio frequencies a heterodyne oscillator and detector may be used to produce an a-f beat note, which can then be observed by any of the methods described. Radio-frequency oscillators may be modulated at an a.f., usually 1 kc, and the bridge output observed on a radio receiver. All-wave receivers cover the frequency range from 10 kc to 30 Mc.

Vacuum-tube amplifiers are used with all types of null detectors to give increased sensitivity. The amount of amplification necessary to give any desired accuracy of balance may be determined by Eq. (15) when the generator is connected across resistive ratio arms. If the generator is placed across unlike arms, one resistive and one reactive, the expression becomes

$$\frac{e}{E} = \frac{A/B}{1 + \left(\frac{A}{B}\right)^2} d \qquad (25)$$

At the most Eqs. (15) and (25) differ by only a factor of 2. These two equations hold exactly for the larger component of impedance, provided that the square of the ratio of the small to the large component is negligible compared to unity. The value of e/E for the smaller component is then less than that for the larger component by their ratio. The vibration and a-c moving-coil galvanometers are about equally sensitive, with a minimum detectable voltage of 20 μv, although a moving-coil galvanometer can be built with a sensitivity of 0.1 μv. Head telephones come next with a minimum detectable voltage of 400 μv. Then in turn come "magic-eye" tubes at 20 mv, vacuum tube and rectifier voltmeters at 100 mv, and cathode-ray tubes at 1 volt.

A considerable amount of selectivity is desirable in a null detector to eliminate the effect of harmonics in the generator and harmonics produced by non-linearity of the unknown impedance. This can be provided by a tuned circuit in the amplifier or by the degenerative feedback amplifier described by Scott.[1] This latter amplifier is particularly valuable because it can be made continuously adjustable over the entire a-f range. The former gives a discrimination of 25 db against the second harmonic and the latter 40 db. The vibration galvanometer is extremely selective and offers about 70 db against the second harmonic. The a-c galvanometer is phase sensitive and responds only to that component of the unbalance voltage which is in phase with its field. It can therefore be made to respond to only one component of bridge balance at a time by connecting its field to a suitable phase-shifting network. The cathode-ray tube can be used in a somewhat similar manner by applying the bridge voltage to

[1] SCOTT, H. H., *Proc. I.R.E.*, **26**, No. 2, 226–235.

its horizontal deflecting plates as a sweep circuit through a phase shifting network.[1] The general pattern appearing on the screen is a tilted ellipse, which at balance reduces to a horizontal line. The phase of the sweep voltage can be so adjusted that one component of bridge balance opens the ellipse while the other tilts it.

43. Bridge Transformers. Transformers are used to match the impedance of a bridge to that of the generator or detector and to isolate the bridge electrostatically. One junction point of the bridge, usually that between the two impedances being compared, is grounded, except when direct impedances are measured.[2] The capacitances to ground of the transformer, generator, or detector not connected to this grounded junction are placed across the two bridge arms whose junction point is grounded. The effect of the ground capacitances of the generator or detector connected to the transformer may be removed by placing a grounded shield between the primary and secondary windings. An impedance bridge with such a transformer connected across its ratio arms is shown in Fig. 40. The terminal capacitances C_{TU} and C_{TS} are placed across the bridge arms U and S. They are usually of the order of several hundred micromicrofarads and may therefore introduce serious errors. The direct capacitance between the two windings may be reduced to a few tenths of 1 $\mu\mu$f.

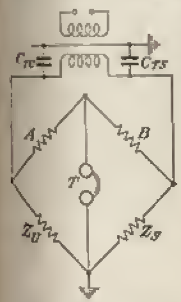


FIG. 40.—Impedance bridge with transformer.

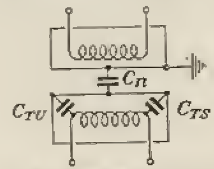


FIG. 41.—Bridge-transformer capacitances.

The effect of the terminal capacitances C_{TU} and C_{TS} may be reduced or either one made zero by the addition of a second shield. The two shields are symmetrically placed around the two windings, as shown in Fig. 41. The capacitance C_n between the two shields may be made much smaller than the terminal capacitances and is in series with them. The resultant terminal capacitances may be placed across either bridge arm U or S by connecting the shield around the secondary winding to one terminal of that winding. The effect of the terminal capacitances may be removed entirely from the arms U and S and placed across the ratio arms A and B by introducing a third shield between the two winding shields and connecting it to the junction of the ratio arms.

44. Bridge Errors. Reactances introduced into the arms of a bridge by the wiring of the bridge and by the generator and detector cause the more serious errors in bridge measurements. These residual reactances may be inductances in series with the bridge arms and capacitances in parallel with them. The effect of such residuals in the ratio arms may be seen by rewriting Eq. (22) of Art. 41 in the approximate form

[1] LAMSON, H., *Rev. Sci. Inst.*, **9**, No. 9, 272–275.
[2] See Art. 46.

$$\left.\begin{aligned}\frac{A}{B} = \frac{U}{S}\left[1 + (Q_A - Q_B)\frac{1}{D_U}\right] &= \frac{X_U}{X_S}[1 - (Q_A - Q_B)D_U] \\ D_S - D_U &= Q_A - Q_B\end{aligned}\right\} \quad (26)$$

where the *storage factors* Q_A and Q_B and the *dissipation factors* D_U and D_S are of the form

$$Q = \frac{X}{R} \quad \text{and} \quad D = \frac{R}{X} \quad (27)$$

The errors introduced are proportional to the difference of the storage factors of the ratio arms, multiplied by the dissipation factor of the impedance arms for the reactance component, and divided by the dissipation factor for the resistance component. For impedances with small dissipation factors the error is confined to the resistance component, for impedances with large dissipation factors to the reactance component.

Residual reactances in the impedance arms produce at low frequencies errors proportional to their ratio with similar reactances in these arms. Series inductance introduces large errors in measurements of small inductances; parallel capacitance in measurements of small capacitances.

The effect of residual reactances increases with frequency, the storage factor of the ratio arms being of the form $Q = \omega L/R$ for series inductance and $Q = R\omega C$ for parallel capacitance. Hence bridges designed for operation at frequencies much above 100 kc must have equal ratio arms because of the difficulty of equalizing their storage factors. When residual inductance in the impedance arms is in series with a capacitance the effective capacitance of the combination is

$$\hat{C} = \frac{C}{1 - \omega^2 LC} \quad (28)$$

which increases indefinitely as the resonant frequency is approached. For an inductance of 1 μh, the approximate value for a constant-inductance three-dial decade resistor, and a capacitance of 1,000 $\mu\mu$f the resonant frequency is 5 Mc. Even the lowest inductance which a 1,000-$\mu\mu$f condenser can have, 0.006 μh, gives a resonant frequency of 65 Mc.

The errors introduced into bridge measurements by reactances in the ratio arms may be minimized by the use of substitution methods. The effect of capacitances to ground and the effect of the reactance of the leads to the known and unknown reactances may also be greatly reduced. Both reactances are connected in the same arm of the bridge, a similar reactance being placed in the other arm. Two bridge balances are obtained, one with the unknown reactance in circuit, the second with it disconnected and its impedance replaced by the known variable reactance and the added resistance. Inductances are connected in series, placing them far enough apart to reduce their mutual inductance to a negligible amount, and the unknown is removed by shorting. Capacitances are connected in parallel, and the unknown is removed by disconnecting its high-potential terminal. Both condensers must be completely shielded and their grounded terminals connected together.

Distinguishing the values for the second balance, when the unknown reactance has been removed, by primes, the values of the unknown reactance are given by the change in reactance of the variable standards.

$$\begin{aligned} L_U &= L_S' - L_S & C_U &= C_S' - C_S \\ &= \Delta L_S & &= \Delta C_S \end{aligned} \quad (29)$$

The corresponding expressions for the resistances are

$$\begin{aligned} U &= S' - S + R' - R & U &= (R' - R)\left(\frac{C_S'}{C_U'}\right)^2 \\ &= \Delta S + \Delta R & &= \Delta R\left(\frac{C_S'}{C_U'}\right)^2 \end{aligned} \quad (30)$$

The squared terms appearing in the expression for the condenser resistance result from the law by which the series resistance of condensers connected in parallel is found.

$$R = \frac{R_1C_1^2 + R_2C_2^2 + \cdots}{(C_1 + C_2 + \cdots)^2} = \frac{\sum_1^n R_m C_m^2}{\left(\sum_1^n C_m\right)^2} \quad (31)$$

The terms containing the resistance of the standard condenser have disappeared because the quantity RC^2 for an air condenser is a constant, independent of the setting of the condenser. This follows from the more general law that for an air condenser, in which the losses occurring in the solid dielectric are independent of the setting of the plates and for which the power factor of the solid dielectric is independent of frequency, the quantity $R\omega C^2$ is constant. This law holds with increasing frequency until the losses due to skin effect in the plates and supports become appreciable.

The series resistance of the plates and supports of a well-designed air condenser is of the order of 0.02 ohm at a frequency of 1 Mc.[1] This resistance varies as the square root of the frequency because even at 1 Mc the skin effect is complete. By shortening the leads and by connecting to the stator and rotor at several points, this series resistance can be reduced to 0.005 ohm at 1 Mc.

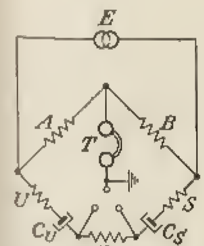


FIG. 42.—Series-resistance bridge.

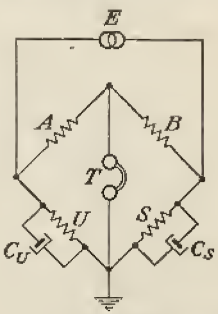


FIG. 43.—Parallel-resistance bridge.

46. Resistance Balance. When two impedances are compared on a four-impedance bridge, the conditions of balance [Eq. (24) of Art. 41] demand that their dissipation factors be equal. Since this will not in general be the case, means must be provided for attaining the resistance balance. The simplest method is that of adding a resistance in series with that impedance having the lower dissipation factor. The connections for a capacitance bridge with the added resistance so arranged

[1] FIELD, R. F., and D. B. SINCLAIR, *Proc. I.R.E.*, **24**, No. 2, 255–274.

that it may be placed in either impedance arm is shown in Fig. 42. The method gives the series resistance and reactance of the unknown impedance and can be used for dissipation factors less than unity. Neither of the impedances, although essentially at ground potential, can be grounded.

Added resistances may be placed in parallel with the two impedance arms as shown in Fig. 45. This method gives the parallel resistance and reactance of the unknown impedance and is best adapted to the measurement of impedances having dissipation factors greater than unity. For small dissipation factors the shunting effect of the parallel resistances is such as to reduce markedly the sensitivity of the bridge balance. One terminal of each impedance is grounded.

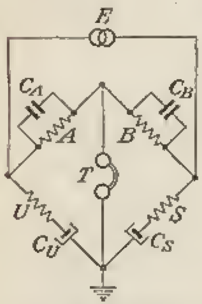


FIG. 44.—Thomas' method.

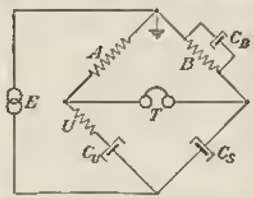


FIG. 45.—Schering bridge.

The resistance balance may also be made by adding suitable reactances to the ratio arms. Rosa in 1907 suggested the use of series inductance, while Thomas in 1914 used parallel capacitance, as shown in Fig. 44. The balance equations are

$$C_U = \frac{B}{A}C_S \text{ (approx). and } U = \frac{A}{B}S + A\left(\frac{C_B}{C_S} - \frac{C_A}{C_U}\right) \qquad (3$$

whence

$$D_U = D_S + Q_A - Q_B$$

Schering in 1920 used a parallel capacitance across one ratio arm in a high-voltage bridge connected as shown in Fig. 45. The generator was connected from the junction of the resistance arms to the junction of the capacitance arms, both to minimize the power losses in the ratio arms and to keep constant the voltage applied to the unknown condenser. The junction of the resistance arms was grounded in order to keep the ratio arms and the detector at a low voltage with respect to ground.

Any bridge, in which the resistive balance is made by adding capacitance across a ratio arm, is now called a *Schering bridge* regardless of the position of the ground or the generator connections. If the junction of the capacitance arms is grounded, it is called an *inverted Schering bridge*. When the generator is connected across the ratio arms, it is called a *conjugate Schering bridge*.

46. Direct Capacitance. Any capacitance having terminal capacitances to a surrounding shield or to ground may be represented as a three-terminal capacitance, as shown in Fig. 46. The capacitance C_D between the terminals 1 and 2 is called the *direct capacitance*. The total capacitance between these terminals is the sum of the direct capacitance C_D, and the two terminal capacitances C_{T1} and C_{T2} in series. The direct capacitance may be measured on a bridge by connecting the shield to either of the junction points of the bridge, to which the direct capacitance is not connected. These two connections are shown in Fig. 47.

Errors due to placing the terminal capacitances across the bridge arms greatly limit the usefulness of these connections. When the shield is connected to the junction of the ratio arms, the terminal capacitance C_{T1} is placed across the arm A and produces an error $A\omega C_{T1}$ in the determination of the dissipation factor of the direct capacitance C_D. The terminal capacitance C_{T2} and any capacitance of the shield to ground are placed across the detector T. When the shield is connected to the junction of the arms B and S, the terminal capacitance C_{T2} is placed across the impedance arm S and produces an error in the determination of the direct capacitance C_D unless the standard capacitance C_S is very large compared to C_{T2}. Any capacitance of the shield to ground is also placed across C_S, while the terminal capacitance C_{T1} is placed across the generator E. If the direct capacitance C_D is not surrounded by a shield, the terminal capacitances C_{T1} and C_{T2} are to ground, and neither of these methods is applicable.

FIG. 46.—Three-terminal capacitance.

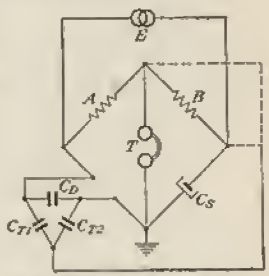


FIG. 47.—Measurement of direct capacitance.

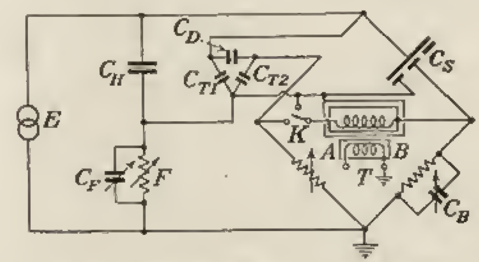


FIG. 48.—Schering bridge with guard circuit.

47. Guard Circuit. The use of a guard circuit enables both direct capacitance and its dissipation factor to be measured correctly, because the terminal capacitances are not connected across any of the bridge arms. A Schering bridge with guard circuit and shielded output transformer is shown in Fig. 48. The terminal capacitance C_{T1} is placed across the guard capacitance C_H, while the other terminal capacitance C_{T2} couples the guard circuit to the junction of the bridge arms A and C_D. The standard condenser C_S is also a three-terminal condenser. The advantages of this construction are that all losses in the insulating supports can be carried to the guard circuit and that no capacitance will be added across ratio arm B. The guard capacitance of this condenser is thus placed across guard capacitance C_H. Frequently this capacitance and the capacitance C_{T1} make up C_H entirely, and it becomes unnecessar to provide an extra high-voltage condenser. The transformer has third shield mentioned in Art. 43; consequently no ground capacitan are placed across the ratio arms. Instead the capacitance between t third shield and the bridge winding shield couples the guard circuit to junction of the bridge arms B and S.

Because of the existence of the capacitances coupling the guard circ to the bridge, the conditions of balance of the bridge involve the balanc of the guard circuit.

$$\frac{Z_A}{Z_U} = \frac{Z_B}{Z_S} = \frac{Z_F}{Z_H}. \qquad (3$$

This is done by disconnecting one terminal of the output transformer fr the bridge by means of switch K and transferring it to the guard circu The new bridge circuit formed by the arms B, S, F, H is then balan to satisfy the right half of Eq. (33), by adjusting the guard circuit. Su cessive balances of bridge and gua circuits must be made until both par of Eq. (33) are satisfied. The accura with which the guard circuit must balanced in order that no apprecia error is introduced into the bridge b ance depends both upon the magnitud of the coupling capacitances betwe the guard circuit and the bridge a also upon the degree with which th bear the same ratio to each other the capacitances C_S and C_D. The c cuit formed by these coupling capac tances is called the *coupling circ* Its relation to the guard circuit is show in Fig. 49. By definition the guar circuit is that circuit which is connect across the generator, while the coupli circuit is connected across the detector. Either circuit can therefore composed of similar or dissimilar elements. The circuit devised by W ner in 1911 for the same purpose and called a *Wagner ground* was alwa composed of similar elements and connected across the ratio arms, wh the generator was also connected. By the above definition it was a gua circuit.

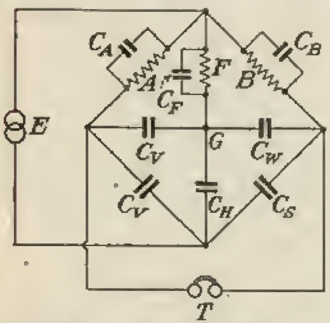


Fig. 49.—Schering bridge with guard circuit and coupling circuit.

Balsbaugh[1] has shown that for the network of Fig. 49 the conditio of balance are either those given in Eq. (33) or those given in Eq. (3

$$\frac{Z_A}{Z_B} = \frac{Z_U}{Z_S} = \frac{Z_V}{Z_W} \qquad (3$$

If either the guard circuit or the coupling circuit is partially balanced the other circuit need be only partially balanced in order to introduce n appreciable error in the bridge balance. Balance of both circuits m be conveniently made without disconnecting either generator or detect by connecting their junction G to those corners of the bridge which pla the guard or coupling circuit in parallel with similar elements of c

[1] Balsbaugh and Herzenberg, *Jour. Franklin Inst.*, **218**, No. 1, 49–67.

parable impedance. Ground may be placed at any corner of the bridge or at the junction G of guard and coupling circuits. This latter point is best in many respects because it simplifies the mechanical construction of the bridge and avoids the need of any insulated shields. All elements used in the bridge must, however, be three-terminal impedances. Grounding the point G also simplifies the construction of the three-terminal measuring cell by making it unnecessary to provide an insulated shield inside the outer grounded case.

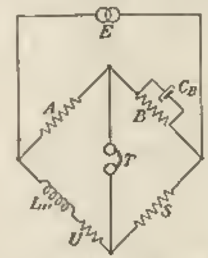


Fig. 50.—Maxwell bridge.

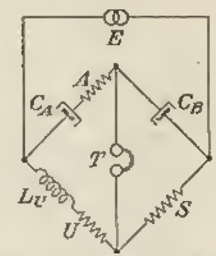


Fig. 51.— Owen bridge.

48. Comparison of Inductances and Capacitances. An inductance and a capacitance may be compared directly by suitably placing them in the four-impedance network. The connections for the Maxwell bridge are shown in Fig. 50.

The balance equations are

$$L_U = ASC_B \qquad \text{and} \qquad U = \frac{A}{B}S \qquad (35)$$

whence

$$Q_U = Q_B$$

Losses in the condenser C_B enter only into the resistance balance and may be made negligible by suitable choice of resistance A. The resistance and reactance balances are not independent unless condenser C_B is continuously variable or resistance is added in series with the unknown inductor.

In the Owen bridge an inductance is compared with a capacitance in the manner shown in Fig. 51.

The balance equations are

$$L_U = ASC_B \qquad \text{and} \qquad U = \frac{C_B}{C_A}S \qquad (36)$$

whence

$$Q_U = Q_A$$

The resistance balance is made either by having condenser C_A continuously variable or by adding resistance in series with the unknown inductor.

The Hay bridge may be considered the complement of the Maxwell bridge with the resistance and capacitance in the B arm connected in series instead of in parallel. On this account, however, it is not independent of frequency. The connections are shown in Fig. 52. The conditions of balance are

$$L_U = \frac{ASC_B}{1 + B^2\omega^2C_B^2} \qquad \text{and} \qquad U = \frac{ABS\omega^2C_B^2}{1 + B^2\omega^2C_B^2} \qquad (37)$$

whence

$$Q_U = \frac{1}{D_B}$$

When Q_U is greater than 10, the error in the expression for L_U caused by neglecting the frequency term in the denominator is less than 1 per cent. The two bridge balances are not independent unless the condenser C_B is continuously variable or resistance is added in series with the unknown conductor.

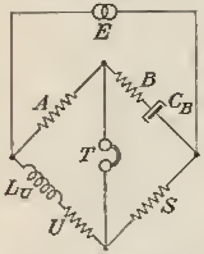


FIG. 52.—Hay bridge.

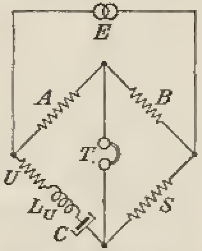


FIG. 53.—Resonance bridge.

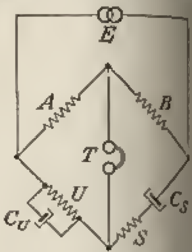


FIG. 54.—Wien bridge.

49. The resonance bridge shown in Fig. 53 is the simplest bridge in which inductance, capacitance, and frequency enter. At balance the arm containing the reactances is resonated to the applied frequency and becomes a pure resistance. The bridge is then an all-resistance equal-arm bridge. For this reason it may be used at high frequencies to measure the resistance and inductance of a reactor.

The balance equations are

$$\omega^2 = \frac{1}{L_U C_U} \qquad \text{and} \qquad U = \frac{A}{B}S \tag{38}$$

This bridge is frequently used to measure frequency, usually in the a-f range. A variable inductor is used, and the condenser may be varied in steps. A range from 200 cycles to 4 kc may be covered in three ranges. The frequency scale is irregular, owing to the characteristics of variable inductors, and the various ranges cannot be made multiples of one another. Owing to the large stray field of the variable inductor, its magnetic pickup is considerable. A resistance balance must be provided to allow for the variation of the resistance of the tuned arm with frequency.

It is equally possible in the resonance bridge to place the unknown inductor and condenser in parallel. Equation (38) still holds except that U will be the equivalent series resistance of the parallel circuit.

50. Wien Bridge. Capacitances may be measured in terms of resistance and frequency with the Wien bridge, shown in Fig. 54. The balance equations expressed in their simplest form are

$$\omega^2 = \frac{1}{USC_UC_S} \qquad \text{and} \qquad \frac{C_U}{C_S} = \frac{B}{A} - \frac{S}{U} \tag{39}$$

Solving for the two capacitances,

$$C_U{}^2 = \frac{BU - AC}{AU^2S\omega^2} \qquad \text{and} \qquad C_S{}^2 = \frac{A}{(BU - AS)S\omega^2} \tag{40}$$

The bridge is valuable because the standards of frequency and resistance are known to a greater accuracy than the standard of capacitance. Ferguson and Bartlett[1] have developed this method to its greatest precision. Their estimated accuracy for the determination of capacitance by this method is 0.003 per cent.

The Wien bridge also furnishes a very convenient means for measuring frequency in the a-f range. The two capacitances are made equal, while the two ratio arms are made such that B is twice A. The two resistances U and S are made variable over a suitable range but are also kept equal. Thus the resistance balance is always satisfied and the reactance balance reduces to

$$f = \frac{1}{2\pi UC_U} \tag{41}$$

In a commercial frequency meter the resistances U and S are wound on tapered cards so shaped that the frequency scale is logarithmic. This gives a constant fractional accuracy of reading. There are three frequency ranges, obtained from three different pairs of condensers, each covering a range of 10 to 1 in frequency. The same calibration serves for all ranges. The frequency limits attained are 20 cycles and 20 kc.

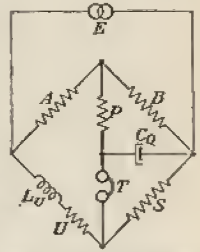


FIG. 55.—Anderson bridge.

51. Six-impedance Network. The six-impedance network was developed by Anderson to provide a modification of the Maxwell bridge which would render the two balance conditions independent even with a fixed capacitance. The connections are shown in Fig. 55.

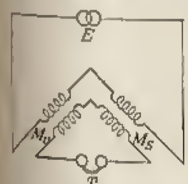


FIG. 56.—Felici mutual-inductance balance.

The general balance condition for the six-impedance network is

$$Z_Q(Z_BZ_U - Z_AZ_S) = Z_P[Z_P(Z_A + Z_B) + Z_AZ_B] \tag{42}$$

For the Anderson bridge this reduces to

$$L_U = SC_Q\left[P\left(1 + \frac{A}{B}\right) + A\right] \text{ and } U = \frac{A}{B}S \tag{43}$$

The effect of losses in the condenser C_Q is usually small.

52. Mutual-inductance Balances. Two mutual inductances may be compared by means of the *Felici mutual-inductance balance* shown in Fig. 56. The known mutual inductance must be variable. For the usual condition of balance, zero voltage across the null detector, the two mutual inductances are equal.

$$M_U = M_S \tag{44}$$

They must be so connected that their induced secondary voltages are in opposition. Mutual inductance between them should be avoided.

53. Four-impedance Network with Mutual Inductances. A mutual inductance may be compared with a self-inductance on a four-impedance bridge by placing it between one arm and either an input or output lead of the bridge, as shown in Fig. 57.

[1] FERGUSON and BARTLETT, *Bell System Tech. Jour.*, 7. No. 3, 420–437.

The general balance equation for this network is

$$Z_A Z_S - Z_B Z_U - j\omega M(Z_A + Z_B) = 0$$

For Campbell's arrangement of this bridge the two conditions of balance become

$$L_U = \frac{A}{B}L_S - \left(1 + \frac{A}{B}M\right) \quad \text{and} \quad U = \frac{A}{B}S$$

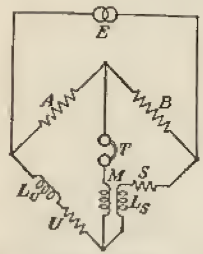


FIG. 57.—Comparison of mutual with self-inductance.

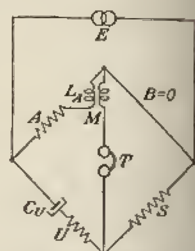


FIG. 58.—Carey Foster mutual-inductance bridge.

A substitution method is usually adopted so that the inductance and resistance of that portion of the mutual inductance connected in the S arm need not be known. When the ratio arms are equal, the extra balancing inductance represented by L_U of Fig. 57 may be eliminated by providing a center tap in one branch of the mutual inductance. This connection is usually referred to as the Heaviside equal-arm bridge.

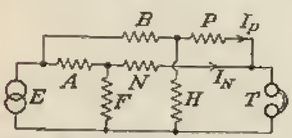


FIG. 59.—Parallel T network.

A mutual inductance may be compared with a capacitance by means of the *Carey Foster bridge*, shown in Fig. 58. The conditions of balance are

$$C_U = \frac{M}{AS} \quad \text{and} \quad U = S\left(\frac{L_A}{M} - 1\right)$$

The impedance of the B arm is made zero in order to make the balance independent of frequency. The method suffers because the resistance and self-inductance of the mutual inductance enter into the expressions for the unknown capacitance and its resistance, respectively. Capacitance between the two windings of the mutual inductance causes the voltage induced in its secondary to have a phase angle with reference to the primary current different from 90 deg. This reduces the calculated resistance of the condenser and frequently yields negative values, especially fo. large mica condensers. The method is perhaps better suited for the measurement of a mutual inductance in terms of a known condenser.

T NETWORKS

54. Two or more T networks connected in parallel provide a method of null balance which in many respects is equivalent to an a-c bridge circuit.[1] The connections for two T networks are shown in Fig. 59. The most important feature of the network is that generator and detector have a common terminal, which can be grounded. Hence no shielded transformer is necessary. This is a considerable convenience at low frequencies and makes it possible to use the network at high frequencies up to at least 30 Mc.

[1] TUTTLE, W. N., *Proc. I.R.E.*, **28**, No. 1, 23–29.

The condition for a null deflection of the detector is that the currents in the output circuits of the two networks shall be equal and opposite.

$$I_P + I_N = 0 \tag{48}$$

These currents are best evaluated by considering the transfer impedances, which are defined as the ratios of the input voltage to the output current.

$$Z_{TP} = \frac{E}{I_P} = Z_A + Z_N + \frac{Z_A Z_N}{Z_F}$$

$$Z_{TN} = \frac{E}{I_N} = Z_B + Z_P + \frac{Z_B Z_P}{Z_H} \tag{49}$$

Hence

$$Z_A + Z_N + \frac{Z_A Z_N}{Z_F} + Z_B + Z_P + \frac{Z_B Z_P}{Z_H} = 0 \tag{50}$$

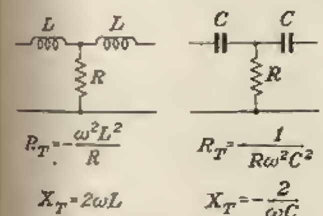


FIG. 60.—T network having negative transfer resistance.

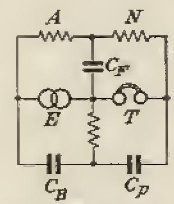


FIG. 61.—Parallel T network equivalent to Wien bridge.

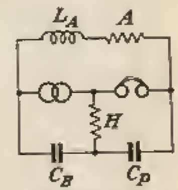


FIG. 62.—Bridged T network.

Under somewhat restricted conditions this equation can be satisfied because the impedances are complex quantities. While any of the terms of Eq. (50) can contain negative reactances, only the product terms can have a negative resistance. The only two T networks having a negative resistance component of transfer impedance are shown in Fig. 60. One of these networks or a modification must be used in every parallel T network which can be balanced.

55. Parallel T Networks. The parallel T network shown in Fig. 61 is equivalent to the Wien bridge[1] and has similar balance equations.

$$\omega^2 = \frac{C_H + C_P}{ANC_B C_P C_F} \quad \text{and} \quad \frac{C_B + C_P}{C_F} = \frac{1}{H}\frac{AN}{A + N} \tag{51}$$

When both of the T networks are made symmetrical and when in addition C_F is made twice C_B and A is made twice H, the resistance balance is always satisfied and the reactance balance reduces to

$$f = \frac{1}{2\pi A C_B} \tag{52}$$

which is identical in form with Eq. (41).

[1] See Art. 50.

56. Bridged T Networks. When the shunt arm of one of the T networks is made infinite, the circuit is called a *bridged T network*. The circuit shown in Fig. 62 is very convenient for measuring an inductance in terms of capacitance, resistance, and frequency. It is equivalent to the resonance bridge (Art. 49). The balance equations are

$$L_A = \frac{2}{\omega^2 C_B} \qquad \text{and} \qquad A = \frac{1}{H\omega^2 C_B^2} \tag{53}$$

whence

$$Q_A = 2H\omega C_B.$$

At balance the full generator voltage appears across the inductance. When the junction of generator and detector is grounded, the terminal capacitances of the inductor are placed across generator and detector, and the direct impedance of the inductor is measured.

References

Bolton, E. P.: "Electrical Measuring Instruments," E. P. Dutton & Co., Inc., New York, 1923.
Brown, H. A.: "Radio-frequency Electrical Measurements," 2d ed., McGraw-Hill Book Company, Inc., New York, 1938.
Cambridge Instrument Company, Catalogue.
General Radio Company, Catalogue.
Hague, B.: "Alternating-current Bridge Methods," 4th ed., Sir Isaac Pitman & Sons, Ltd., London.
Hund, A.: "High-frequency Measurements," McGraw-Hill Book Company, Inc., New York, 1933.
Jansky, C. M.: "Electrical Meters," McGraw-Hill Book Company, Inc., New York.
Kennelly, A. E.: "Electrical Vibration Instruments," The Macmillan Company, New York.
Laws, F. A.: "Electrical Measurements," 2d ed., McGraw-Hill Book Company, Inc., New York, 1938.
Leeds & Northrup Company, Catalogue.
Moullin, E. B.: "Radio Frequency Measurements," 3d ed., Charles Griffin & Company, Ltd., London.
Starling, S. G.: "Electricity and Magnetism," Longmans, Green & Company, London, 1924.
Terman, F. E.: "Measurements in Radio Engineering," McGraw-Hill Book Company, Inc., New York, 1935.
Weston Electrical Instrument Corporation, Catalogue.

SECTION 8

VACUUM TUBES

By J. M. Stinchfield, B.S.[1]

1. Electrons. The electron is a negatively charged particle of electricity. In 1897 J. J. Thomson discovered that the cathode rays passing from the cathode, to the anode in a gaseous discharge, were moving, negatively charged, particles. He measured the ratio of the charge e to the mass m of these particles and termed them corpuscles. Thomson's corpuscles are now commonly known as electrons. The cathode rays or streams of electrons are deflected by either magnetic or electrostatic fields. They exert mechanical force sufficient to turn a vane in a vacuum or to heat the object they strike.

2. Electrons in an Electrostatic Field. An electrostatic field exerts a force upon an electron. If the field intensity is X and the charge on the electron e, the force f acting on the electron is

$$f = Xe \tag{1}$$

If the mass of the electron is m, the acceleration a will be

$$a = \frac{Xe}{m} \tag{2}$$

The force and acceleration on the electron will change if the field intensity changes. The force is in the direction of the field at the point considered, the electron tending to move toward the positive.

In a uniform field the work W done on an electron in moving between two points distance s apart will be

$$\begin{aligned} W &= fs \\ &= Xes \end{aligned} \tag{3}$$

Since Xs is also the potential difference between the two points, calling this potential difference V, the work done on the electron is

$$W = Ve$$

If the field is not uniform the line integral of the force and distance regardless of the path between the two points will give the work done. The work done on a unit charge moved between two points defines the potential difference between the two points. The work done on an electron moved between two points of potential difference V will be

$$W = Ve \tag{4}$$

[1] Engineering Department, RCA Manufacturing Co., Inc., Radiotron Division.

If the velocity of an electron is changed by an amount v in passing between two points, the change in kinetic energy will be

$$\frac{mv^2}{2}$$

The change in potential energy or work done in passing between the two points will be

$$Ve$$

The change in kinetic energy is equal to the change in potential energy, and

$$Ve = \frac{mv^2}{2}$$

The velocity acquired by an electron in passing between two points of potential difference V is

$$v = \sqrt{\frac{2Ve}{m}}$$

The potential V is in absolute e.s.u. in the relations above. The potential difference in volts divided by 300 is the potential difference in absolute e.s.u.

The ratio of the charge e to the mass m of the electron is

$$\frac{e}{m} = \frac{4.774 \times 10^{-10}}{8.999 \times 10^{-28}} = 5.305 \times 10^{17} \text{ e.s.u. per gm}$$

The electron velocities corresponding to various potential differences are shown in the table. When the velocity becomes greater than about one-tenth the velocity of light, the apparent mass of the electron increases enough to cause a small error. The error in using Eq. (7) is less than one-half of 1 per cent for potential differences less than 300 volts.

Volts	Velocity, Centimeters per Second
1	0.00595×10^{10}
5	0.0133
10	0.0188
20	0.0266
30	0.0326
40	0.0376
50	0.0421
60	0.0461
70	0.0498
80	0.0532
90	0.0564
100	0.0595
200	0.0841
300	0.103
400	0.119
500	0.133
1,000	0.188

Volts	Velocity, Centimeters per Second
10,000	0.586×10^{10}
100,000	1.64
1,000,000	2.82

3. Electrons in an Electromagnetic Field. An electron moving with a velocity v in an electromagnetic field of intensity H is acted on by a force

$$f = Hev \tag{8}$$

The direction of the force is at right angles to both the direction of the field H and the direction of motion of the electron.

The force f is effective in producing an acceleration:

$$a = \frac{Hev}{m} \tag{9}$$

The acceleration is at right angles to the direction of motion. If the electron moves unimpeded and the field H is uniform, the path will be circular and of radius

$$r = \frac{v^2}{a} = \frac{mv}{eH} \tag{10}$$

4. Current Due to a Stream of Electrons. A current i is defined by the quantity of electricity q flowing per unit of time. If there are n electrons per unit of volume in a certain space, the quantity of electricity q in this space is ne per unit of volume. If these electrons are moved with a velocity v, the quantity flowing per unit of time is the current

$$i = nev \tag{11}$$

This is the current per unit of area at right angles to the direction of flow.

5. Space Charge Due to a Cloud of Electrons. If in a given space there are n electrons per unit of volume, the volume density of electrification is

$$\rho = ne \tag{12}$$

The potential distribution in the given space due to the electrons is given by

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi\rho \tag{13}$$

For the case of large parallel plates, only the distance x between plates need be considered. Equation (13) simplifies to

$$\frac{\partial^2 V}{\partial x^2} = -4\pi\rho \tag{14}$$

If a current i is flowing and the electrons move with uniform velocity v the space charge or volume density of electrification is

$$\rho = \left(\frac{i}{v}\right) \tag{15}$$

6. Emission of Electrons. Certain internal forces existing at the surfaces of substances prevent the escape of the free electrons unless a certain amount of energy is supplied to the surface. In the usual t[ype] of radio tube, the electron-emitting filament material is supplied w[ith] the heat energy of an electrical current sufficient to cause the des[ired] electron emission. Emission excited by heat energy is known as th[ermi]*ionic emission.*

Electron emission may be produced by electrons impinging u[pon] substances with sufficient velocity. For example the electrons emit[ted] by the hot filament of a radio tube may be accelerated toward the pla[te] by a positive voltage. If a great enough velocity is reached each electro[n] will have sufficient energy to release one or more electrons from the pla[te]. This is known as *secondary emission.*

The energy supplied by light is sufficient to cause emission from so[me] substances. This is the type of emission employed in photoelectric ce[lls] and is known as *photoelectric* emission.

Strong electric fields acting on gases or vapors may cause the g[as] particles to collide with sufficient energy to release electrons from the g[as]. This process is known as *ionization.* In this case both the electron a[nd] the remaining positively charged gas ion are mobile, so that the electro[n] moves toward the positive and the gas ion toward the negative electrode from which the field originates.

7. Thermionic Emission. The emission of electrons from me[tals] heated to a certain temperature is a characteristic property of the me[tal]. From consideration of thermodynamics and the kinetic theory of gases Richardson obtained an equation for thermionic emission.

$$I_s = A_1 T^{1/2} \epsilon^{-\frac{b_1}{T}} \qquad (1[6])$$

where I_s = emission current in amperes per square centimeter
A_1 = a constant for the emitting substance
T = absolute temperature in degrees Kelvin
ϵ = base of Napierian logarithms
b_1 = a constant depending upon the nature of the emitting surface

A similar equation giving equivalent results was derived by Dushman:

$$I_s = A_2 T^2 \epsilon^{-\frac{b_2}{T}} \qquad (17)$$

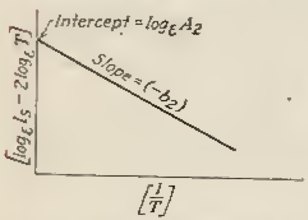


Fig. 1.—Determination of constants in emission equation.

where I_s = electron emission in amperes per square centimeter
T = absolute temperature of the emitter in degrees Kelvin (t + 273)
ϵ = base of Napierian logarithms (2.718)
b_2 = a constant for the material

The constants A_2 and b_2 of Eq. (17) can be determined for a given material in the following manner:

$$\log_\epsilon [I_s] = \log_\epsilon \left[A_2 T^2 \epsilon^{-\frac{b_2}{T}} \right]$$

$$[\log_\epsilon I_s - 2 \log_\epsilon T] = \left[\log_\epsilon A_2 - \frac{b_2}{T} \right]$$

Readings of the emission current from the substance at different temperatures are obtained. Values of $[\log I_s - 2 \log_\epsilon T]$ are plotted against $[1/T]$. The result should be a straight line. The intercept of this line with the vertical axis gives the value of $\log_\epsilon A_2$, the slope gives the value of $(-b_2)$.

Equations (16) and (17) are experimentally indistinguishable within the usual range of temperatures. When the constants are known for Eq. (16) the constants for Eq. (17) may be calculated from the following approximate relations.

$$b_2 = \left[b_1 - 1.5 \frac{T_1 + T_2}{2} \right] \qquad (18)$$

$$A_2 = [0.223 A_1 T^{-1.5}] \qquad (19)$$

For Non-homogeneous Emitters. For thoriated tungsten and oxide-coated emitters the emission constants depend to a considerable extent on the processing as well as on the materials. The curves below show typical data relative to pure metallic emitters.[1]

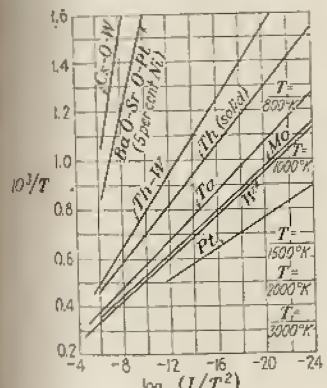


Fig. 1*a*.

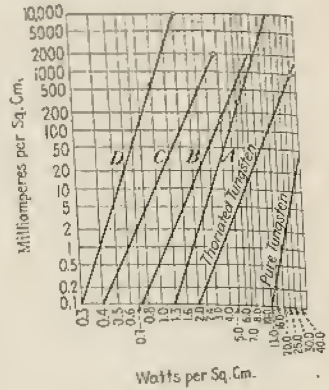


Fig. 1*b*.

Fig. 1*a*.—Emission of coated filament compared to that from pure metals. *A.I.E.E. reprint.*)

Fig. 1*b*.—Emission from coated filament *vs.* power input. *A* to *D* represent different examples from several sources. ("*A Science Series for Engineers,*" *A.I.E.E. reprint.*)

A filament coated with a mixture of the oxides of barium and strontium on a core of 95 per cent platinum and 5 per cent nickel has the following characteristics:

Electrical Resistivity of the Core.

$= 0.000022(1 + 0.00208t - 0.000,000,46t^2)$ ohm cm
t = temperature in degrees centigrade

Thermal Emissivity (Ratio to Black Body).

$= [0.4 + 0.00025T]$

where T = degrees Kelvin lies between 800° and 1200°K.

[1] Dushman, Saul, "A Science Series for Engineers," *A.I.E.E.*, reprint.

The electron emission in zero field is given by the equation

$$I_s = 0.01T^2\epsilon^{-\frac{11,600}{T}}$$

where T = degrees Kelvin
I_s = emission current in amperes per square centimeter

For an anode potential of 150 volts and a current limited by space charge to 0.010 amp. per square centimeter the average life is

$$= 0.000015\epsilon^{\frac{22,000}{T}} \text{ hr.}$$

The following values are those most probable when the anode potential equals 150 volts and the electric field is zero:

T	I_s	p_r	p_e	Life
900	20	2.3	0.02	730,000
950	45	3.0	0.045	170,000
1,000	90	3.7	0.09	55,000
1,050	170	4.6	0.17	20,000
1,100	310	5.6	0.31	7,400

T = temperature in degrees Kelvin
I_s = emission current in milliamperes per square centimeter
p_r = power thermally radiated in watts per square centimeter
p_e = power absorbed by electron emission in watts per square centimeter
Life = most probable average life in hours

8. Contact Potential. The rate of emission of electrons from different substances and the contact differences of potential are closely related.

The contact potential depends only upon the materials of the electrodes and their temperature, but not upon size, shape, or position of the electrodes.

For example, an electron in escaping from the inner to the outer surface of substance A will do work equal to W_A so that its potential is changed to V_A. Similarly the work for an electron to escape from the surface B is W_B and the potential change V_B. Hence in moving an electron from substance A across a space to substance B the work done will be

$$[W_A + (V_A - V_B)e - W_B]$$

This is the algebraic summation of the work done and would be equal to zero, except for the work done at the junction of the two substances in the return connection. This later potential difference is known as the *Peltier effect* and is negligible in comparison with the other effects.

$$W_A = \phi_A e$$
$$W_B = \phi_B e$$
$$(V_A - V_B)e = W_B - W_A = (\phi_B - \phi_A)e$$
$$(V_A - V_B) = (\phi_B - \phi_A)$$

$(V_A - V_B)$ is called the *contact potential difference* between the two substances, and by Eq. (24) it is equal to the difference in the work *function*, or electron affinity ϕ of the two substances.

9. Work Function. When a quantity of electricity q is moved through a potential difference V the work done equals qV. Work must be done when an electron is removed from a surface. If the work done per electron is W_1, the electron charge e, and the potential difference ϕ is required to supply an amount of energy equal to W_1, then,

$$W_1 = \phi e \tag{25}$$

$$\phi = \frac{W_1}{e} = \frac{k_0 b}{e} = (8.62 \times 10^{-5} b) \text{ volts} \tag{26}$$

ϕ is called the *electron affinity* of the substance and is equal to the work function (W_1/e). The smaller the quantity ϕ the easier it will be for an electron to escape from the cathode. A low value of ϕ indicates a large electron emission for a given temperature.

The following table gives the electron affinity or work function of several substances expressed in volts:

Substance	ϕ
Tungsten	4.52
Platinum	4.4
Tantalum	4.3
Molybdenum	4.3
Carbon	4.1
Silver	4.1
Copper	4.0
Bismuth	3.7
Tin	3.8
Iron	3.7
Zinc	3.4
Thorium	3.4
Aluminum	3.0
Magnesium	2.7
Nickel	2.8
Titanium	2.4
Lithium	2.35
Sodium	1.82
Mercury	4.4
Calcium	3.4

10. Filament Calculations. The dimensions of filaments designed to operate at a given voltage and temperature, and to furnish a certain total emission current are related to the physical properties of the material.

Suppose that the required total emission current is I_B ma. From the power-emission chart for the type of filament material being used, find I_s the emission current in milliamperes per square centimeter for a given power input p watts per square centimeter corresponding to good life performance, or to temperature T.

The total surface area of the required filament: $A = (I_B/I_s)$.
The total power input to the filament: $pA = E_f I_f = P_f$ watts.
At a voltage E_f the filament current $I_f = (pA/E_f)$.

Filament resistance at the operating temperature: $R_f = (E_f/I_f)$.

The resistance of a circular filament: $R = \left[\rho \frac{A}{2\pi^2 r^3}\right]$

where A = area of the filament surface
r = radius of the filament
ρ = specific resistance of the filament material. ρ must be known as a function of the temperature.

The resistance of a rectangular filament is given by

$$R = \left[\rho \frac{A}{2S_1S_2(S_1 + S_2)}\right]$$

where A = area of the filament surface
S_1 = thickness of the filament
S_2 = width of the filament
ρ = specific resistance of the filament material at temperature T

11. Filament-current Filament-radius Relation. For a given type of filament material operating at a specified temperature and filament voltage, the radius or filament cross section is uniquely related to the filament current

For a circular filament: $I_f = [(2p/\rho)^{1/2}\pi r^{3/2}]$

For a rectangular filament: $I_f = (2p/\rho)^{1/2} \cdot [S_1S_2(S_1 + S_2)]^{1/2}$.

For a square filament: $I_f = (2p/\rho)^{1/2} \cdot 2^{1/2} \cdot S_1^{3/2}$.

12. Filament-voltage Filament-dimensions for a Constant Temperature. For a given filament material to be operated at a given temperature, the filament voltage is related to the filament length and sectional dimensions as follows:

Circular filament: $E_f = (2p\rho)^{1/2}\dfrac{l}{r^{1/2}}$

Rectangular filament: $E_f = (2p\rho)^{1/2}\left(\dfrac{1}{S_1} + \dfrac{1}{S_2}\right)^{1/2} \cdot l$

13. Lead-loss Correction. The cooling effect of the leads connected to a filament decreases the emission from the parts near the junction. The voltage drop in these parts of the filament is also less.

Langmuir and Dushman give the following correction formulas for a V-shaped filament cooled by large leads. The decrease in voltage due to the cooling effect of the two end leads is

$$\Delta V = 0.00026(T - 400) \text{ volts}$$

T = degrees Kelvin of the central portion of the filament.

The correction for the effect on the electron emission is given in terms of the voltage of a length of uncooled filament which would give the same effect as the decrease caused by the cooling of the leads. The correction for the two leads is $\Delta V_H = 2(0.00017T\phi - 0.05)$ volts. ϕ is a number which depends upon the temperature coefficient of the quantity H, which may represent any property of the metal, such as candlepower, electron emissivity, etc. For the case of electron emission the exponent of the temperature coefficient is $N = \left(2 + \dfrac{b_0}{T}\right)$

Dushman's coefficient for the material b_0 and the temperature T in degrees Kelvin being known, N is calculated.

N	0.5	1.0	2.0	2.5	5.0	10.	20.	30.	50.
ϕ	0.48	0.85	1.23	1.44	1.72	2.10	2.47	2.69	2.95

N is related to ϕ as shown by the data above which may be plotted as a curve. Knowing ϕ the correction ΔV_H is determined.

The electron emission per unit area after taking into account the lead-loss correction is

$$I = \left(\frac{i}{S}f\right)$$

where i = observed total emission from any given filament
S = total filament area

The correction factor f is given by

$$f = \left[\frac{V + \Delta V}{V + \Delta V - \Delta V_H}\right]$$

Dushman gives curves of ΔV and ΔV_H plotted against temperature for different values of b_0.

$V + \Delta V$ corresponds to the corrected voltage drop along the filament.

14. Effect of Space Charge. The equations of Richardson and Dushman for thermionic emission give the total electron current, with

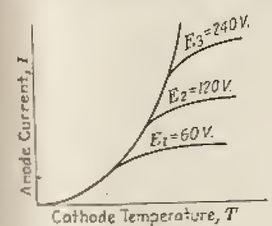


FIG. 2.—Space-charge effect in limiting emission.

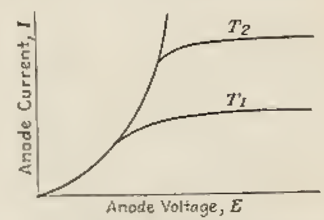


FIG. 3.—Saturation at constant temperatures.

zero field strength at the surface of the cathode. If the electrons are allowed to accumulate just outside the surface they form a negative cloud. If the electrons are drawn to a positive electrode both the negative cloud and to a less degree the cathode surface fields are changed.

Langmuir found that if the voltage applied to the anode was not sufficiently high a temperature increase of the cathode did not increase the current indefinitely. This effect is shown in Fig. 2. It is due to the repelling effect of the negative cloud of electrons surrounding the cathode and is known as the *space-charge effect*, or volume density of electrification. Figure 3 shows this effect with constant-cathode temperatures and variable-anode voltage.

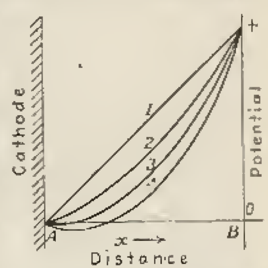


FIG. 4.—Distribution of potential in cathode-plate space.

The theory of these effects is as follows: The distribution of the potential between two large parallel plates is directly proportional to the distance starting from the low and increasing to the high potential plate. If plate A emits low-velocity electrons (assumed zero) spontaneously, and if plate B is positive with respect to A, electrons will be drawn over to B. Starting with a low temperature T, the distribution of potential between A and B will be uniform as shown by the straight line 1 in Fig. 4. Increasing the temperature of A will cause an electron current of I amp. per square centimeter to flow to B. Laplace's equation connecting the potential distribution with the volume density of electrification ρ is

$$\Delta V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi\rho \qquad (27)$$

For large parallel planes Eq. (27) may be simplified to

$$\frac{d^2 V}{dx^2} = -4\pi\rho \qquad (28)$$

If ρ is constant and negative, the potential distribution will be a parabolic curve as shown by curve 2 in Fig. 4. A further increase in the temperature of A will cause the parabola to take the form of curve 3 having a horizontal tangent at A. In this case the potential gradient at the cathode is zero ($dV/dx = 0$), and a further increase of temperature will not increase the electron current to B. This accounts for the effect shown in Fig. 2.

In the above discussion the electrons were assumed to be emitted with no initial velocity. Usually small initial velocities exist, so that a slightly negative gradient is necessary at A in order to prevent an increase in current. Curve 4 of Fig. 4 shows the effect of the initial velocities of emission on the potential distribution at the temperature for which a further increase in temperature will not increase the anode current.

15. Schottky Effect. Richardson's and Dushman's equations for the thermionic emission from a substance at a given temperature assume that the electric field strength is zero at the cathode. In actual practice a definite potential is used. This effect of the potential gradient at the cathode on the observed emission current is called the *Schottky effect*.

Dushman gives the correction for the Schottky effect as follows:

I_0 = electron emission in zero field
I_v = observed emission at an anode voltage V

Then

$$I_v = I_0 \epsilon^{\frac{4.39\sqrt{kV}}{T}}$$

where k = a constant whose value depends upon the relative geometrical arrangement of anode and cathode
T = temperature in degrees Kelvin
ϵ = base of Napierian logarithms.

16. Electron Current between Parallel Plates. When the cathode is a large flat surface A and the plate, or anode, B is a parallel surface, the plate current per square centimeter of surface not too near the edges of the plates is given by the equation

$$i = 2.34 \times 10^{-6} \frac{V^{3/2}}{x^2} \qquad (29)$$

where i = maximum current density in amperes per square centimeter
x = distance between plates in centimeters
V = potential difference between A and B in volts

This equation assumes that the initial velocities of the electrons leaving A are zero. If the potential of B is large relative to one or two volts, the initial velocities of the electrons can be neglected.

Equation (29) assumes that the anode potential is positive with respect to A so that some current is flowing but that the anode potential is below

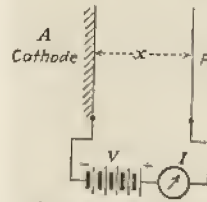


FIG. 5.—Electron current between parallel plates.

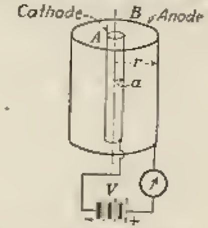


FIG. 6.—Cylindrical structure.

the value necessary to give the full current emitted at A. When the anode potential is great enough to draw over all of the electrons emitted at A, the current (saturation current) I_s is given by the Richardson-Dushman equation.

17. Electron Current between Concentric Cylinders. Given two concentric cylinders A and B (Fig. 6) having radii of a and r cm and of infinite length. Langmuir's equation for the electron current to the plate B is given by the relation

$$i = 14.7 \times 10^{-6} \frac{V^{3/2}}{r\beta^2}$$

where i = current in amperes per centimeter length
V = potential between A and B in volts
r = radius of the anode in centimeters
a = radius of the cathode in centimeters
β = a factor which varies with the ratio of (r/a)

r/a	β^2	r/a	β^2
1.00	0.000	20	1.072
2.00	0.279	50	1.094
3.00	0.517	100	1.078
4.00	0.667	200	1.056
5.00	0.767	500	1.031
10.00	0.978	1,000	1.017
		∞	1.000

When the inner cylinder is a small wire of less than one-tenth the diameter of the plate, the error is small if β is neglected, and the approximate equation is

$$i = \left[14.7 \times 10^{-6} \frac{V^{3/2}}{r}\right]$$

18. Electron Current with Any Shape Electrodes. Langmuir has demonstrated that under the assumption on which the above equations were derived the current will vary as the three-halves power of the poten-tial difference V regardless of the shape of the electrodes. The deriva-tion of the equations neglects the initial velocities of the electrons an the potential gradient at the cathode.

19. Two-electrode Vacuum Tubes. The three-halves power equa-tion for the plate current of a two-electrode tube is quite accurate whe the voltage between cathode and plate is large with respect to the effec of (1) initial velocities of emission; (2) voltage drop in the filament o cathode; (3) contact potential between cathode and plate and th emission of electrons from the cathode is large and the plate voltag well below the value for saturation current. The electrodes are assum to be in good vacuum, so that the effects of gas are negligible.

In the case of thoriated-tungsten or oxide-coated filaments only fraction of the total cathode surface is active so that the saturatic current may be reached at a plate voltage below the theoretical.

The current is calculated from the formula

$$i = kV^{3/2}$$

where k is the space-charge constant of the tube for a given type of tu structure and depends only upon the geometrical configuration witho regard to the dimensions of the tube. The value of k for infinite para plates is $\left(2.34 \times 10^{-6}\frac{A}{x^2}\right)$

where A = the area of the plate in square centimeters
x = the distance from the cathode plate to the anode plate in cen meters

For concentric cylinders, $k = \left(14.7 \times 10^{-6}\frac{l}{r\beta^2}\right)$

l = length of the cylinders
r = radius of the outer cylinder or anode
β = a function of (r/a) (see table on page 241)

20. Effect of Initial Velocities—Parallel Plates. If the effect of the initi velocity of the electrons is included and they have a Maxwellian distributio

$$i = 2.34\frac{A}{(x_a - x_m)^2}(V_a - V_m)^{1.5}\left(1 + 0.0247\sqrt{\frac{T}{V_a}}\right)$$

where i = total plate current in amperes
A = area of one surface of the anode in square centimeters
T = temperature of the cathode in degrees Kelvin
V_a = potential of the anode above that of the cathode volts
V_m = minimum potential of the space between cathode and anode w respect to the cathode
x_a = distance from cathode to anode in centimeters
x_m = distance from cathode to V_m in centimeters

21. Effect of Magnetic Field. Initial velocities = 0. For coaxial cylind

$$i = kV_a^{1.5}, \text{ if } V_a > V'$$
$$i = 0, \quad \text{if } V_a < V'$$
$$k = \text{same as above}$$

$$V' = \left[0.0221H^2r_o^2 + 0.0188I^2\left(\log_{10}\frac{r_o}{r_i}\right)^2\right]$$

H = strength of magnetic field externally applied parallel to axis cylindrical electrodes
I = current flowing through the inner cylindrical electrode parallel to axis
r_o = radius of the outer electrode
r_i = radius of the inner electrode

22. Characteristics of Typical Commercial Diodes.

Type	i_f	E_f	E_m	i_m	P_n	k
ThW	1.25	7.5	550	0.065	0.0075	1.2
ThW	3.25	10	1,500	0.20	0.050	1.7
ThW	3.85	11	2,500	0.25	0.250	1.1
PW	14.7	11	16,000	0.166	1.00	0.5
PW	24.5	22	17,500	0.833	5.00	1.0
PW	52	22	18,000	3.0	20.00	1.1
PW	10	10	20,000	0.10	....	0.10
PW	10	10	85,000	0.10	....	0.11
PW	32	9	75,000	0.25	....	0.25
PW	10	10	150,000	0.100	....	0.11
PW	32	12.5	150,000	0.25	....	0.11

i_f, E_f = filament current, voltage (amperes and volts)
E_m = maximum effective a-c input voltage (volts)
i_m = maximum rectified tube current (amperes)
P_n = nominal power rating (kilowatts)
ThW, PW = thoriated tungsten, and pure tungsten, filament
k = 0.0001 amp. per volt$^{1.5}$

23. Effect of the Grid. When a wire mesh or similar electrode having openings through which electrons may pass is placed between the cathode and the plate, it exerts a large controlling effect on the flow of electrons to the plate. The meshlike electrode between cathode and plate is termed a *grid*.

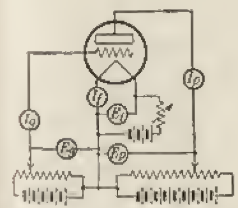


FIG. 7.—Circuit for measuring static characteristics.

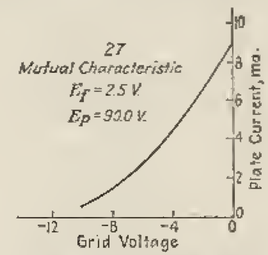


FIG. 8.—Typical grid-voltage plate-current characteristic.

When the grid is connected to a source of voltage, the electrons are attracted if the grid is positive with respect to the cathode and repelled if it is negative. The close proximity of the grid to the space charge surrounding the cathode increases its effectiveness in controlling the electron flow.

In most useful applications the tubes are operated with sufficient electron emission and with plate and grid voltages low enough so that the space charge surrounding the cathode is ample to permit large momentary increases in the electron flow to the plate.

The effect of a large positive plate voltage in drawing the electrons to the plate can be reduced by a relatively small negative voltage applied to the grid. The electrons being negative will avoid the negative grid so that no current will flow in the grid circuit. If the negative grid voltage is not too large with respect to the plate voltage, electrons will be drawn through the openings in the grid mesh to the positive plate.

The resulting plate current is controlled by the grid, although no current flows in the grid circuit. Zero power in the grid circuit can thus control a considerable amount of power in the plate circuit. Voltage variations of the grid produce corresponding variations of the plate current. The extent to which the plate-current variations are faithful reproductions of the grid-voltage variations depends upon the steady polarizing voltages (A, B, and C voltages) applied to the tube and the range of the voltage variations.

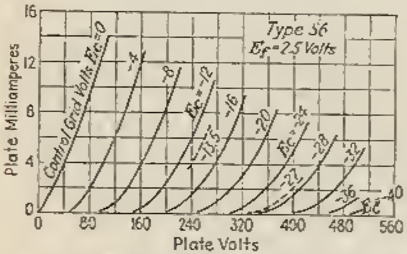


FIG. 9.—Plate characteristics of typical triode tubes.

CHARACTERISTICS OF THE THREE-ELECTRODE TUBE

24. Static Characteristics. The effects of various d-c voltages applied to the electrodes of a tube are shown by curves called the *static characteristics.*

The *mutual* or *transfer* characteristic of the tube shows the effect of the grid voltage upon the plate current. The term mutual or transfer indicates that the voltage in one circuit controls the current in another circuit.

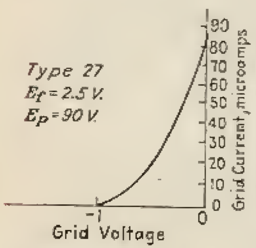


FIG. 10.—Grid-current grid-voltage characteristic.

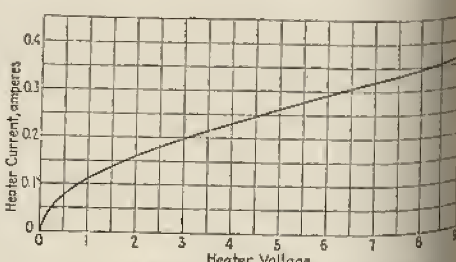


FIG. 11.—Filament characteristic of 6C6.

The *plate characteristic* represents the relation between plate current and plate voltage.

The *grid characteristic* shows the grid current-grid voltage relation. Electron flow to the grid starts in the region of zero grid voltage. The exact point at which grid current starts is determined by the initial velocities of emission and the contact potential of the grid to cathode. The net effect is equivalent to a small positive or negative bias usually not greater than one volt.

In Fig. 10 the inherent bias in the tube is nearly 0.9 volt positive, so that the grid must be biased negative by 0.9 volt to secure an effective zero grid voltage.

The filament-voltage filament-current curve obtained with plate and grid terminals disconnected is termed the *filament characteristic.* The characteristic refers to the heater filament when the tube is of the indirectly heated cathode type.

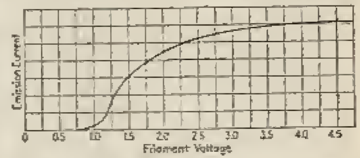


FIG. 12.—Filament characteristics of 1C6.

25. Normal Emisson and Emission Characteristic. The *normal emission current* is ordinarily obtained as a single reading at rated filament voltage. The circuit arrangement for this test is shown in Fig. 13. A definite voltage (50 volts is commonly used) is applied between the cathode and all other electrodes as the anode. A switch is arranged in the circuit so that the voltage is applied only long enough to obtain the emission-current reading. This test should not be made at rated filament voltages on large power tubes where the heating would be excessive; or on certain low-filament-current types; or on certain oxide-coated filament tubes. An emission check on all these types can be made by observing the filament voltage required to give a certain small value of emission current (values of 3 or 5 ma are generally used).

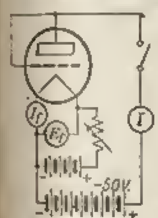


FIG. 13.—Measurement of emission characteristic.

The normal emission test, even though applied only momentarily, usually causes some liberation of gas and heating of the electrodes. Hence it is desirable to complete other tests before this test is made or to allow sufficient time after this test, operating with normal voltages, to clear up gas and to return the temperature to normal.

When the effect of filament voltage on normal emission current is of interest, readings, obtained as above but for different filament voltages, are plotted as a curve.

The *emission characteristic* shows the true (total) emission current for a range of cathode heating power.

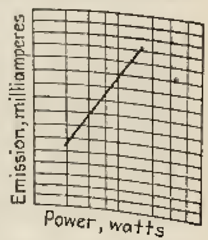


FIG. 14.—Emission curve.

To avoid the effects of space charge, heating of grids and anode, liberation of gas, and such extraneous effects, the readings are taken only with low cathode-heating power, and the emission for normal heating power is obtained by extrapolation. A usual procedure is to read the cathode-heating power for emission currents of 0.1, 0.2, 0.5, 1.0, 2.0, and 5.0 ma with 50 volts positive on the common electrode connection with respect to the cathode. The data are plotted on a special coordinate paper devised by C. J. Davisson. If the emission follows Richardson's temperature equation and the power is radiated according to the Stefan-Boltzmann law of radiation, the curve appears as a straight line. The extension of this straight line shows the emission current for normal or other values of cathode-heating power.

If the curve of the experimental data plotted in Davisson coordinates is not a straight line, this may be caused by one or more of the following conditions:

1. Departure from the Stefan-Boltzmann cooling (bends downward).
2. Anode voltage too low to draw off all the electrons (bends downward).
3. Effect of cooling due to heat of evaporation of cooling of electrons (bends downward). The cooling due to electron evaporation amounts to approximately ϕI_s watts, where I_s represents the emission current in amperes and ϕ represents the work function of the cathode in volts. This effect may be considerable in transmitting tubes where the currents are high, and in tungsten-filament tubes where the work function is large.
4. Poor vacuum (gas ionization effects) (bends upward).
5. Heating of the anode by the emission current (bends upward).
6. Progressive change in activity of the cathode.

A method for reading emission currents which is applicable in general consists in the use of a commutator for applying the voltage recurrently for only a small time interval. By means of an oscillograph the emission current is read as the peak current during the interval the voltage is applied. By this method the heating effects can be kept low.

CALCULATION OF THE SPACE CURRENT OF THE THREE-ELECTRODE TUBE

26. The space current I of a three-electrode tube is equal to the sum of the plate current I_p and the grid current I_g; $I = (I_p + I_g)$. The three-electrode tube is calculated as an equivalent diode $I = k(E_p + \mu E_g)^{3/2}$. The grid voltage E_g is equivalent to a plate voltage μE_g. μ is the amplification factor of the tube.

27. Plane-parallel Elements. For a structure with plane-parallel elements with the filament symmetrically placed between grids and plates:

$$k = 2.34 \times 10^{-6} \times \frac{A}{(\alpha + \beta)^{1/2}[\alpha + \beta(\mu + 1)]^{3/2}}$$

$$\mu = \frac{2\pi a n}{\log_e \dfrac{1}{2\pi r n}}$$

$$I = 2.34 \times 10^{-6} \frac{A}{(\alpha + \beta)^{1/2}} \left[\frac{E_p + \mu E_g}{\alpha + \beta(\mu + 1)}\right]^{3/2}$$

where I = total space current in amperes
α = distance from plate to grid in centimeters
β = distance from grid to filament in centimeters
n = number of grid wires per centimeter length of the structure
r = radius of the grid wires
A = effective plate area.

28. Concentric Elements. For a structure with a cylindrical anode and grid and a coaxial strand of filament,

$$k = 14.7 \times 10^{-6} \frac{L R_p^{1/2}}{[(R_p - R_g) + R_g(\mu + 1)]^{3/2}}$$

$$\mu = \frac{2\pi n R_g^2 \left(\dfrac{1}{R_g} - \dfrac{1}{R_p}\right)}{\log_e \dfrac{1}{2\pi r n}}$$

$$I = 14.7 \times 10^{-6} \frac{L}{R_p}\left[\frac{(R_p - R_f)(E_p + \mu E_g)}{(R_p - R_f) + (R_g - R_f)(\mu + 1)}\right]$$

If R_f is very much smaller than R_p and R_g, the equation can be written approximately

$$I = 14.7 \times 10^{-6} L R_p^{1/2}\left[\frac{E_p + \mu E_g}{(R_p - R_g) + R_g(\mu + 1)}\right]^{3/2}$$

where L = length of the structure in centimeters
R_f = radius of the filament in centimeters
R_p = radius of the plate in centimeters
R_g = radius of the grid in centimeters.

The above relations are useful in the design of the structures. The k should be determined for the type of tube structure. The μ and the current-voltage characteristics remain the same if all dimensions are changed proportionately. The plate current equals the space current when the grid current is zero.

29. Amplification Factor. The amplification factor is a measure of the effectiveness of the grid voltage relative to that of the plate voltage upon the plate current. It is the ratio of the change in plate voltage to a change in grid voltage in the opposite polarity, under the condition that the plate current remains unchanged. As most precisely used, the term refers to infinitesimal changes as indicated by the defining equation:

$$\mu = -\frac{\partial e_p}{\partial e_g}; \qquad i_p = \text{constant}$$

The amplification factor is indicated by the horizontal spacing of the plate characteristic or mutual characteristic curves of the tube. Since horizontal lines represent constant plate current, the plate voltage spacing divided by the grid-voltage spacing of the curve is the amplification factor. The amplification factor of three-electrode tubes is nearly constant for a constant plate current. In the region near zero plate current or near the full emission current of the filament, the amplification factor changes greatly with voltage.

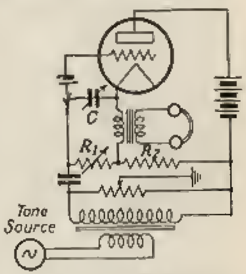


FIG. 15.—Measurement of amplification factor.

30. Measurement of Amplification Factor. An a-c bridge circuit shown schematically in Fig. 15 may be used to measure μ. The resistance R_1 is adjusted for zero sound in the phones. The amplification factor is given by

$$\mu = \frac{R_2}{R_1}$$

Owing to tube capacities or other reactances in the circuit it is usually necessary to provide a means for adjusting the phase of the grid and plate a-c voltages for complete balancing out of the sound in the phones. This phase balance is secured with condenser C in Fig. 15. The d-c voltage drop in R_2 should be allowed for when setting the plate voltage. The adjustable ground connection is convenient in eliminating the unbalancing effects of capacity to ground. The a-c tone voltage should be as small as practical. The phones can be preceded by a suitable amplifier.

CALCULATION OF THE AMPLIFICATION FACTOR

31. Plane-parallel Electrodes. When the diameter of the grid wires is large compared to their spacing the formula derived by Vodges and Elder is most accurate. Figure 16 shows a cross section of the electrodes. The amplification factor is

$$\mu = \left[\frac{2\pi ns - \log_\epsilon \frac{1}{2}(\epsilon^{2\pi nr} + \epsilon^{-2\pi nr})}{\log_\epsilon (\epsilon^{2\pi nr} + \epsilon^{-2\pi nr}) - \log_\epsilon (\epsilon^{2\pi nr} - \epsilon^{-2\pi nr})}\right]$$

where r = radius of the grid wire in centimeters

n = number of grid wires per centimeter length of structure

s = distance from plate to grid in centimeters.

When the diameter of the grid wires is small compared to their spacing, the equation above simplifies to

$$\mu = \frac{2\pi ns}{\log_\epsilon \left(\frac{1}{2\pi nr}\right)}$$

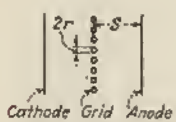


FIG. 16.—Tube with plane-parallel electrodes.

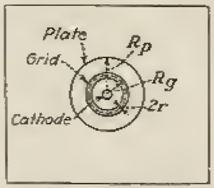


FIG. 17.—Tube with concentrically arranged electrodes.

32. Concentric Cylindrical Electrodes. The amplification factor of the cylindrical structure shown in Fig. 17 is given by

$$\mu = \frac{2\pi nR_g \log_\epsilon (R_p/R_g) - \log_\epsilon \frac{1}{2}(\epsilon^{2\pi nr} + \epsilon^{-2\pi nr})}{\log_\epsilon (\epsilon^{2\pi nr} + \epsilon^{-2\pi nr}) - \log_\epsilon (\epsilon^{2\pi nr} - \epsilon^{-2\pi nr})}$$

where R_p = radius of the anode in centimeters

R_g = radius of the grid in centimeters

r = radius of the grid wires in centimeters

n = number of grid wires (turns) per centimeter length of structure

When the diameter of the grid wires is small compared with their spacing the equation simplifies to

$$\mu = \frac{2\pi nR_g \log_\epsilon (R_p/R_g)}{\log_\epsilon \left(\frac{1}{2\pi nr}\right)}$$

33. Plate Resistance and Plate Conductance. The plate resistance r_p is defined by the equation

$$r_p = \frac{1}{S_p} = \frac{\partial e_p}{\partial i_p}$$

It is the reciprocal of the plate conductance S_p.

The plate conductance is the ratio of the change in plate current to the change in plate voltage producing it, all other electrode voltages being maintained constant. As most precisely used, the term refers to infinitesimal changes as indicated by the defining equation

$$S_p = \frac{\partial i_p}{\partial e_p}$$

The plate conductance is given by the slope of the plate-characteristic curves of the tube. When readings are taken on the characteristic curves, the current and voltage increments should be made as small as convenient. The plate resistance is the reciprocal slope of the plate-characteristic curve. The numerical value of the plate resistance changes with the applied d-c operating voltages.

34. Measurement of the Plate Resistance. The plate resistance or plate conductance can be measured directly with the aid of a bridge type of circuit. When the bridge in Fig. 18 is balanced for minimum sound in the phones, the plate resistance of the tube is

$$r_p = R_2R_3/R_1$$

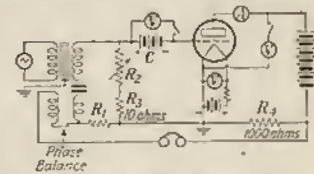


FIG. 18.—Measurement of plate resistance.

The alternating voltage (tone) applied to the bridge should be as small as practical. The use of an amplifier preceding the phones increases the sensitivity and accuracy of these measurements. The effects of small capacities are sometimes troublesome in circuits of this type. The electrode capacity of the tube causes some phase shift resulting in a poor balance. The phase balance variometer balances the small out-of-phase component permitting a closer adjustment to the null point. The capacity to ground can be balanced by suitable shielding or by means of a Wagner earth connection.

35. Calculation of the Plate Resistance. The plate resistance of a tube depends upon the operating voltages as well as the structural parameters. Within certain limits it is inversely proportional to the area of the anode and also to the area of the cathode. Decreasing the distance between filament and plate decreases the plate resistance. Since it is desirable to make (μ/r_p) large, the grid to plate distance controlling μ should not be decreased too much. This requires that the grid be placed near the filament to lower the plate resistance. When the grid is too near to the filament, it will be heated. Small amounts of grid emission current resulting from too high grid temperature have an objectionable effect on the operation of the tube.

The plate resistance of a tube may be calculated from the plate-current plate-voltage relation. For a structure with plane-parallel elements in which the filament is symmetrically placed between grids and plates, the plate resistance is

$$r_p = \frac{(\alpha + \beta)^{\frac{1}{2}}[\alpha + \beta(\mu + 1)]^{\frac{3}{2}}}{A(E_p + \mu E_g)^{\frac{1}{2}}} \times 10^6$$

where r_p = plate resistance in ohms
α = distance from plate to grid in centimeters
β = distance from grid to filament in centimeters
μ = amplification factor
E_p = plate voltage
E_g = grid voltage
A = a constant depending on the cathode area, or anode area, and type of structure. For typical filament-type tubes $A = 1.8L$ where L is the length of the filament in centimeters.

The grid voltage E_g is conveniently made zero and the plate voltage taken equal to the value giving normal plate current.

36. Transconductance. The grid-plate transconductance is defined by the relation

$$S_m = S_{gp} = \frac{\partial i_p}{\partial e_g}$$

It is the ratio of the change in plate current to the change in grid voltage, under the condition that all other voltages remain constant. It is also equal to the ratio of the amplification factor μ to the plate resistance r_p of the tube:

$$S_m = \frac{\mu}{r_p}$$

The transconductance determines the plate-current change per volt applied to the grid. It is evident that this is the most important characteristic of a tube. It is a figure of merit of the tube and enters into the calculations of the performance of the tube. It is a direct measure of the amplifying properties of the tube operating into a load impedance which is small with respect to the plate resistance. With high impedance loads the amplification factor and plate resistance are considered separately in determining the tube performance. The transconductance may be determined graphically from the slope of the mutual characteristic curve of the tube. Direct measurements are usually most convenient when many readings are required.

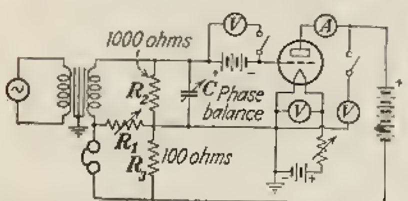


FIG. 19.—Measurement of transconductance.

37. Measurement of Transconductance. The transconductance can be measured directly in the circuit shown in Fig. 19. The resistance R_2 and the phase balance C are adjusted until the sound in the phones is balanced out. The transconductance is given by

$$S_m = \frac{R_1}{R_2R_3}\left(1 + \frac{R_3}{r_p}\right) = \frac{R_1}{R_2R_3} \text{ (approx.)}$$

38. Calculation of the Transconductance. The transconductance S_m is equal to the ratio of the amplification factor μ to the plate resistance r_p. Each of these factors can be calculated with a fair degree of accuracy

for certain types of structures. The amplification factor depends almost entirely upon the structure of the grid and the grid-plate distance. The plate resistance depends upon the amplification factor, the surface areas of the cathode and anode, the grid-filament distance, and the applied d-c operating voltages. The transconductance depends upon all these factors.

39. Grid-current Coefficients. When the grid is not biased with sufficient negative voltage and the tube operation extends into the positive range of grid voltage, an electron current will flow to the grid. Under these conditions the current in the grid circuit may change the effective grid voltage. When it is desirable to include these effects in determining the performance of the tube, the coefficients relative to the grid current are useful.

The *grid conductance* S_g, or its reciprocal the grid resistance r_g, is defined by the equation

$$S_{gg} = S_g = \frac{\partial i_g}{\partial e_g}$$

$$r_g = \frac{1}{S_g} = \frac{\partial e_g}{\partial i_g}$$

The grid conductance S_g is the ratio of the change in the grid current to the change in grid voltage producing it, other electrode potentials being maintained constant. As most precisely used, the term refers to infinitesimal changes, as indicated by the defining equation.

The coefficient showing the relative effectiveness of grid and plate voltages on the grid current has been variously termed *reflex factor*, *inverse amplification factor*, and *inverse factor*. Recent I.R.E. standards term this coefficient the *plate-grid mu factor*. It is the ratio of the change in grid voltage to the change in plate voltage required to maintain a constant value of grid current. As most precisely used, the term refers to infinitesimal changes as indicated by the defining equation

$$\mu_{ggp} = \mu_n = -\frac{\partial e_g}{\partial e_p};\ i_g = \text{constant}$$

The coefficient showing the effect of plate voltage on the grid current has been termed *inverse mutual conductance*, or the *plate-grid transconductance* (note that this is not the grid-plate transconductance. The difference in these terms can be easily remembered, since the words grid and plate appear in the same order as the direction of action in the tube). It is the ratio of the change in grid current to the change in plate voltage producing it, all other electrode voltages being maintained constant. As most precisely used, the term refers to infinitesimal changes, as indicated by the defining equation

$$S_{gp} = S_n = \frac{\partial i_g}{\partial e_p}$$

Grid-current coefficients of the tube may be determined graphically from the static characteristic curves or measured directly in bridge circuits similar to those employed for plate-current coefficients.

40. Higher-order Coefficients. Tube coefficients in most common use are the amplification factor, plate resistance or conductance, and transconductance. These are the first-order plate-current coefficients of triode. They determine the amplifying properties of the tube and enter into nearly all applications of the tube.

When the tube is operated so that detection, modulation, distortion, cross modulation, frequency conversion, and such effects are of importance, it is necessary to use second-order, third-order, and higher order coefficients in addition to the first-order coefficients to determine the performance of the tube. For example, in the case of plate-circuit detection the tube coefficient determining this effect is the second derivative of the plate current with respect to the grid voltage. The first derivative, or first-order coefficient, is the transconductance which is

$$\frac{\partial i_p}{\partial e_g} = S_m$$

The second derivative, or second-order coefficient, is

$$\frac{\partial^2 i_p}{\partial e_g^2} = \frac{\partial S_m}{\partial e_g}$$

The d-c plate-current change with signal voltage and second-harmonic distortion are also determined by the second-order coefficient.

Cross modulation and modulation distortion in the r-f stages of a receiver are determined by the third-order coefficient

$$\frac{\partial^3 i_p}{\partial e_g^3} = \frac{\partial^2 S_m}{\partial e_g^2}$$

The third-harmonic distortion in a tube is also determined by the third-order coefficient. The fifth-harmonic distortion would be determined by the fifth-order coefficient,

$$\left(\frac{\partial^5 i_p}{\partial e_g^5}\right)$$

Higher order coefficients are usually obtained graphically from the current-voltage characteristics of the tube. When the analytical expression for the current is known, the coefficients may be obtained by differentiation. The measurement of an effect depending principally on one coefficient may be used as a measure of the coefficient.

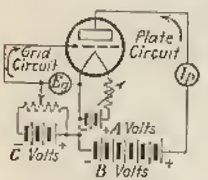


FIG. 20.—Triode circuit.

41. Mechanism of the Three-electrode Amplifier. Figure 20 represents a triode connected to a suitable source of A, B, and C voltage. A meter I_p is connected in the plate circuit for reading plate current. A potentiometer is connected across the C voltage. The grid voltage E_g will be changed as the slider is changed on the potentiometer. If the slide moves toward the positive, the plate current increases; if toward the negative, the plate current decreases. The plate currents corresponding to different grid voltages are plotted as in curve 1 in Fig. 21. This is a mutual characteristic curve of the tube.

Suppose that the slide is varied in some definite manner. For example, start to count time from zero on curve 2 in Fig. 21. With the slider

initially at 5 volts the plate current is 3 ma. Move the slider steadily in the negative direction, until say, in 3 sec. the grid voltage is 9 volts. The plate current will be 0.5 ma. Now start the slider in the positive direction, moving at the same steady rate. At the end of 6 sec. the slider has returned to its original position. If you continue the motion of the slider in the positive direction, at the end of 9 sec. the grid voltage is -1.0 volt, and the plate current is 6.5 ma. If the slider is started in the negative direction at the same rate, the grid voltage will be -5 volts at the end of 12 sec., thus completing the cycle.

Curve 3 shows the plate-current change corresponding to the grid-voltage change with time. If the slider is connected to a mechanism arranged to continue this motion, the plate current would contain an a.c. of 1 cycle in 12 sec. or 5 cycles per min. The wave form of the a.c. will be as shown in curve 3. It is superimposed upon the d-c plate current.

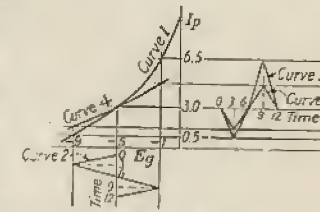


FIG. 21.—Mechanism of amplification.

The positive and negative peaks of the plate current as measured from the initial 3-ma point are not equal, although the grid-voltage peaks are equal. In this case the plate current is not a faithful reproduction of the input voltage.

If a resistance is connected in the plate circuit, the effective plate voltage is reduced as the plate current increases. The plate current at E_g equals -5 volts can be brought to the initial 3-ma point by a suitable increase in the B voltage to compensate for the voltage lost in the resistance. Starting with the same initial 3-ma point, the resulting characteristic with a resistance load is shown by the curve 4 in Fig. 21. The same alternating grid-voltage curve 2 produces the plate-current curve 5. The positive and negative plate-current peaks of curve 5 as measured from the initial point are almost identical. The distortion has been eliminated, and the voltage developed across the resistance can be used to operate a succeeding stage of amplification or other device.

The potentiometer and slider of Fig. 20 can be replaced with a fixed grid-bias voltage and an a-c voltage. The tube will operate as described above except that a-c cycles usually occur so rapidly that the plate-current (d.c.) meter cannot follow them. A meter showing the effective value (r-m-s) of the a.c. can be used to measure the current. The a.c. can be heard when connected to a loud-speaker, if it is within the audible range of frequencies. The wave form of the a.c. can be seen when connected to an oscillograph.

42. General-purpose Triodes. General-purpose triodes are used for detection, for voltage amplification, and in general in circuits where a low-power triode tube is needed.

Some of the available types of cathodes are as follows: a filament type with low current suitable for operation with dry-cell batteries; a filament type with higher current used with storage batteries (filament types of tubes requiring relatively high current and operated with a-c supply are used in the power output stage); a heater-cathode type operating on 2.5 volts a-c supply; a heater-cathode type operating on 6.3 volts for direct connection to the storage battery of an automobile, for use in series-connected d-c line or universal a-c, d-c circuits, and for use with 6.3 volts a-c supply.

A medium amplification factor (6 to about 15 or 20) is characteristic of the general-purpose type. The high-amplification-factor tubes are especially suitable for use in resistance-coupled a-f circuits. The plate characteristics are relatively low plate current and medium or high plate resistance. The grid-plate transconductance is usually not so high as obtained with power amplifier triodes.

The medium-plate-resistance types are suitable for use in transformer-coupled a-f amplifier circuits, in grid-leak detector circuits, and in general in circuits where a medium-plate-resistance, medium-amplification-factor triode tube is suitable.

The high-amplification-factor type having high plate resistance can be used with resistance-coupled (or impedance-coupled) circuits for a-f voltage amplification. This type is suitable for use as a grid-biased detector with resistance-coupled output. The medium-amplification-factor types also can be used as grid-biased detectors when a resistance-coupled or high-impedance output circuit is used.

Operating plate voltages below 250 volts are usual unless exceptionally large amplitude output voltages are required. The operating plate voltage must be large enough to accommodate this maximum output voltage. The grid-bias voltage and the plate load impedance are usually chosen to give low distortion and maximum output.

43. Power Amplifier Triodes. Power amplifier triodes are used when more power is needed than can be obtained from the ordinary amplifier triodes or where lower plate resistance or higher transconductance is desired. For the power output stage in radio receivers, for operating relays, lighting small signal lamps, and in general for delivering voltage and power in low-resistance loads these types are used. The low plate resistance is an advantage when a flat amplification characteristic over a wide range of frequencies is desired. In some instances, for example, in operating a low-resistance relay where a large plate-current change per volt on the grid is desired, a power triode with high transconductance is used. When adequate signal voltage is available and an insensitive relay is used or when positive action is of first importance, a tube with maximum plate current would be more important than a high transconductance.

For operating loud-speakers, the transformer primary carries the d-c plate current plus the alternating current due to the signal. In this case a low d-c plate current causes less tendency to saturate the core when a single tube is used and less loss in the winding resistance when a push-pull stage is used. For loud-speaker and other applications where appreciable power with low distortion is desired, a power amplifier triode is used.

An important characteristic of the power amplifier triode is that the distortion decreases to a low value and the power output decreases only at a slow rate as the load resistance increases beyond a value equal to the plate resistance of the tube. For low distortion (about 5 per cent second harmonic) it is usual to operate with a load resistance equal to twice the plate resistance of the tube.

Power amplifier triodes are characterized by high plate current, low plate resistance, low amplification factor, high transconductance, and moderate to high power output depending on the maximum plate voltage and plate current or the power dissipation permissible in the tube.

Typical power amplifier triode tubes for radio receivers and similar low-power usage have a range of plate current for the various types from 12.3 to 60 ma; plate resistances from 800 to 5,000 ohms; amplification factor from 3.0 to 8.0; transconductance from 1,050 to 5,250 micromhos. The rated maximum plate voltage ranges from 180 to 450 volts. The bias voltage, which is a measure of the signal voltage required for full output, ranges from minus 30 volts to minus 84 volts. The power output ranges from 0.375 to 4.6 watts.

For higher power output per tube either pentodes, class B tubes, or the larger high-voltage power tubes are used.

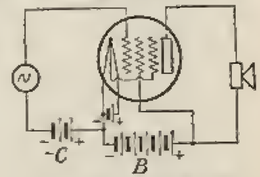


FIG. 22.—Connections of pentode for power output tube.

44. Power Amplifier Tetrodes and Pentodes. A power amplifier tetrode is similar to a power output pentode except that the tetrode does not have a suppressor grid. The electrodes are cathode, control grid, screen grid, and plate. The construction is such that the secondary emission from the plate cannot reach the screen grid. The plate characteristic curves are similar to those for a pentode tube without the secondary emission dip which is characteristic of amplifier (screen-grid) tetrodes. The operating conditions are similar to those used for power output pentodes.

Power amplifier pentodes are high-efficiency power output tubes. They are capable of higher power output with less plate voltage, less power input, and less signal voltage than are triode power amplifier tubes.

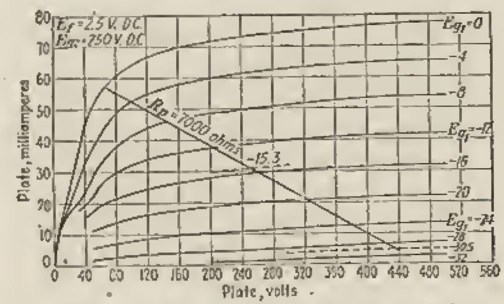


FIG. 23.—Load characteristic of 47 pentode.

Circuits using pentode power amplifier tubes must be more carefully designed to obtain low distortion than are circuits using triode power amplifier tubes.

The electrodes in a power amplifier pentode are cathode, control grid, screen grid, suppressor grid, and plate. The cathode may be either a filament or a unipotential heater type. The control grid connects to a negative bias and the signal voltages. The screen grid connects to the plus B voltage usually of the same value as used on the plate. The screen grid is by-passed with a condenser between it and the cathode. The suppressor grid is usually connected to the cathode inside of the tube.

This grid prevents the screen grid from collecting secondary emission electrons from the plate and thus eliminates the dip in the plate-characteristic curves which appear in the screen-grid types of tubes.

Power amplifier pentodes are used in the power output stage of radio receivers and for operating relays and other devices where high mutual conductance and high plate resistance are desired. Owing to its high plate resistance, it is useful in circuits requiring a constant-current characteristic. For example, for distortionless magnetic deflection of a cathode-ray tube at all frequencies, the current through the deflecting coils should be directly proportional to the signal voltage. When a pentode power amplifier is used, a distortionless pattern results over a range of frequencies for which the deflecting coil impedance is low enough to utilize the pentode constant-current characteristic.

Typical power amplifier pentodes have a plate current from 22 to 34 ma, transconductance from 1,200 to 2,500 micromhos, plate resistance from 35,000 to 100,000 ohms, amplification factor from 80 to 220, and power output from 1.4 to 3.4 watts. The maximum plate voltage ratings range from 135 to 250 volts. The grid-bias voltage which is approximately equal to the peak signal voltage for full output ranges from minus 12 to minus 25 volts.

Pentodes for r-f amplification at high frequencies have been made available. Tubes of this type (6AB7, 6AC7, 1851) have high transconductance (5,000–9,000 micromhos).

45. Dual-grid and Triple-grid Power Amplifiers. Tubes of this class have a cathode, two grids, and a plate. When the two grids are connected together and used as a single grid, the resulting characteristics are suitable for use as a class B power output tube. When the inner grid is used as the control grid and the outer grid is connected to the plate, the resulting characteristics are suitable for class A power amplification, suitable for driving the class B stage.

The characteristics of typical tubes have for class B operation a quiescent plate current of 2 and 6 ma, plate-voltage ratings of 180 and 400 volts, and class B a-f power output for two tubes of 3.5 and 20 watts. For class A operation the maximum plate-voltage ratings are 135 and 250 volts. The corresponding grid-bias voltages are −20 and −33 volts, the amplification factors 4.7 and 5.6, plate resistance 4,175 and 2,380 ohms, the transconductance 1,125 and 2,350 micromhos, and the class A power output 0.17 and 1.25 watts.

The triple-grid power amplifier tube is a universal type of power amplifier tube. With various connections of the grids it may be used as a class A triode, class B triode, or class A pentode power amplifier.

46. Class B Twin Amplifiers. Class B twin-amplifier tubes as the name implies consist of two triode class B a-f amplifier structures in a single bulb.

Like other special class B tubes these tubes operate in a push-pull circuit with zero control-grid bias voltage. The initial plate current of typical tubes ranges from 10 to 17.5 ma. For maximum plate voltages ranging from 135 to 300 volts, the power output of these small-sized tubes ranges from 2.1 to 10 watts. A small power amplifier tube is used to drive the class B tube.

47. Calculation of Power Output and Distortion. To calculate the power output and distortion of a power tube, draw a line on the $I_p - E_p$ characteristic curves representing the load resistance. The line is

drawn through the operating point with the reciprocal slope (voltage to current ratio) equal to the resistance of the load.

A pure sine wave (or cosine wave) signal voltage is assumed to be effective on the grid. At certain values of bias voltage E_c corresponding to selected points on the signal voltage wave, the plate current is noted. With these values of plate current the power output and distortion are calculated as shown by the following example for the type 47 tube:

$E_c = 0$	$= 0$	$I_{max.}$	$= 0.0585$
$E_c = 0.293E$	$= -4.47$	I_x	$= 0.0527$
$E_c = E$	$= -15.25$	I_{po}	$= 0.0320$
$E_c = 1.707E$	$= -26.03$	I_y	$= 0.0107$
$E_c = 2E$	$= -30.50$	$I_{min.}$	$= 0.0052$

Static operating point is $E_B = E_{c2} = 250$ volts, $E_{c1} = -15.25$ volts, $E_f = 2.5$ volts d.c., $I_{po} = 32.0$ ma. Load resistance = 7,000 ohms. The plate current corresponding to values of bias voltage not shown on the $I_p - E_p$ curves can be obtained by plotting a curve of the known values of I_p *versus* E_c from which intermediate points may be read.

$$e_g = E \cos \omega t$$
$$i_p = I_0 + I_1 \cos \omega t + I_2 \cos 2\,\omega t + I_3 \cos 3\,\omega t$$
$$I_0 = +\tfrac{1}{8}[I_{max.} + I_{min.} + 2(I_x + I_{po} + I_y)]$$
$$I_1 = +\tfrac{1}{4}[I_{max.} - I_{min.} + \sqrt{2}(I_x - I_y)]$$
$$= \tfrac{1}{4}[0.0585 - 0.0052 + 1.414(0.0527 - 0.0107)] = 0.0282$$
$$I_2 = +\tfrac{1}{4}[I_{max.} + I_{min.} - 2I_{po}]$$
$$= \tfrac{1}{4}[0.0585 + 0.0052 - 2 \times 0.0320] = -0.00007$$
$$I_3 = +\tfrac{1}{4}[I_{max.} - I_{min.} - \sqrt{2}(I_x - I_y)]$$
$$= \tfrac{1}{4}[0.0585 - 0.0052 - 1.414(0.0527 - 0.0107)] = -0.0015$$
$$\text{Power output} = \tfrac{1}{2}I_1^2R = \tfrac{1}{2}(0.0282)^2 \times 7{,}000 = 2.77 \text{ watts}$$

$$\text{Percentage second harmonic} = \frac{I_2}{I_1} \times 100 \text{ per cent} = \frac{0.00007}{0.0282} \times 100 \text{ per cent} = 0.25 \text{ per cent}$$

$$\text{Percentage third harmonic} = \frac{I_3}{I_1} \times 100 \text{ per cent} = \frac{0.0015}{0.0282} \times 100 \text{ per cent} = 5.3 \text{ per cent}$$

The power output and distortion with various load impedances are shown in Fig. 24. The second harmonic distortion is a minimum near the rated 7,000-ohm. load. The harmonic distortion increases with the load. The total distortion is the vector sum of the second and third harmonics, since the magnitude of the higher frequency components is small. The power output for minimum distortion is near the maximum obtainable.

48. Screen-grid Amplifiers. The screen-grid amplifier tube possesses properties that make it markedly superior to a triode for amplification of r-f or a-f voltages. It is also a good detector tube.

Owing to the low value of control grid to plate capacitance in a screen-grid tube (about 0.01 $\mu\mu$f), the feedback is negligible, and stable operation results without the use of critically balanced circuits. Also the screen grid has the effect of greatly increasing the plate resistance, and, since the transconductance is not decreased, the effective value of amplification factor ($\mu = R_p S_m$) is very large. In use, the high plate resistance puts less shunt-load resistance across any tuned circuit to which it is connected. The result is a more sharply tuned circuit with higher over-all impedance. The net result is higher voltage amplification and greater selectivity. For example, with triode tubes a voltage amplification of 20 per stage is considered high at broadcast frequencies, while with screen-grid tubes a gain in excess of 100 per stage is easily obtained. At intermediate frequencies a gain of 200 to 400 per stage is readily obtained.

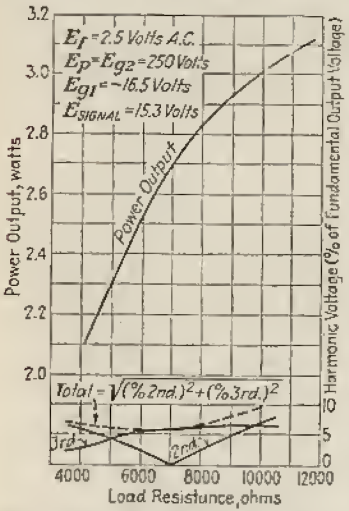


FIG. 24.—Output characteristics of pentode power tube.

The screen-grid tube has a cathode, two grids, and a plate. The inner grid is used as the control grid, to which signal and bias voltages are applied. The outer grid serves as an electrostatic screen between the plate and the inner structure. It is operated at a fixed positive potential ordinarily not higher than about one-half to one-third of the plate voltage.

Typical tubes have plate currents ranging from 1.7 to 4.0 ma, plate resistance from 0.3 to 1.2 megohms, transconductance from 500 to 1,080 micromhos, and grid to plate capacitance from 0.02 to 0.007 $\mu\mu$f.

49. Triple-grid Detector Amplifiers. Triple-grid detector amplifier types have three grids, a cathode, and a plate. Although the three grids all have external terminals to permit various connections in circuits, these tubes are most frequently operated as pentode voltage amplifiers. With this connection the inner grid functions as the control grid, the second grid as the screen grid, and the outer grid as the suppressor grid.

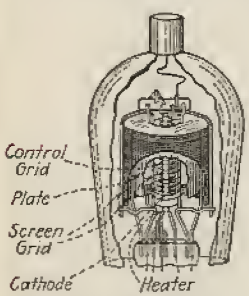


FIG. 25.—Structure of screen-grid tube.

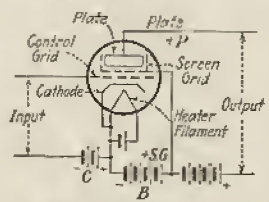


FIG. 26.—Circuit for screen-grid tube.

The operating characteristics are like those of a screen-grid tube except that certain improvements in performance result. The plate

resistance is higher and the grid-plate capacitance is lower than for screen-grid tubes. Owing to the presence of the suppressor grid, the same

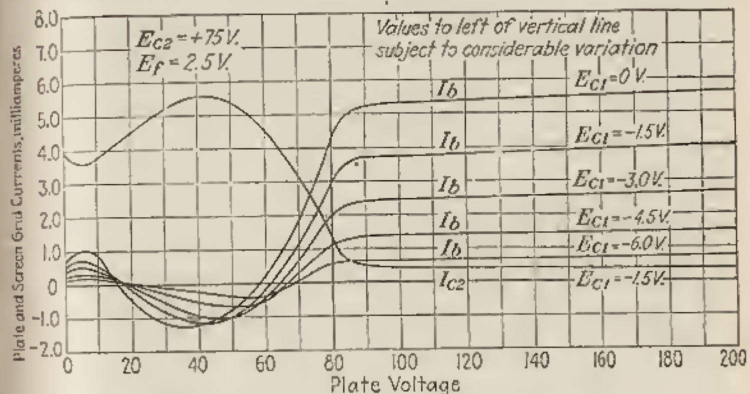


FIG. 27.—Type-24 screen-grid characteristics.

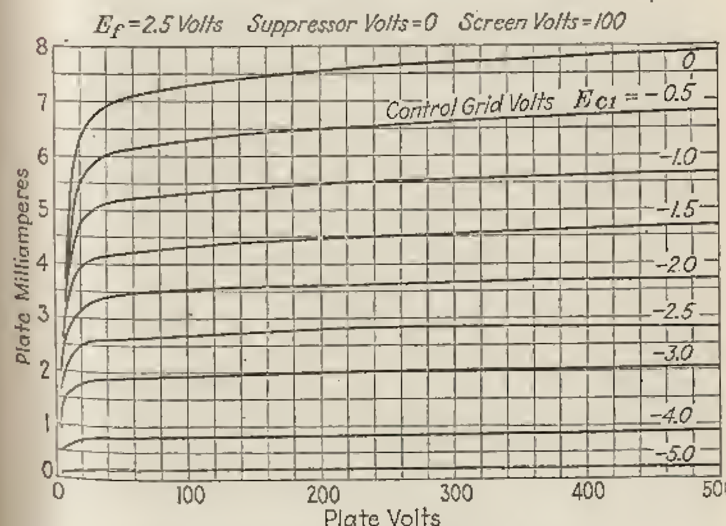


FIG. 28.—Average plate characteristics, type 57.

voltage can be used on the plate and screen grid. This is possible because there is no secondary emission kink in the plate-characteristic curves. This is an advantage, for example, when operating with a

100-volt supply since the use of 100 volts on the screen grid produces high transconductance and also permits higher signal voltages on the control grid. When large amplitude output voltages are required, this connection permits utilization of nearly the entire range of plate voltage. In some r-f circuits the suppressor grid is used for modulation. In one circuit, that of an electron-coupled oscillator, the suppressor grid is grounded so that it functions as an electrostatic screen.

When used as a voltage amplifier for audio frequencies, high gain, large amplitude output, and low distortion can be obtained with this type of

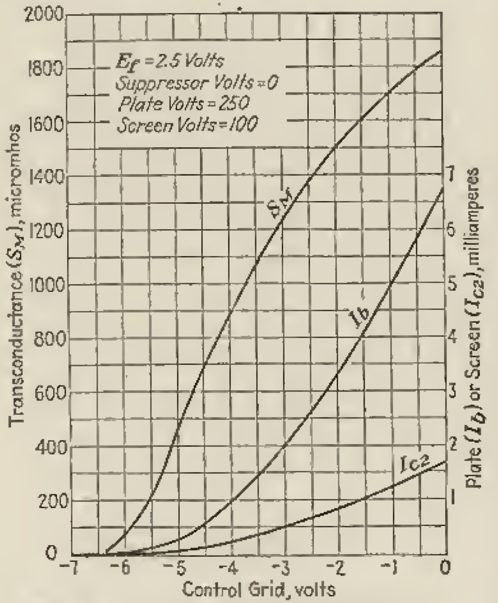


Fig. 29.—Average characteristics, type 57.

tube. Operating characteristics of the 57, for example, are as follows: plate-supply voltage, 250 volts; screen voltage, 50 volts; grid bias, minus 2.1 volts; self-bias resistor, 3,500 ohms; plate-load resistor, 250,000 ohms; grid resistor of following stage, 0.5 megohm; plate current, 0.48 ma. peak output, 60 to 70 volts, voltage amplification, 100.

As a detector, owing to the sharp cutoff, the sensitivity is high, and the distortion low. A high-resistance plate load is used. A suitable condition for operating the type 57 is the same as shown above for a-f amplification.

Typical tubes of this class operate with 250 volts on the plate, 100 volts on the screen grid, and minus 3 volts on the control grid. Operating conditions for small r-f voltages are a plate resistance of 1.5 megohms or more, plate current of 2.0 to 2.3 ma, transconductance of 1,225 to 1,350 micromhos, and grid-plate capacitance of 0.007 to 0.010 $\mu\mu$f.

60. Screen-grid Supercontrol Amplifiers. The screen-grid supercontrol amplifier tube differs from the ordinary screen-grid tube in that it has a remote plate-current cutoff characteristic (variable-mu effect) instead of the usual cutoff characteristic of the detector amplifier type of tube. The supercontrol type is designed for use in r-f and i-f circuits where the stage gain is to be controlled by means of grid-bias voltage. It is effective in reducing cross modulation and modulation distortion over a large range of signal voltages. A change in grid-bias voltage from minus 3 volts to minus 40 volts changes the transconductance from 1,050 to 15 micromhos. This corresponds to a change in gain of approximately 70 to 1 per stage. At the minus 40-volt bias point a signal amplitude of approximately 10 volts can be accommodated without serious distortion.

Supercontrol r-f amplifier pentodes with internally connected suppressor grids are operated the same as screen-grid supercontrol amplifier tubes. The plate resistance of this type is somewhat higher. The secondary emission kink in the plate characteristics is eliminated so that screen grid and plate may be operated on the same voltage when low-voltage operation is desired.

61. Triple-grid Supercontrol Amplifiers. The triple-grid supercontrol amplifier types like the triple-grid detector amplifier types have three grids, a cathode, and a plate. This type is particularly suited for use as an r-f and i-f amplifier. With the usual connection of the three grids a pentode type of characteristic results. The operating characteristics are similar to the triple-grid detector amplifier tube except for somewhat lower plate resistance, higher transconductance, higher plate current, and a remote plate-current cutoff characteristic. The remote cutoff characteristic permits a large range of control of amplification of r-f voltages without cross modulation or modulation distortion. It is useful also as a first detector in superheterodyne circuits but is not generally satisfactory for use as the second detector or for use as an a-f amplifier. For these latter applications the sharp plate-current cutoff detector amplifier type should be used.

The characteristics of typical tubes of this type show higher plate current for low bias voltages than for the detector amplifier triple-grid tubes. The plate-characteristic curves show a continuously decreasing effect of grid-bias voltage on plate current as the negative bias voltage is increased (variable-mu effect). This gradual decrease in plate-current and large bias voltage required for plate-current cutoff permits the use of large signal voltages while the tube is biased to reduce amplification without distortion or cross modulation of the r-f and i-f voltages. The plate resistance of this type tends to be less than for the sharp cutoff type. The values of 0.6 to 0.8 megohm are high enough to prevent excessive loading of the tuned circuits. Voltage amplification greater than 100 at broadcast frequencies and from 200 to 400 at intermediate frequencies is readily obtained.

In operation the grid-bias voltage (E_{c_1}) of the 6D6 can be made variable from minus 3 volts to minus 40 volts for gain control of r-f or i-f stages. As a mixer tube a grid bias of minus 10 volts is used for an oscillator voltage of 7 peak volts. Consideration should be given to the amplitude of the signal voltages to be expected in each stage, and the bias-voltage range should be limited accordingly. The signal voltage should never cause the grid to swing far enough in the positive direction to permit grid current to flow, nor far enough in the negative direction to exceed the plate-current cutoff point.

52. Duplex-diode Triodes. The duplex-diode triode tubes have an amplifier triode and two small diodes in a single bulb. Usually the cathode of all units has a common connection. The diodes are small units used with high-resistance loads (peak currents less than approximately 0.5 ma) for detection and gain-regulating circuits. The triode is of the general type of detector amplifier triodes.

Typical tubes of this type have triodes with amplification factors of 8.3 and 100, plate resistance of 7,500 and 91,000 ohms, and transconductance of 1,100 micromhos. The medium-amplification-factor type can be used as a transformer-coupled a-f amplifier, with one diode for detection and the other as an automatic volume control. Various other uses in circuits will be evident. The high-amplification-factor type is suitable for use as a resistance-coupled a-f amplifier, with one diode as a detector and the other for gain control or various other circuit arrangements.

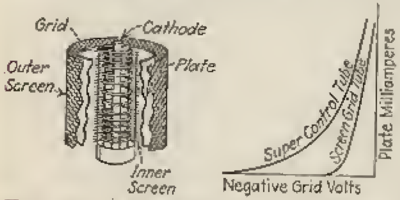


FIG. 30.—Variable-mu or supercontrol tube.

53. Duplex-diode Pentodes. These types, like the duplex-diode triode types, have two small diodes for use as detectors or gain control, and a pentode voltage amplifier unit in a single bulb. The pentode unit may be used for either r-f or a-f amplification. Thus the pentode may operate as an i-f amplifier supplying signal to the diode units functioning as detector and gain-control units, or the pentode may function as a resistance-coupled a-f amplifier following the diode units.

54. Triode Pentode. This tube exemplified by the type 6F7 has a pentode voltage amplifier unit and a small triode unit in a single bulb. The two units operate independently except that a common cathode connection is used. The principal advantage is economy of space; the disadvantage is that failure of one unit requires replacement of the entire tube.

55. Pentagrid Converters. The pentagrid-converter tube has a cathode, five grids, and an anode. It is designed to perform the combined functions of oscillator and first detector in a superheterodyne circuit. The electrodes, starting from the cathode and counting outward (the usual method for designating grids by number), are first (No. 1 grid) the oscillator control grid; next (No. 2 grid) the oscillator anode; grids 3 and 5, connected together within the tube, are used to accelerate the electron stream from the cathode (similar to the operation of the screen grid in screen-grid and pentode tubes); and grid 4 operates as the signal control grid. The grids 3 and 5 shield grid 4 from the inner and the plate electrodes and give the tube a high plate resistance. The high plate resistance permits the use of high-impedance loads resulting in high gain and selectivity.

In operation the electron stream is initially modulated at oscillator frequency by the inner electrodes. The incoming r-f signal, applied to grid 4, further modulates the electron stream, thus producing components of plate current, the frequencies of which are the various combinations of the oscillator and signal frequencies. Since the primary circuit of the first i-f stage is designed for resonance at the i.f. (equal to the difference

between the oscillator and signal frequencies), only the desired i.f. will be present in the secondary of the i-f transformer.

In use, the oscillator coils are designed with a little greater coupling between grid and oscillator anode coils than is commonly used with triode oscillators. A ratio of mutual inductance between these coils to the

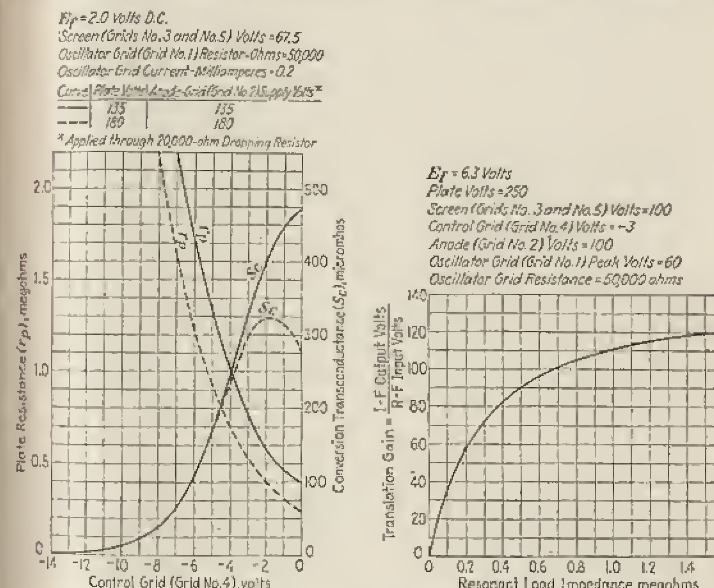


FIG. 31.—Characteristics of pentagrid converters 1C6, left; 6A7, right.

inductance of the grid coil (tuned coil) of 0.25 to 0.40 is satisfactory. Higher values of coupling may cause difficulty in tracking the oscillator frequency to the signal frequency.

The translation gain is given by the relation

$$A = \frac{aS_cZr_p}{(Z + r_p)}$$

where a = voltage ratio of i-f transformer
S_c = conversion transconductance
Z = effective impedance of i-f transformer
r_p = plate resistance of the tube.

With transformers ordinarily used, a translation gain of approximately 60 or with special high-impedance transformers a gain of 100 can be readily obtained.

The characteristics of typical tubes of this type are as shown in the table on page 264.

TYPICAL PENTAGRID CONVERTERS

Type	E_f	I_f	E_p	$E_{c_{3-5}}$	E_{c_2}	E_{c_4}	E_{c_1}	I_p	I_{3-5}	I_{c_2}	I_{c_1}	R_p	S_c
1A6	2.0	0.06	180	67.5	135	−3	50,000 Ω	1.3	2.4	2.3	0.2	0.5	300
1C6	2.0	0.12	180	67.5	135	−3	50,000 Ω	1.5	2.0	3.3	0.2	0.75	325
2A7	2.5	0.8	250	100	200	−3	50,000 Ω	3.5	2.2	4.0	0.7	0.36	520
6A7	6.3	0.3	250	100	200	−3	....	3.5	2.2	4.0	0.7	0.36	520

56. Metal Tubes. Metal radio tubes employ a metallic envelope instead of a glass envelope for maintaining a vacuum in the space surrounding the electrodes of the tube. The structure of the tube electrodes is similar to that used in glass-bulb tubes. The metallic envelope offers the following advantages: elimination of additional tube shields; better shielding of tube electrodes from stray fields than with metal-coated glass bulbs or shield cans; greater mechanical strength; and a smaller size.

This construction also permits better dissipation of the heat developed at the anode.

The metal radio tubes have the octal type of base having a central lug which aids in locating the tube in the socket. There are eight pin positions on the base, the same spacings being used on all types except that pins are omitted or included as needed. This permits the use of one type

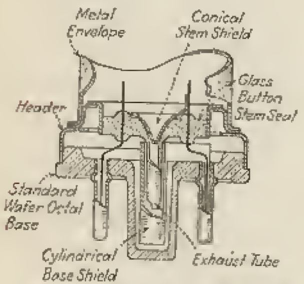


FIG. 32.—Metal tube base construction (shields used in single-ended types).

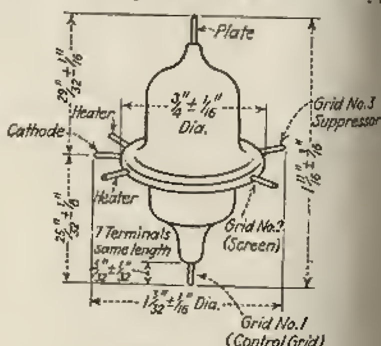


FIG. 33.—Acorn pentode.

of socket for a greater number of tube types. This is of considerable advantage, for example, in testing tubes where the large number of sockets and electrode combinations unduly complicates the equipment.

The characteristics of the metal radio tubes are similar to other tubes of the same general type.

ULTRA-HIGH-FREQUENCY TUBES

57. Receiving Types. At frequencies above 60 Mc (wave lengths below 5 meters) conventional tubes and circuits give poor performance. By means of tubes specially designed for ultra-high frequencies, the performance can be greatly improved. For low-power circuits and for receiving circuits, these special tubes of unusually small dimensions are used. These tubes permit the use of conventional circuits in the frequency range of 60 to 300 Mc and higher.

ACORN RECEIVING TUBES

Type	E_f	I_f	Type cathode	E_p	E_{ct}	E_{c1}	I_p	I_{ct}	R_p	μ	S_m	Capacitance, in micromicrofarads		
												G-P	G-C	P-C
Triode (Detector, Amplifier, Oscillator)														
955	6.3	0.15	{H-C O}	180	.....	− 5.0	4.5	...	12,500	25	2,000	1.4	1.0	0.6
957	1.25	0.05	{F O}	135	.....	− 5.0	2.0	...	24,600	16	650			
958	1.25	0.10	{F O}	135	.....	− 7.5	3.0	...	10,000	12	1,200			
Pentode (Detector, Amplifier)														
954	6.3	0.15	{H-C O}	250	100	− 3.0	2.0	0.7	1.5×10^6	2,000	1,400	G-P 0.007	Input 3	Output 3
959	1.25	0.05	{F O}	135	67.5	− 3.0	1.7	0.4	0.8×10^6	480	600			
Pentode (Super-control R-f Amplifier, Mixer)														
956	6.3	0.15	{H-C O}	250	100	− 3.0	5.5	1.8	0.8×10^6	1,440	1,800	0.007	2.7	3.5
				250	100	−45	...	...	....	....	2			

A small glass bulb with the electrode connections sealed directly through the center and end portions of the bulb is used. There is no base on these tubes. The electrode terminals appear directly on the bulb and are made strong enough for insertion in a socket. The electrodes are similar to those in other types of tubes except that, owing to the unusually small dimensions, special design and construction are required.

Some of the advantages of this type of tube are low electrode capacitance, low electrode connecting lead inductance, small electron transit time, and small space requirement.

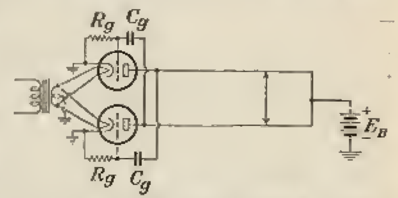


FIG. 34.—Tuned transmission line push-pull oscillator.

High-frequency circuits consisting of small coils with condenser tuning have been used with these tubes for frequencies as high as 300 Mc. Special care is required to reduce inductance of the connecting leads and to obtain good by-passing. Better results at these frequencies and higher can be obtained with the relatively large distributed circuits of the transmission-line type. An example of this type of circuit is the push-pull oscillator shown in Fig. 34. In this circuit good stability is obtained with freedom from "dead spots" often observed with single-tube oscillators tuned over a wide range of ultra-high frequencies.

Performance comparable with other types of receiving tubes is readily obtained when these tubes are used as r-f or a-f amplifiers, detectors, mixers, or oscillators in all of the lower frequency ranges.

TYPICAL BEAM-POWER AMPLIFIER TUBES

Type	E_f	I_f	Type cathode	E_p	E_{c2}	E_{c1}	I_p	I_{c2}	Plate load	Sm	Po	Bulb	Base
Receiving Types													
6L6	6.3	0.9	H-C / O	350	250	−18	54	2.5	4,200	5,200	10.8	MT-10	Octal 7 pin
6V6	6.3	0.45	H-C / O	315	225	−13	34	2.2	8,500	3,750	5.5	MT-8	Octal 7 pin
6Y6-G	6.3	1.25	H-C / O	200	135	−14	61	2.2	2,600	7,100	6.0	ST-14	Octal 7 pin
25L6	25.0	0.3	H-C / O	110	110	−7.5	49	4	2,000	8,200	2.2	MT-8	Octal 7 pin
35A5-LT	35.0	0.15	H-C / O	110	110	−7.5	40	3	2,500	5,800	1.5	T-9	Octal 8 pin
Transmitting Types													
807	6.3	0.9	H-C / O	600	275	−78	100	9		6,000	37	ST-16	Medium 5 pin
1614	6.3	0.9	H-C / O	375	200	−35	88	9		6,050	17	MT-10	Octal 7 pin
1619	2.5	2.0	F / O	400	300	−55	75	10.5		4,500	19.5	MT-10	Octal 7 pin
814	10.0	3.25	F / TT	1,600	300	−90	150	24		3,300	160	T-16	Medium 5 pin
813	10.0	5.0	F / TT	2,000	400	−90	180	15		3,750	260	T-20	Giant 7 pin
828	10.0	3.25	F / TT	1,500	400	−100	180	28		4,500	200	T-16	Medium 5 pin
Push-pull Type													
832	6.3	0.8	H-C / O	400	250	−60	90	18		3,000	22	T-16	Special

58. Beam-power Tubes. Beam-power tubes employ the principle of directing the electrons into beams to obtain improved tube performance. These tubes have a cathode, control grid, screen grid, and plate. Beam-forming plates, located between screen grid and plate, assist in limiting the spread of the beams.

The wires of the control grid and screen are aligned and so spaced that the electron current from the cathode is focused into a series of beams passing between the grid wires. This reduces the screen-grid current. It also makes it possible to use a close-spaced plate with suppressor action

supplied by space-charge effects produced between the screen and the plate.

The resulting plate-current *versus* plate-voltage characteristic curves are practically of the ideal pentode type. The knee of the curve occurs at a low voltage permitting increased plate voltage swing. The plate current is also increased by the amount that the screen-grid current is reduced. Thus the screen dissipation is reduced, the power output is increased, and efficiency increased. A high value of power sensitivity is obtained with these tubes.

The characteristics of typical beam-power tubes are as shown in the table on page 266.

59. Ultra-high-frequency Transmitting Types. Power-amplifier tubes and oscillator tubes of conventional design show a rapid decrease in power output and efficiency as the operating frequency is increased in the region beyond about 50 Mc.

ULTRA-HIGH-FREQUENCY TRANSMITTING TUBES

Type	E_f	I_f	Type cathode	Maximum plate dissipation, watts	E_p	E_{c1}	I_p	μ	Po	Capacitance, micro-microfarad G-P	G-C	P-C	Bulb	Base
RCA 834	7.5	3.25	F / TT	50	1,250	−225	90	10.5	75	2.6	2.2	0.6	S-21	Medium 4 pin
WE 304A	7.5	3.25	... ...	50	1,250	.....	100	11	..	2.5	2.0	0.67	S-21	Medium 4 pin
RCA 1628	3.5	3.25	F / TT	40	1,000	−65	50	23	35	2.0	2.0	0.4	T-8	Special
WE 316A	2.0	3.5	F / TT	30	400	.....	75	5.0	...	1.8	1.0	0.75	Special	Special
Experimental small double lead (W.E.) tube	1	4	... ...	25	300	.....	...	9	...	0.9	1.0	0.7	Special	Special
RCA 887	11	24	F / T	1,000	3,000	−500	400	10	800	6.9	2.5	2.7	Water-cooled anode	Special
RCA 888	11	24	F / T	1,000	3,000	−300	400	30	800	7.8	2.8	2.5	Water-cooled anode	Special

The performance of tubes of this type has been improved and extended into the u-h-f range by methods of design similar to those used in the Acorn receiving tubes. The use of short heavy lead wires is effective in reducing lead inductance. Close spaced electrodes reduce the transit time of the electrons between the electrodes and permit high mutual conductance with small cathode area. Small-sized electrodes keep the electrode capacitance low. Dielectric losses are kept at a minimum by elimination of the base, by sealing the leads through a good quality of glass, and by supporting the electrodes with a minimum of dielectric materials.

Because of the small size and close spacing of the electrodes, these tubes must be designed to withstand high temperatures or to dissipate a considerable amount of power on the electrodes in order to obtain a high power rating.

In some types such as the RCA 1628 and the WE 316A this has been accomplished by the use of special materials, notably tantalum, for the grid and plate electrodes.

Other types such as the RCA 887 and RCA 888 employ a water or forced air cooling of a copper anode which forms part of the external envelope. The grid is made of tantalum to withstand high temperature and is cooled by conduction and by proximity to the water-cooled anode.

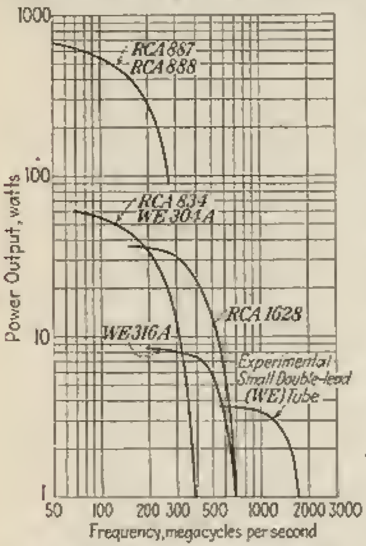


FIG. 35.—Performance capability of ultra-high-frequency power tubes.

60. Ultra-high-frequency Tubes Employing New Principles of Operation. Various types of u-h-f tubes have been described in the literature which employ principles differing in certain respects from the conventional tubes. Notable among these are the velocity-modulated tubes, the Klystron, and the inductive-output tubes.

Velocity-modulated tubes employ the input signal voltage to change the velocity of the electrons in a constant-current electron beam. The velocity-modulation is converted into a current (space charge) modulation by means of a drift tube, a retarding field, or deflection. The space-charge modulated current may be utilized with a conventional plate output electrode or with the newer inductive-output circuits.

The advantages of velocity modulation are reduced input-loading (the transit-time loading effect such as occurs at the grid of conventional tubes) and freedom from critically close-spaced grid electrodes. The transconductance is, however, much lower than can be obtained with the conventional control grid.

The inductive-output tubes employ conventional control-grid modulation of the current but direct the beam of electrons through a cavity resonator in such a manner that the electron beam induces current in the cavity resonator circuit. The electron beam current is collected at low voltage, thus keeping (plate) losses at a minimum.

The Klystron tube consists of two cavity resonators, one (input) arranged to produce a velocity modulation of a constant-current beam of electrons, the other (output) to absorb energy from the electron beam after it has been converted to current modulation.

GAS-FILLED TUBES

There are a variety of useful functions performed by the many types of gas-filled tubes. These tubes, after evacuation, are filled to a low pressure with an inert gas such as argon, neon, or krypton or with mercury vapor.

There are two principle classes of gas-filled tubes, according to the type of discharge occurring. In one class a *hot cathode* emits electrons in sufficient quantity to carry the current. The gas ions act only to reduce the space charge, thus allowing a large current to flow with small voltage drop in the tube. In tubes of the other class is a *cold cathode* with a self-sustaining gas discharge, having a high-voltage gradient close to the cathode and a low-voltage gradient throughout a relatively long positive column.

Examples of the first class of gas-filled tubes are the hot-cathode mercury-vapor rectifier tubes and the hot-cathode gas-triode tubes known as thyratrons.

In the second class of gas-filled tubes are the voltage regulator tubes, the cold-cathode gas rectifier tubes, the cold-cathode gas-triode relay tubes, and the a-c surge and protector tubes.

For hot-cathode mercury-vapor rectifiers see the section on Rectifiers and Power-supply Systems.

61. Cold-cathode Gas-filled Rectifiers. Cold-cathode rectifier tubes for low-power applications are usually filled with an inert gas such as helium or argon. A starting voltage of a few hundred volts is ordinarily required to start the discharge. The voltage drop in the tube falls to a relatively low value when current is flowing, but the voltage drop and tube losses are higher than for hot-cathode mercury-vapor tubes.

These tubes are used in circuits where the saving in filament power is important.

It is sometimes necessary to take precautions to avoid radiation of h-f noise generated by the breakdown surge in the tube. Small chokes in the plate leads, by-pass condensers from each plate to the transformer center tap, or a shield around the tube and circuit may be required.

The following are typical cold-cathode rectifiers:

Type	D-c output			Tube voltage drop average	Starting voltage per plate	Peak voltage, plate to plate	Peak plate current max., ma
	Current		Voltage max.				
	Max., ma	Min., ma					
BR	50	..	...	..	...	600	200
BH	60	..	300	..	...	1,000	400
BA	180..	..	200	..	...	1,000	1,000
OZ4 } OZ4G }	75	30	300	24	300	1,000	200

62. Negative-grid Gas Triodes. If a grid electrode is introduced between the cathode and anode of a suitably designed gas-filled tube, the starting of the discharge can be controlled. If the grid is sufficiently negative completely to cut off all electron flow from the cathode, the gas in the tube remains in its normally unionized condition.

When the grid voltage is made less negative or the plate voltage more positive, a point is reached at which electrons begin to escape. These electrons produce ionization in the gas, which in turn helps more electrons to escape, so that the process is cumulative. The current builds up within a few microseconds to a value limited only by the impedance in the external circuit.

After the gas becomes ionized, the grid ordinarily has no further effect on the flow of the plate current. The grid is said to be covered with a *sheath* of positive ions which neutralize the negative field of the grid.

Typical Hot-cathode Gas Triodes (also Known as Thyratrons and Grid-glow Tubes)

Type	E_f	I_f	Type cathode	Plate current, ma		Plate voltage maximum
				Average	Peak	
Negative-grid Gas-filled Tubes						
884	6.3	0.6	O H-C	2–3	300	300
885	2.5	1.4	O H-C	2–3	300	300
FG-178	2.5	2.25	O F	125	500	310
ELC1A	2.5	6.0	O F	400	2,500	170
FG-81	2.5	5.0	O F	500	2,000	180
Negative-grid Mercury-vapor Tubes						
KU-636	2.5	6.0	O F	100	300	750
FG-65	2.5	2.0	O F	125	500	1,000
FG-17	2.5	5.0	O F	500	2,000	2,500
KU-627	2.5	6.0	O F	640	4,000	2,500
KU-638	2.5	6.0	O F	640	2,500	10,000
FG-27	5.0	7.0	O F	2,500	10,000	1,000
FG-57	5.0	4.5	O H-C	2,500	15,000	1,000
KU-628	5.0	11.5	O F	4,000	16,000	2,500
DKU-623	5.0	20.	O F	10,000	40,000	2,500
FG-29	5.0	17.5	O H-C	12,500	75,000	3,500

If the plate current is stopped long enough for ionization to subside (usually about 0.001 sec. or less), the grid will again exert control. Thus when a gas triode is used as a (60-cycle) rectifier, the output current can be controlled, since the control of the grid is reestablished during the negative half cycle.

The control characteristic for a gas triode may be a single curve relating plate voltage to grid voltage at which the discharge starts. For large negative grid voltages this is usually a straight line, since the ratio of

the voltages is nearly constant. Near zero grid voltage the characteristic shows appreciable curvature.

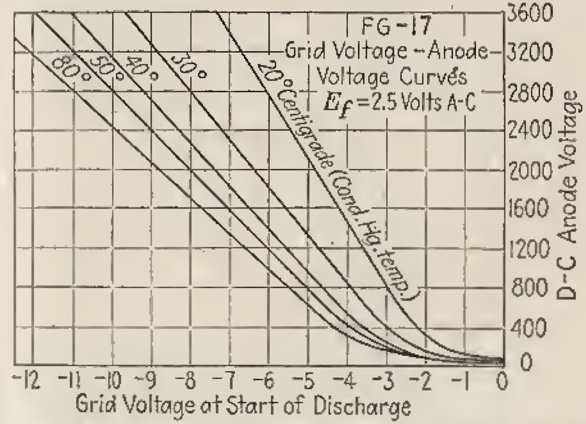


Fig. 36.—Typical mercury-vapor grid-controlled rectifier characteristics.

Mercury-vapor tubes show a different control characteristic curve for different temperatures of the condensed mercury.

Typical Hot-cathode Positive-grid Gas Triodes

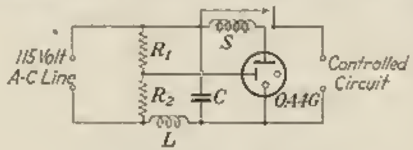
Type	E_f	I_f	Type cathode	Plate current, ma		Plate voltage maximum
				Average	Peak	
KU-610	2.5	6.5	O F	400	800	750
FG-33	5.0	4.5	O H-C	2,500	15,000	1,000
FG-67	5.0	4.5	O H-C	2,500	15,000	1,000*
FG-118	5.0	20.	O H-C	12,500	75,000	10,000

* Inverter.

Cold-cathode Gas-triode Tubes

Type	Plate current, ma		A-c plate voltage r.m.s.	A-c starter electrode voltage minimum peak	Remarks
	Average	Peak			
OA4-G	25	100	105–130	110	Starts with 55 peak r-f volts plus 70 peak a-c volts
FG-157	10	50	220	. . .	A-c positive control
KU-618	15	100	. . . .	. . .	Positive control

63. Positive-grid Hot-cathode Gas Triodes. Gas-triode tubes designed for positive grid control are used in circuits where it is desirable to eliminate the negative grid voltage. The cathode to grid region is so



L,C = High-Q Tuned Circuit for r-f Signal
R_1 = 15,000 ohms (½ watt)
R_2 = 10,000 " "
S = Relay

FIG. 37.—Cold-cathode tube remote-control circuit.

shielded from the anode field in these tubes that a small positive voltage is required to give the electrons enough velocity to ionize the gas. Owing to the high grid power required, these tubes are not used so generally as are the negative-grid tubes.

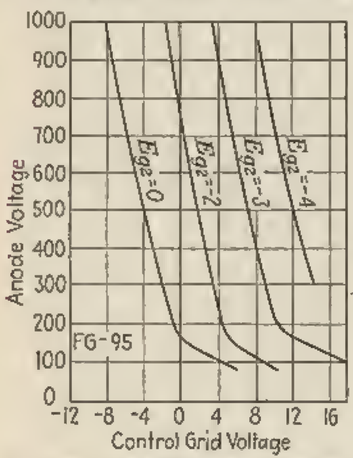


FIG. 38.—Characteristics of shield-grid thyratron.

64. Cold-cathode Gas Triodes. These tubes are most useful for controlling relays or other applications where it is desirable to keep the power consumption low during long stand-by periods.

They are usually designed for positive-grid control. An electrostatic impulse picked up on a suitable electrode or a voltage developed in a tuned circuit can be used for control.

A circuit for remote control by means of h-f impulses over the power circuit is shown in Fig. 37.

65. Gas Tetrode Tubes. Gas tetrode tubes have a second grid or shield grid. By the use of two grids the current to the control grid can be reduced and the sensitivity of control increased. The start of the discharge in these tubes is determined by the relative value of the voltages on the two grids as well as by the plate voltage. The control characteristics for a tube of this type are shown in Fig. 38.

66. Pool-cathode Tubes. Tubes having a pool of mercury as the cathode are termed *pool-cathode tubes*. They are cold-cathode tubes with a self-sustained discharge. High current densities obtained in these tubes produce a low internal voltage drop. The electrons are said to be emitted from the "spot" on the cathode by "field emission" due to the high voltage gradient occurring close to the cathode surface.

TYPICAL HOT-CATHODE GAS TETRODE TUBES

Type	E_f	I_f	Type cathode	Plate current, ma		Plate voltage peak	
				Average	Peak	Forward	Inverse
Negative-grid Gas-filled Tubes							
2051	6.3	0.6	O, H-C	75	375	350	700
2050	6.3	0.6	O, H-C	100	500	650	1,300
FG-98	2.5	5.0	O, F	500	2,000	180	180
FG-154	5.0	7.0	O, F	2,500	10,000	1,000	1,000
FG-180	5.0	18.0	O, F	6,400	25,000	1,000	1,000
Negative-grid Mercury-vapor Tubes							
FG-97	2.5	5.0	O, F	500	2,000	1,000	1,000
FG-95	5.0	4.5	O, H-C	2,500	15,000	1,000	1,000
FG-105, FG-172	5.0	11.0	O, H-C	6,400	40,000	1,000	1,000

Pool-cathode tubes are used for power installations where high currents are required. They can be used for spot welders and other apparatus requiring extremely high peak currents.

With the "immersion starter" (ignitron) type tube the starting of the discharge can be controlled by means of an electrode of high resistance material (Carborundum, Glowbar, or Thyrite) immersed in the mercury pool. Voltage applied to this electrode produces sparking which starts the main discharge.

POOL-CATHODE IMMERSION-STARTER TUBES

Type	Current, amperes		Voltage peak	
	Average	Peak	Forward	Inverse
FG-139	15	1,000	1,000	1,000
KU-637	20	1,000	750	750
KU-639	50	2,000	750	750
FG-179	75	5,000	1,000	1,000

67. Surge and Protector Tubes. These tubes are two-electrode gas-discharge tubes. They are connected across a line or circuit for protection against excess voltage. When the voltage exceeds the breakdown rating of the protector tube, a discharge takes place which limits the voltage.

The following are typical tubes of this type:

A-c Surge and Protector Tubes

Type	Breakdown voltage	Current, amperes	
		Average	Peak
PJ-20	70-160	0.25	200
DKX-642	400	....	50

68. Ballast Tubes and Voltage Regulator Tubes. A ballast tube or current regulator tube is used as a resistance connected in series with a load in which it is desired to maintain constant current. It consists of a wire filament enclosed in a bulb filled with a gas. The temperature-resistance characteristic of the filament is such that ballasting action is obtained.

Within the limits of its useful operating range a small change in current is accompanied by a relatively large change in its terminal voltage. When connected in series with a load, any change in the applied voltage is absorbed largely by the ballast tube, hence the current and voltage at the load remain approximately constant. Since it has a limited range of operation, it must be designed for a definite value of current and voltage.

The ballasting action changes slowly and may require several minutes to reach equilibrium. Consequently it is used for compensating slow changes, such as line voltage changes occurring during different parts of the day and not for momentary fluctuations.

Because of the high operating temperature of the bulb, in use the precaution is usually taken to enclose the ballast tube in a wire gauze or perforated metal shield (the soft glass bulb may develop a strain crack, and explode, especially if it accidentally comes in contact with a cold metallic object).

Typical Ballast Tubes

Type	Current range	Voltage range	Maximum ambient temperature, degrees Fahrenheit	Over-all dimensions, inches
896	0.225–0.275	5–8	150	1 3/16 × 3 [illegible]
7A	0.50 –0.53	3–10	150	1 5/16 × 3 [illegible]
B6	0.96 –1.00	15–21	150	1 1/16 × 3 [illegible]
B4	1.24 –1.36	105–125	150	2 11/16 × 9 [illegible]
876	1.63 –1.77	40–60	150	2 1/16 × 7 [illegible]
886	1.97 –2.13	40–60	150	2 1/16 × 7 [illegible]

A voltage regulator tube is a gas-discharge tube. It has, in its simplest form, two electrodes between which a self-maintained gas discharge takes place. The voltage across the discharge remains approximately constant for a considerable range of the discharge current.

When connected in parallel with a load, small variations in the applied voltage (from a source with sufficient resistance) or changes in the load current are absorbed by a change in the current in the voltage regulator tube, its terminal voltage remaining constant.

It is designed for a definite operating voltage. The current must remain between rated maximum and minimum values. A starting voltage somewhat higher than the operating voltage must be exceeded in order to initiate the discharge.

It is effective in regulating momentary fluctuations as well as for steady conditions. Because of this rapid response it serves not only to regulate the voltage but also to by-pass ripple voltage. It is sometimes used in place of a by-pass condenser at very low frequencies where the size of a condenser would be prohibitive.

Typical Voltage Regulator Tubes

Type	Voltage		Current milliamperes			Over-all dimensions, inches
	Operating	Starting minimum	Peak	Maximum	Minimum	
991	48–67	87	3.0	2.0	0.4	5/8 × 1 15/16
874	90	125	...	50	10	2 3/16 × 5 5/8
VR105-30	105	137	...	30	5	1 9/16 × 4 1/8
VR150-30	150	180	...	30	5	1 9/16 × 4 1/8
KX-641	110	...	...	...	...	2 × 8 3/4

CATHODE-RAY TUBES

A cathode-ray tube is an electron tube in which a beam of electrons (cathode rays) is focused and deflected so that patterns (wave forms or pictures) are formed. The patterns may be made visible on a fluorescent *screen* such as is employed in an oscillograph or in television viewing (receiving) tubes, or may be used with mosaics or other means such as are used in television pickup (transmitting) tubes.

69. Principles of Operation. Early types of cathode-ray tubes employed gas at a low pressure to assist in focusing the beam and in some high-voltage types to generate electrons by means of a discharge in the gas.

In modern high-vacuum tubes the electrons emitted by a thermionic cathode are focused into a beam by means of either electrostatic or electromagnetic fields applied at one or more positions along the beam.

Fields applied near the cathode (ordinarily by the electrodes in the tube) perform the functions of accelerating and controlling the electron flow, concentrating the electrons into a small area (called the *crossover point*), and forming a beam. The beam passes through a final focusing field which focuses the beam to a *spot*. By deflection of the beam the spot is made to move, thus tracing patterns in accordance with the applied deflecting fields.

The electrode structure from cathode to final focusing field is commonly known as the *electron gun*.

The deflecting fields are usually applied in the region beyond the electron gun, *i.e.*, between the final focusing field and the screen (or the surface on which the spot is focused).

For maximum deflection sensitivity the distance from deflecting field to screen should be large. Thus the deflecting fields are usually applied as near to the final focusing field as is permissible without excessive distortion.

Tubes employing electrostatic fields ordinarily have electrodes for the purpose within the tube. Voltages applied to the electrode terminals produce the electrostatic field.

Since electromagnetic fields (low frequency) pass through glass with negligible distortion, it is most convenient to use external coils for tubes employing electromagnetic fields. Current through a coil arranged in the proper position produces the electromagnetic field.

There are available tubes employing electrostatic (final) focusing, electromagnetic (final) focusing, electrostatic deflection (deflection plates), electromagnetic deflection (deflection coils), and in some cases a combination of these.

70. Screen Size. The viewing screen of standard types of cathode-ray tubes ranges in size from 1 to 12 in. in diameter. Experimental tubes ranging up to about 30 in. in diameter have been demonstrated. Owing to the tremendous atmospheric pressure (14.7 lb. per square inch) on large bulbs these tubes are sometimes made of metal.

Screen sizes ranging up to 5 in. are commonly used for laboratory oscillographs or for viewing by two or three persons. For viewing by larger groups, classroom demonstrations, lectures, etc., a 9-in. or larger screen is desirable.

71. Screen Material. Screen materials might be classified according to color of fluorescence, to persistence (which is the time required for the phosphorescent afterglow to disappear), or to efficiency.

Medium-persistence screens are available in green, yellow, and white fluorescent colors. The green (willemite) screen is probably most satisfactory for general use. Its efficiency, including visibility, is high. Stationary patterns can be readily photographed. The yellow and white screens are less efficient than the green but are preferred because of color for television use.

Short-persistence screens of a blue color are used for photographic recording. The short persistence permits continuous moving-film recording. Ordinary blue-sensitive photographic emulsions can be used with these screens. Long-persistence screens of a bluish color are useful for observing the complete trace of a phenomenon that occurs slowly or for direct comparison of the traces on the screen after the beam deflection has ceased. Because of the lower intensity of the persistent image it is viewed best in subdued light.

72. Operating Voltages. The high-voltage supply for (sealed-off high-vacuum) cathode-ray tubes ranges from 250 to 1,000 volts for a few low-voltage oscillograph tubes and to about 15,000 volts for tubes used for recording transient phenomena.

For oscillograph tubes, operating voltages of 1,000 to 3,500 volts are satisfactory for most purposes. Higher voltages are useful when additional brightness is needed to speed up photographic recording. For television viewing, voltages of 6,000 to 7,000 volts are commonly used to increase the brightness and detail of the pictures.

73. Types of Deflection. The electrostatic deflection tubes are used most for general oscillograph purposes. Since almost negligible power is required by the deflection plates, they can be connected across almost any circuit in which it is desired to observe the voltage variations. The deflection is directly proportional to the voltage.

TYPICAL CATHODE-RAY TUBES

Type No.	Screen				Focus type	Deflection			Anode high voltage (range)	Grid cut-off (bias) voltage
	Size diameter, inches	Color	Persistence	Use		Type	Sensitivity for minimum anode voltage			
							D_1D_2, mm/v	D_3D_4, mm/v		
1803-P4	12	W	M	Tel.	es	em	.....	.....	6,000– 7,000	
1804-P4	9	W	M	Tel.	es	em	.....	.....	6,000– 7,000	
903	9	G	M	Os.	es	em	.....	.....	1,000– 7,000	−120
914	9	G	M	Os.	es	es-4	0.204	0.260	2,500– 7,000	
1800	9	Y	M	Tel.	es	em	.....	.....	3,000– 7,000	− 75
7AP4	7	W	M	Tel.	es	em	.....	.....	3,500	
904	5	G	M	Os.	es	es-2-em	0.40	.....	1,000– 4,600	−140
905	5	G	M	Os.	es	es-4	0.38	0.46	1,000– 2,000	− 60
907	5	B	S	Os. Photo	es	es-4	0.38	0.46	1,000– 2,000	− 60
909	5	B	L	Os. View	es	es-4	0.38	0.46	1,000– 2,000	− 60
912	5	G	M	Os.	es	es-4	0.053	0.102	5,000–15,000	−125
1801	5	Y	M	Tel.	es	em	.....	.....	2,500– 3,000	− 35
1802-P1	5	G	M	Tel.	es	es-4	0.50	0.55	1,200– 2,000	
1802-P4	5	W	M	Tel.	es	es-4	0.40	0.44	1,500– 2,000	
906/906-P1	3	G	M	Os.	es	es-3	0.55	0.58	600– 1,200	
906-P4	3	W	M	Tel.	es	es-3	0.55	0.58	600– 1,200	
908	3	B	S	Os.	es	es-3	0.55	0.58	600– 1,200	
910	3	B	L	Os.	es	es-3	0.55	0.58	600– 1,200	
911	3	G	M	Os.	es	es-3	0.55	0.58	600– 1,200	
902	2	G	M	Os.	es	es-3	0.28	0.33	400– 600	− 80
913	1	G	M	Os.	es	es-3	0.15	0.21	250– 500	− 90
1849	$3^{9}/_{16} \times 4^{3}/_{4}$	Iconoscope (film)			es	em	.....	.....	1,200	− 20
1850	$3^{9}/_{16} \times 4^{3}/_{4}$	Iconoscope (direct)			es	em	.....	.....	1,200	− 30
1898	3	Monoscope (Pic.)			es	es-3	(125)	(135)	750– 1,300	− 70
1899	5	Monoscope (500)			es	em	.....	.....	1,000– 1,700	− 60

The magnetic deflection tubes are preferred for television work. Good television pictures can be produced with electrostatic deflection tubes, but the required deflection voltage is too high when the high-voltage anode (and screen) is operated above approximately 2,000 volts.

74. Deflection Sensitivity. The deflection sensitivity for the electrostatic types at rated minimum anode voltage ranges from 0.08 to 0.58 mm per volt. For an anode voltage of 1,000 volts the sensitivities range from about 38 volts per inch deflection up to 680 volts per inch. Most types have sensitivities in the range 38 to 80 volts per inch. The sensitivity is decreased in proportion to the increase in anode voltage.

The deflection sensitivity of magnetic-deflection types depends upon the ampere turns in the deflecting magnet and upon the length and arrangement of the coils.

75. Modulation Characteristics. For television applications the modulation characteristics are important. The beam current should focus into a spot almost as small as the width of one line in the picture. The spot size should not change as the beam current is modulated to change the picture brightness. A value of transconductance which is high permits small signal voltages from the video amplifier. Electrostatically focused tubes should have sufficiently good regulation in the voltage supply for the second (focusing) anode to accommodate the second anode current modulation without defocusing.

The maximum beam current which can be obtained without the above effects causing loss of picture detail determines the maximum picture brightness.

PHOTOELECTRIC TUBES

76. The Photoelectric Effect. Certain metals, notably the alkali metals, have the property of releasing electrons when irradiated with light of certain wave lengths, notably the wave lengths corresponding to the shorter end of the visible spectrum (violet and ultraviolet). This property is the basis upon which phototubes operate. These are cold-cathode tubes in which the electron flow is controlled by the intensity of illumination permitted to fall upon a light-sensitive surface. There are two types in general use, the high-vacuum tubes and those in which there is some gas. The latter are more sensitive, but there is not the linear relation existing between light intensity and current flow that is characteristic of the vacuum types.

Phototubes have found application in sound motion pictures, translating variations in film density (or a variable area of blackened film) into sound variations, and in industry where they perform certain control functions through the medium of a beam of light. In the laboratory phototubes are often used as a means of measuring intensity of illumination either for its own sake or as an intermediate method of measuring some other quantity.

Since phototubes are not used to any extent in radio communication, they will not be discussed further in this volume. For further data see Zworykin and Wilson, "Photocells and Their Application," John Wiley & Sons, Inc.; Henney, "Electron Tubes in Industry," McGraw-Hill Book Company, Inc.; Fink, "Engineering Electronics," McGraw-Hill Book Company, Inc.

INTERELECTRODE CAPACITANCE

77. Tube-equivalent Network. The capacitances between the grid, plate, and filament of a triode are illustrated in Fig. 39 and also the

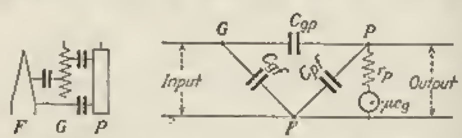


FIG. 39.—Interelectrode capacitance network.

equivalent mesh network. These are the direct interelectrode capacitances of the tube. In general, an n-electrode tube has N direct interelectrode capacitances, where

$$N = \frac{n}{2}(n - 1)$$

The direct interelectrode capacitance is the standard method of specifying the tube capacitances. It is preferred to the older methods of measurement with one electrode floating or between one electrode and the other electrodes connected together. Either of these methods leads to results which are not independent of the particular arrangement of apparatus. The direct interelectrode capacitance is the same regardless of the type of measuring circuit. The capacitance of the socket and socket connections is not included. The tube is usually measured with the cathode cold. When the cathode is heated and voltages applied, the capacitance may change a small amount.

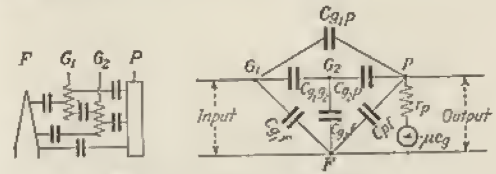


FIG. 40.—Tetrode network.

The three direct capacitances of a triode are grid-plate capacitance (C_{gp}), grid-cathode capacitance (C_{gf}), and plate-cathode capacitance (C_{pf}). The grid-plate capacitance allows energy feedback from the plate to the grid circuit having an important effect on the stability and input impedance. The grid-cathode capacitance and the plate-cathode capacitance shunt the input and output load impedances having some effect on the tuning or frequency characteristics.

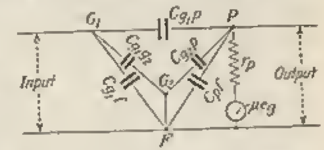


FIG. 41.—Equivalent network of screen-grid tube.

The direct interelectrode capacitances of a tetrode are represented in Fig. 40. The six direct capacitances form a three-mesh network. When the tetrode is connected as a screen-grid tube, the screen grid G_2 is effectively grounded. The three-mesh network is reduced to an equivalent single-mesh triode network. The screen-grid cathode capacitance (C_{g_2f}) is effectively short-circuited by a large by-pass condenser. The control-grid to screen-grid capacitance ($C_{g_1g_2}$) is in parallel with the control-grid to cathode capacitance (C_{g_1f}). The screen-grid to plate capacitance (C_{g_2p}) is in parallel with the plate-to-cathode capacitance (C_{pf}). The equivalent network is shown in Fig. 41.

The capacitances of a screen-grid tube are usually stated as the maximum grid-plate capacitance (C_{g_1p}), the average input capacitance

$$(C_{g_1f} + C_{g_1g_2})$$

and the average output capacitance ($C_{pf} + C_{g_2p}$).

78. Measurement of Interelectrode Capacitance. The direct interelectrode capacitance can be measured with the bridge circuit of Fig. 42. The electrodes to be measured are connected to terminals AB. The remaining electrodes and any shields are connected to ground terminal G. When the bridge is balanced, the capacitance is

$$C_{AB} = C_{gp} = \frac{R_1 C}{R_2}$$

The resistance R corrects the phase and balances the effect of the capacitance across R_2.

Any leakage resistance R_{AB} across C_{AB} will cause an error. If the leakage resistance R_{AB} is known, the capacitance C_{AB} is given by the relation

$$C_{AB} = \frac{R_1 C}{R_2} \cdot \sqrt{1 - \frac{1}{\omega^2 \left(\frac{R_1 C}{R_2}\right)^2 R^2_{AB}}}$$

For example if $(R_1C/R_2) = 5.0$ $\mu\mu$f, the frequency is 1,000 cycles, and R_{AB} is 100 megohms; the correction factor is approximately 0.95 and $C_{AB} = 4.75$ $\mu\mu$f.

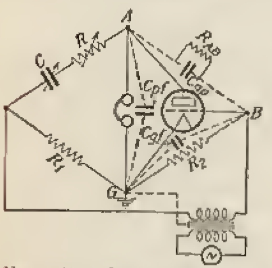


FIG. 42.—Measurement of tube capacitances.

79. Radio-frequency Method. An r-f method of measuring the direct interelectrode capacitances is shown schematically in Fig. 43. The r-f oscillator supplies sufficient voltage to cause a current through C_2 which

FIG. 43.—Method of measuring tube capacitances.

can be measured with the thermocouple TC. The capacitance C_1 does not affect the measured current if the voltage E is held constant. The reactance of capacitance C_3 is high with respect to the low-resistance thermocouple. The indicating microammeter I has one side grounded. An r-f choke L and by-pass condenser C keep r-f currents out of the meter I. When the voltage E and current I are known, the capacitance C_2 is given by

$$C_2 = \frac{I}{\omega E}$$

If a standard variable capacitance of slightly greater range than C_2 is available, a substitution method can be used. The standard capacitance is connected across C_2. It should be enclosed in a grounded shield. The small capacitance to the shield is in parallel with C_1 and C_3.

In use, the meter reading I is noted with the tube in place. The tube is then removed, and the standard capacitance is increased until the same meter reading I is obtained. The difference in the two readings of the standard capacitance is the value of the tube capacitance C_2. The r-f voltage E should be constant. The absolute value of the voltage and current need not be known. A thermocouple with a filter and meter connected in series with a small capacitance across the oscillator terminals can be used as the voltage indicator.

80. Grid-plate Capacitance of Screen-grid Tubes. The direct grid-plate capacitance of screen-grid tubes is a small fraction of a micro-microfarad. Bridge measurements are not generally satisfactory. The r-f substitution method is convenient for this purpose. Figure 44 is the schematic circuit. C is a standard capacitance having a range equal to

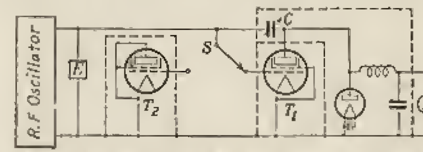


FIG. 44.—Measurement of screen-grid plate-grid capacitance.

the range of capacitances to be measured. Coaxial cylinder capacitors can be constructed accurately covering an extremely small capacitance range. The thermocouple current indicator should be replaced with a sensitive indicator such as a tube rectifier or carborundum crystal. The plate of the tube should be shielded from the grid. A balancing tube T_2 of the same type as the tube T_1 being measured serves to maintain the tube input capacitance load on the oscillator. The low-capacity switch S is first thrown to the tube T_1 under test, and the reading of the meter noted. The switch is then thrown to the balance tube T_2 and the standard condenser C adjusted to give the same reading on the meter. The grid-plate capacitance is equal to the change in the standard capacitance.

81. Receiving Tube Bases. The bases of all standard types of receiving tubes fit one of the following types of sockets:

WD 4-pin.
Small nub 4-pin.
4-pin.
5-pin.
6-pin.
7-pin small.
7-pin medium.
Octal.
Lock-in types (trade names Loktal and Octalox).
Button-base.

The *WD 4-pin* (used on type 11) and the *small nub 4-pin* (used on type V99) types are now practically obsolete.

The *4-pin socket* accommodates the *small 4-pin base*, the *medium 4-pin base*, and the *tapered small 4-pin base*.

The *5-pin socket* is used for both the *small 5-pin base* and the *medium 5-pin base*.

The *6-pin socket* holds the *small 6-pin base* and the *medium 6-pin base*.

The *7-pin small-type base* requires a *7-pin small-type socket*.

The *7-pin medium-type base* has its pins arranged in a larger diameter circle than the *7-pin small base* and requires a *7-pin medium socket*.

The *Octal base* (first used on metal tubes) has eight equally spaced pins arranged around a central locating lug. On tube types requiring less than 8 pins, some pins may be omitted, but the positions of the other pins remain unchanged. Thus all Octal-base tubes fit the same *Octal socket*. The *small wafer Octal*, the *intermediate-shell Octal*, the *dwarf-shell Octal*, the *small-wafer Octal with sleeve*, the *small-shell Octal*, and the *medium-shell Octal bases* all fit the same type *Octal socket*.

Lock-in type bases (trade names Loktal and Octalox) have a central locating lug with provision for locking the tube in the socket. The eight small pins are equally spaced, with pins omitted when not needed. Special constructional features are employed in various types of these tubes, although they can be used in the same *lock-in type socket*.

The *button base* is an especially small base designed for use on miniature tubes. There are seven small pins which are usually molded directly into the glass. The pins are not equally spaced, thus assuring correct position in the *button-base type socket*.

The cap connection used on some screen-grid tubes may be either the *small cap* used on types such as the 24, 57, etc., or the *miniature cap* such as is used on metal-type tubes.

The *skirted-miniature cap* requires the same size connection as the *miniature cap*.

References

Carson, J. R.: A Theoretical Study of the Three-element Vacuum Tube, *Proc. I.R.E.*, April, 1919.

Dushman, S.: *Gen. Elec. Rev.*, **18**, 156, 1915.

———: Thermionic Emission, *Rev. Modern Phys.*, **2**, October, 1930.

———, and J. W. Ewald: Graphs for Calculation of Electron Emission from Tungsten, Thoriated Tungsten, Molybdenum, and Tantalum, *Gen. Elec. Rev.*, **26**, No. 3, March, 1923.

———, H. N. Rowe, J. Ewald, and C. A. Kidner: *Phys. Rev.*, **25**, 343, 1925.

Fry, T. C.: Potential Distribution between Parallel Plane Electrodes, *Phys. Rev.*, **17**, 441, 1921; **22**, 445, 1923.

Groszkowski, J.: "Les lampes à plusieurs électrodes," Étienne Chiron, éditeur, Paris, 1927.

"International Critical Tables," McGraw-Hill Book Company, Inc.

King, R. W.: Calculation of the Constants of the Three-electrode Vacuum Tube, *Phys. Rev.*, **15**, No. 4, 256, 1920.

Kusunose, Y.: Calculation of Characteristics and the Design of Triodes, *Proc. I.R.E.*, October, 1929.

Langmuir, I.: Emission from Thoriated Tungsten Filaments, *Phys. Rev.*, **22**, 357, 1923.

———: *Phys. Rev.*, **7**, 154, 302, 1916.

———: *Phys. Rev.*, **34**, 401, 1912.

———: The Effect of Space Charge and Residual Gases on Thermionic Currents in High Vacuum, *Phys. Rev.*, **2**, No. 6, 1913.

———, and Karl T. Compton: Electrical Discharges in Gases, *Rev. Modern Phys.*, **3**, April, 1931.

Llewellyn, F. B.: Operation of Thermionic Vacuum Tube Circuits, *Bell System Tech. Jour.*, **5**, July, 1926.

Peterson, E., and H. P. Evans: Modulation in Vacuum Tubes Used as Amplifiers, *Bell System Tech. Jour.*, **6**, July, 1927.

Richardson, O. W.: "The Emission of Electricity from Hot Bodies," Longmans, Green & Co., 1921.

"Smithsonian Physical Tables," 1927, p. 403.

Thomson, J. J.: "Conduction of Electricity through Gases," 3d ed., Cambridge University Press, 1928.

Van der Bijl, H. J.: "The Thermionic Vacuum Tube," McGraw-Hill Book Company, Inc., 1920.

Vodges, F. B., and F. R. Elder: Formulas for the Amplification Constant for Three-electrode Tubes in Which the Diameter of Grid Wires Is Large Compared to the Spacing, *Phys. Rev.*, **24**, December, 1924.

Warner, J. C.: Some Characteristics and Applications of Four-electrode Tubes, *Proc. I.R.E.*, April, 1928.

SECTION 9

VACUUM-TUBE OSCILLATORS

By Robert I. Sarbacher, ScD.[1]

1. Classification of Oscillators. A vacuum-tube oscillator is usually defined as a device which converts power obtained from a d-c source into alternating power. Some of the principal types of vacuum-tube oscillators are listed below.

1. Feedback oscillators.
2. Negative-resistance oscillators.
3. Beat-frequency oscillators (heterodyne).
4. Magnetostriction oscillators.
5. Relaxation oscillators.
6. Magnetron oscillators.
7. Klystron oscillators.
8. Barkhausen-Kurtz oscillators.
9. Mechanical-electronic oscillators.

It is customary to classify oscillators into two groups. The first group is characterized by a definite frequency and by nearly sinusoidal voltage. When such a system is started from rest, it will complete a large number of oscillations before reaching the steady state in which each cycle is identical with the preceding one. The members of this group may be called *harmonic oscillators* and include all the members of the above list except the relaxation oscillators.

Characteristics of the second group are rather indefinite frequency and extremely non-sinusoidal operation. When systems of this type are started from rest, they may reach the steady state in a very few cycles. Such oscillators are referred to as relaxation oscillators.

The harmonic oscillators which comprise the first group are of greater importance and find much wider application than do the relaxation oscillators of the second group. The latter are seldom used directly in communication circuits. The frequency is not very definitely fixed by the circuit elements and so is relatively easily controlled by an external influence.

A system will not oscillate unless the various elements are properly proportioned, even if the configuration is correct. Fundamentally this means that, unless as much energy is delivered to the oscillatory circuit through the tube as is dissipated in each cycle, the oscillations cannot be maintained. For any system to oscillate stably at a definite amplitude, it is necessary that it involve some non-linearity.[2] In some cases the nature of the non-linearity is not obvious, but the effect is always there. The source of the non-linearity may be in the tube, in the resonant circuit, or

[1] Illinois Institute of Technology, Chicago.
[2] Le Corbeiller, P., *I.E.E.*, *Wireless Sec.*, **11**, 292, 1936.

in a special control circuit. In any system in which the tube itself is non-linear, the stabilization is necessarily accompanied by the generation of harmonic currents and voltages, although the effect of these may be reduced by highly selective resonant circuits.

When extremely accurate frequency control is required, low-powered oscillators are used because it is then less difficult to meet the conditions required by a high degree of frequency stabilization. One or more buffer amplifiers may be used under these circumstances to meet the power requirements of the particular application. When frequency stability is not particularly important, high-power oscillators may be used, with which tube efficiencies approaching 90 per cent may be obtained.

2. Feedback Oscillators. Oscillations may be generated with an amplifier that is connected so as to supply its own input voltage in the correct phase and magnitude. This is possible since the power required to supply the input voltage to the amplifier tube is much less than the amplified output. Oscillators operating in this way may be classed as feedback oscillators. Circuits which may be used for this purpose are shown in Fig. 1. It can be shown that, in general, the alternating voltage fed back to the grid of the oscillator tube should be 180 deg. out of phase with the alternating voltage across the plate terminals of the tube. The voltage fed back to the grid must further have an amplitude sufficient to develop the output power necessary to maintain this voltage. In the tuned-plate, tuned-grid, and Meissner oscillators, Figs. 1*a*, 1*b*, and 1*c*, this is achieved through mutual induction between the plate and grid circuits. In the Hartley and Colpitts oscillator circuits, Figs. 1*d* and 1*e*, the grid voltage is obtained by applying a portion of the voltage developed in the resonant circuit to the grid. In the tuned-plate tuned-grid circuit Fig. 1*f*, the energy necessary to develop the grid voltage is fed back to the grid circuit through the plate-grid capacity of the tube.

The frequency at which oscillation occurs approximates very nearly the frequency of the resonant circuit associated with the oscillator. In the case of the Meissner and tuned-plate tuned-grid oscillators it may be shown that, since these circuits have more than one resonant branch, they may operate at either of two frequencies when the coupling between the two circuits exceeds a critical value.

One of the most popular oscillator circuits is the Hartley. This popularity is due partly to the fact that the criterion of oscillation is not at all critical. The amplitude of oscillation is easily controlled by adjustment of the tap on the oscillator coil. For the generation of low audio frequencies, with good wave form, the Hartley oscillator is particularly suitable. This is because the resonant circuit condenser shunts both the coils L_p and L_g and hence gives a lower frequency of oscillation for a given total inductance than either the tuned-plate or tuned-grid oscillators.

The Colpitts oscillator is less convenient to operate as a variable frequency oscillator since it is necessary to vary both C_g and C_p in order to maintain oscillations. However, with this type of oscillator the impedance of both the plate and grid circuits to harmonics is quite low since these circuits are shunted by the condensers C_g and C_p, respectively. This low-impedance path for the harmonic currents results in a reduction in the harmonic voltages generated in the system and hence improves the wave form.

Any of these fundamental oscillator circuits may be modified to employ two tubes in push-pull or in parallel. With parallel operation, parasitic

oscillation which may be developed must be suppressed (see Radio-frequency Amplifier Section). With push-pull operation the harmonic content is decreased and the frequency stability increased over that of

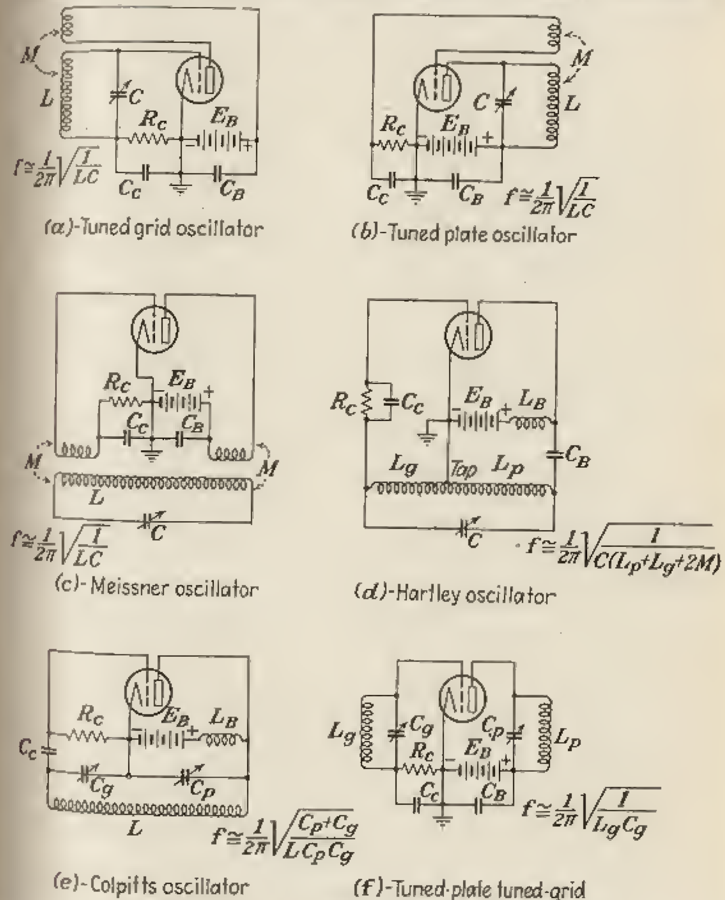


FIG. 1.—Types of feedback oscillator circuits.

the single-tube circuits. Push-pull operation of oscillators is particularly advantageous at high frequencies.

When the plate supply voltage is connected in series with the plate inductances, the connection is called *series feed* (see Figs. 1*a*, 1*b*, 1*c*, 1*f*). When the plate-supply voltage is connected through a choke coil to the plate of the oscillator tube and the oscillating circuit is connected through a blocking condenser to the plate, the connection is called *parallel feed* (see Figs. 1*d*, 1*e*). In practice it is usually desirable to employ parallel feed since with this type of connection the resonant circuit is isolated from the d-c supply voltage.

Fixed bias is rarely used in feedback oscillators. Resistance bias, as shown, is almost always used in order that the oscillator be self-starting and that stable operation, as is discussed in Art. 14, be ensured.

3. Frequency Stabilization. 1. *Causes of Frequency Variation.* There are three major causes which contribute to undesired frequency variation.[1] These are a result of changes in (*a*) tube characteristics, (*b*) circuit parameters, and (*c*) mechanical arrangements of the oscillating system resonant circuit.

Changes in the tube characteristics result in general from changes in (*a*) plate potential, (*b*) grid potential, (*c*) filament potential, (*d*) filament emission due to causes other than (*c*) (such as disintegration of the filament), (*e*) changes in spacing of tube elements, and (*f*) interruptions (keying) of the circuit.

Changes in the values of circuit parameters result from (*a*) changes in temperature of inductances, (*b*) changes in temperature of capacitances, and (*c*) changes in power taken from oscillator.

Changes in the mechanical arrangement of the circuit elements may be caused by (*a*) vibration, (*b*) electromagnetic force, (*c*) electrostatic force, and (*d*) temperature.

2. *Methods of Preventing Frequency Variation.* The plate and grid polarizing potentials may be stabilized by employing voltage-regulating devices. Since the oscillator tube is usually operated so that there is an abundant space charge in the neighborhood of the filament, slight variations in heater voltage and cathode emission have a small effect. The spacing of the electrodes, which may vary slightly with tube temperature, affects the interelectrode capacities. This effect may be minimized by the choice of a larger capacitance in the resonant circuit and the use of circuits in which the resonant circuit capacitance shunts the capacity between the plate and grid. At higher frequencies, where the resonant circuit capacitance becomes of the same order of magnitude as the plate-grid capacitance, this effect is increased.

Changes in the values of circuit parameters such as those caused by temperature variation of inductance and capacitance can be reduced by (*a*) temperature-controlled compensating inductances, and (*b*) temperature-controlled compensating condensers.[2]

It can be shown that the frequency of oscillation will be affected by changes in load unless the power output can be taken from the system without changing the current in the inductance.[3] The use of buffer amplifiers or electron coupling makes it possible to prevent changes in load from affecting the frequency. Electron coupling is discussed in Art. 30.

By careful mechanical and electrical design it is possible to reduce the effects caused by vibration.

Rather than attempt to prevent the variation of the tube characteristics in oscillators in which a high degree of frequency stability is required,

[1] GUNN, R., *Proc. I.R.E.*, **18**, 1560, 1930; LLEWELLYN, *Proc. I.R.E.*, **19**, 2063, 1931.

[2] GUNN, R., *Proc. I.R.E.*, **18**, 1565, 1930; GRIFFITHS, W. H., *Wireless Eng.*, **11**, 234, 1934.

[3] REICH, H. J., "Theory and Application of Electron Tubes," p. 332, McGraw-Hill Book Company, Inc., 1939.

we may design them in such a way that the variation of these factors does not affect the frequency of oscillation. The principal methods of doing this are by the use of the following:

1. Piezoelectric crystals.
2. Magnetostriction rods.
3. Selection filters.
4. Resistance stabilization.
5. Reactance stabilization.
6. Bridge stabilization.

4. Piezoelectric Crystal Oscillators. Oscillators which have the highest frequency stability are those which are controlled by crystals. This control is based upon the piezoelectric effect, which is a means by which a mechanical motion is coupled to an electric circuit. When a piezoelectric material is compressed or stretched in certain directions, electric charge appears on the surfaces of the material that are perpendicular to the axis of strain. Conversely, when such a material is placed between two metallic surfaces and a potential difference applied to them, mechanical strains are set up within the crystal. The amplitude of the voltage produced by mechanical strain may vary from a fractional voltage to several hundred volts.

There are a number of crystalline substances which exhibit this piezoelectric effect; among them are quartz, Rochelle salts, and tourmaline. Of these, quartz is used almost exclusively for controlling the frequency of oscillators because it is mechanically rigid, inexpensive, and has a low temperature coefficient. Tourmaline is sometimes used (although it is more expensive than quartz) because it may be ground to a smaller size and, therefore, have a higher resonant frequency. When crystals are used in electric circuits, they are cut into bars, slabs, and other geometric configurations which bear certain relations to the crystal structure. The frequency at which the crystals vibrate is determined principally by their physical dimensions. Articles 5, 6, 7, and 8 of this section describe the crystals, crystal cuts, methods of temperature control, and methods of mounting.

There are many circuits in which piezoelectric crystals are employed; two representative circuits are shown schematically in Fig. 2; these have been designed by Pierce.[1] Others are described in more detail in later sections. Analysis of circuits containing crystals is greatly simplified by replacing the crystal with its equivalent electric circuit.[2] This equivalent circuit represents the crystal as a series resonant circuit, consisting of a condenser inductance and resistance shunted by the capacitance of the holder. Care must be taken in the choice of the biasing resistor R_c, shown in the circuits of Fig. 2, since this resistance, in addition to its function of controlling the grid polarizing potential, also controls the a.c. which flows through the crystal. If this current exceeds the safe operating value for the crystal,[3] the crystal may vibrate so violently as to shatter itself.

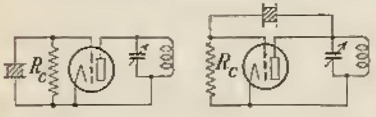
[1] Credit is due G. W. Pierce on many crystal oscillator circuits which have been accredited to others. See his patents U. S. 1789196, filed February, 1924, and U. S. 2133642 through U. S. 2133648 filed between 1926 and 1931.

[2] VAN DYKE, K. S., *Proc. I.R.E.*, **16**, 742, 1928; and MASON, W. P., *Proc. I.R.E.*, **23**, 1252, 1935.

[3] In general, the safe operating value for the current through the crystal may be set approximately at 100 ma. for l-f crystals and about one-half this value for crystals operating above 1 Mc.

If the plate circuit of Fig. 2*a* is inductive, the effective input conductance of the tube is negative, and oscillations may be set up in a resonant circuit connected between grid and filament. To keep the plate resonant circuit inductive, it must be tuned to a frequency slightly higher than that of the crystal. In the circuit of Fig. 2*b*, the crystal is connected between the plate and grid of the oscillator tube. This circuit will oscillate only when the plate circuit is capacitative, and hence the natural frequency of the plate resonant circuit must be slightly lower than that of the crystal.

The resonant curve of a crystal is extremely sharp, and it is this characteristic of the crystal that makes it suitable for use in controlling the frequency of oscillators. The standard measure of the sharpness of resonance of a crystal or an electrical circuit is usually denoted by Q and is numerically equal to the ratio of the total inductive reactance to the total effective resistance of an oscillating circuit. The selectivity Q of the equivalent circuit of a crystal is of the order of magnitude of one hundred times that which can be attained with ordinary inductances and capacitances. In view of this high selectivity, the crystal can oscillate over only a very narrow frequency range. With temperature-controlled crystals, frequency variations of as little as ± 2 parts in 10^7 are not uncommon. With a special circuit described in Art. 13, short-time frequency drift may be kept within ± 6 parts in 10^{10}.



(a) (b)

FIG. 2.—Types of piezoelectric crystal-controlled oscillators.

The output of crystal oscillators may vary from a fraction of a watt to several hundred watts. In applications where extremely constant frequency is required, the oscillators are usually designed for low power output, and one or more buffer amplifiers are used. In this way the crystal current may be kept small and the heating effects due to it minimized. The buffer amplifier also greatly reduces the effect on the oscillator of variations in load. With modern high-gain pentodes, operating in crystal-controlled circuits, reasonably good frequency stability at high power output may be obtained. This stability is usually sufficient for the requirements of amateur communication.

The frequency of negative resistance oscillators may also be controlled by the use of crystals.[1]

5. Piezoelectric Crystals.[2] The occurrence of quartz crystals (the most commonly used of the piezoelectric materials) in the natural state is quite generally known. These crystals, while rarely symmetrical in form, have the general shape of a hexagonal prism, sometimes surmounted on the ends by a hexagonal pyramid. A cross section of a symmetrical crystal is shown in Fig. 3. In this diagram the electric axes (so called because the greatest piezoelectric activity is observed in the direction of these axes) are represented by the lines XX, $X'X'$, and $X''X''$. The other axes, YY, $Y'Y'$, and $Y''Y''$, have been given the name "mechanical axes." Through the point O, perpendicular to the plane of the page,

[1] MacKinnon, K. A., *Proc. I.R.E.*, **20**, 1689, 1932.

[2] See References at end of section.

passes the optic axis (Z-axis) of the crystal. Sections or plates are cut from the crystal for use as highly selective circuit elements.

Crystals cut perpendicular to the X-axis are called X-cut, and crystals cut perpendicular to the Y-axis are called Y-cut or 30-deg. cut (see Fig. 4). Although both the X-cut and Y-cut have been used extensively, they are now largely superseded by more modern cuts which greatly improve the performance of crystals.

6. Piezoelectric Crystal Cuts. One of the objections to the X- and Y-cut crystals is their large temperature coefficient, amounting to -10 to -25 parts per million per degree centigrade for the X-cut, and $+100$ to -20 parts per million per degree centigrade for the Y-cut. When they are used as frequency-control elements, provision must be made to keep their temperature constant. Also these plates often exhibit discontinuous frequency-temperature characteristics. This characteristic of the Y-cut plate can be improved by suitable grinding, while the X-cut plate cannot be improved, and may often be inoperative at the desired frequency of operation.

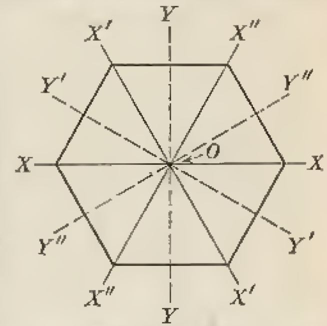
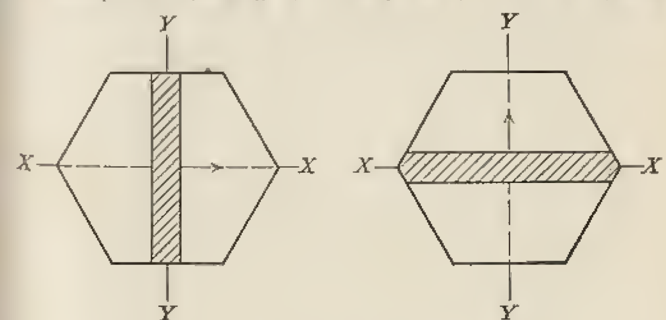


FIG. 3.—Quartz crystal cross section.

From the statement above regarding the range of the temperature coefficient for the Y-cut plate, it might seem possible to get a plate having a zero temperature coefficient. Marrison[1] found this to be the case for the so-called *ring* or *doughnut plate* when operated at a temperature of approximately 40°C. (see Fig. 6). This plate is, however, very difficult



(a) (b)

FIG. 4.—X- and Y-cuts.

to grind and therefore expensive. Moreover, it exhibits a number of spurious resonances near the desired frequency.

More recent work has resulted in the discovery of a number of plates which overcome most of the difficulties encountered with those plates

[1] Marrison, W. A., *Proc. I.R.E.*, **17**, 1103, 1929, and *Bell System Tech. Jour.*, **8**, 493, 1929.

mentioned above. These plates are obtained by cutting the crystal in such a way that at least two faces of the plate are not perpendicular to the crystallographic axes. Some of these plates are considered below.

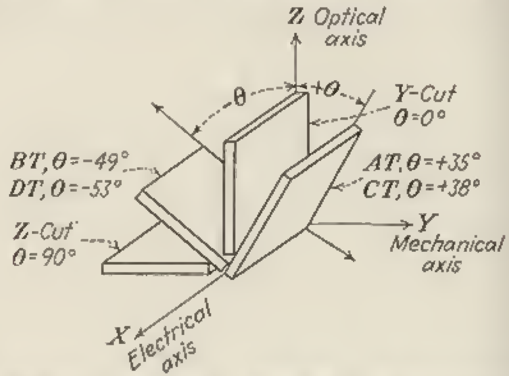


FIG. 5.—Orientation of crystal cuts with respect to the crystallographic axes.

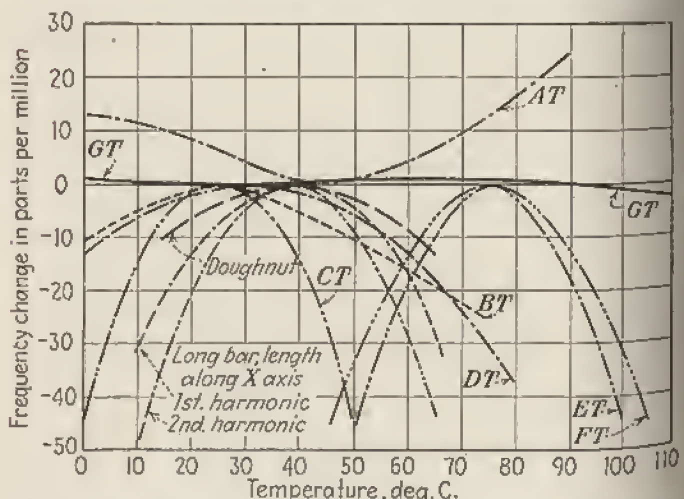


FIG. 6.—The temperature coefficient of frequency for different crystal cuts. (*Courtesy of Bell System Tech. Jour.*)

Two cuts which are suitable for operation above 500 kc are the *AT* and *BT* cuts. These have a zero temperature coefficient when operated at temperatures of approximately 45°C. and 25°C. (see Fig. 6).

The *AT* plate is obtained by cutting the crystal at an angle of rotation about the *X*-axis of 35 deg. (see Fig. 5). The *BT* plate is obtained at an angle of −49 deg. as indicated.

When the thickness of the *AT* and *BT* plates is increased to obtain lower operating frequencies (below about 500 kc), difficulties arise due to coupled modes of vibration. Even though elastic coupling between desired and undesired modes of vibration in these plates is small, it becomes important when the frequencies approach one another, as is the case when the thickness dimension becomes comparable with the other dimensions. To avoid the use of unusually large plates of quartz for lower frequency operation, two new types of quartz crystal elements have been developed. These are known as the *CT* and *DT* plates and are directly related to the h-f low-temperature-coefficient *AT* and *BT* plates. The temperature coefficient of frequency of these new plates may be made zero by operating them at a suitable temperature, as indicated in Fig. 6. These *CT* and *DT* plates are useful as stabilizing elements for oscillators operating between 50 and 500 kc.

The *ET* and *FT* crystal cuts have zero temperature coefficients at higher temperatures than those discussed above. Their useful range, which is from 100 to 1,000 kc, extends to higher frequencies than that of the *CT* and *DT* plates. This is because they operate at a harmonic of the fundamental vibration.

The most recently announced crystal, called the *GT* cut,[1] has a constant frequency over a very wide temperature range. As can be seen in Fig. 6, the shape of the temperature-frequency curve is different from that of the other special cuts. The superiority of this cut, particularly when temperature control is not used, is evident. The *GT* cut is very satisfactory at frequencies near 100 kc.

There is no definitely established frequency limit for quartz plates; the practical limits are being constantly extended. Plates have been used at 20 Mc, and a 1-kc quartz bar has been reported. Quartz plates are rarely called upon to control more than a few watts directly; higher powers are controlled by amplifying the output of the crystal stage.

Several other materials which assume a more or less well-defined crystalline form have been investigated as possibilities for piezoelectric elements. Among these may be mentioned tourmaline and Rochelle salt. The Rochelle salt crystals have, in general, been discarded, although they have found applications in loud-speakers, microphones, and phonograph pickups.

Tourmaline, while it is practically as good as quartz over a great frequency range—and somewhat better than quartz in the range from about 3 to 30 Mc—has the disadvantage of being a semiprecious stone; its cost is, in consequence, out of proportion to its usefulness.

Beyond the range where crystals exert satisfactory control, *i.e.*, about 30 Mc at the present time, special resonant circuits of extremely high selectivity may be used as frequency-control elements (see Art. 23).

7. Temperature Control of Piezoelectric Crystals. Since the resonant frequency of all crystals, particularly of the *X*- and *Y*-cuts, changes with temperature, it is necessary, if a high degree of frequency stability is required, to make some provision to keep the temperature of the crystal constant. In some cases, where every possible precaution is taken to

[1] MASON, W. P., *Bell System Tech. Jour.*, **19**, 74, 1940.

prevent frequency variation, the associated electrical circuit as well as the crystal is maintained at a constant temperature.

Electric ovens suitable for temperature control of crystals are usually designed after the principles given by Marrison.[1] These principles involve the thermal conductivity of the material of which the oven is made, the ambient-temperature range, and the temperature coefficient of the quartz plate.

Briefly stated, the problem is one of accurately determining the temperature at which it is desired to maintain the plate and of causing any slight deviation from this temperature to actuate suitable thermostatic devices, which in turn cause more or less current to flow through the heater associated with the oven.

An example of such a control chamber is given by Marrison as follows:

"It consists of a cylindrical aluminum shell with a wall about one inch thick, with a heater, and with a temperature-responsive element in the wall to control the rate of heating. The aluminum shell has a metal plug that screws into the open end forming a chamber for the crystal which is then completely closed except for a small hole for electrical connections.

"Since aluminum is a good thermal conductor the shell equalizes the temperature throughout the chamber and thus avoids the use of a fluid bath. The main heating coil is wound in a single layer over the whole curved surface of the aluminum cylinder, being separated from it only by the necessary electrical insulation. Auxiliary heating coils are wound also on the ends so as to distribute the heating as uniformly as possible. This, in effect, makes the short cylinder behave like a section from an infinite cylinder. To protect the thermostat from the effect of ambient temperature gradients the heating coil has an outside covering consisting of four layers each of thin felt and sheet copper spirally wound so that alternate layers are of copper and felt, the innermost layer being of felt and the outer one of copper. . . . This covering is very effective in reducing surface gradients since the conductivity in directions parallel to, and perpendicular to, the surface differ by a large ratio."

The thermostat used with these constant-temperature chambers is generally the mercury-column type. This is essentially a thermometer in which contact wires have been fused. At the point on the scale where the operating temperature is located, the glass stem has been drawn out; *i.e.*, if the device is to function at, say 35°C., the stem of the thermometer is constricted and elongated between about 34.5° and 35.5°. One of the contact wires is fused through the glass at the 35° point; the other wire making contact with the mercury at the bulb. This elongation of the stem over a range of 1° or so causes the mercury column to move an appreciably greater distance per fraction of a degree change in temperature.

This type of regulator is very sensitive to minute temperature changes but is expensive, fragile, and cannot carry any appreciable current. For this latter reason, it is customary to utilize the regulator simply to change the grid bias on a vacuum tube; the tube plate circuit includes the winding of a relay which operates with small changes of plate current. This relay, which is generally too small to handle the heater current, actuates still another relay to open or close the heater circuit.

With the advent of the new crystal cuts, the temperature coefficient of frequency is so low that temperature control is normally not required. Some types of service, notably aircraft radio, where ambient temperatures may range from −40°C. to +40°C., still require some kind of temperature regulation, but the requirements are satisfactorily met with a more or less conventional heating chamber and an ordinary bimetallic thermostat.

8. Mountings for Piezoelectric Crystals. There are, in general, two types of crystal holders: those in which the crystal plate is firmly clamped,

[1] MARRISON, W. A., *Proc. I.R.E.*, **16**, 976, 1928. Also see CLAPP, J. K., *Proc. I.R.E.*, **8**, 2003, 1930.

and those employing an air gap between the plate and one of the electrodes. In recent high-precision work, crystals with the electrodes directly plated on them have been used. The holders for plated crystals are relatively simple contacting devices.

The use of a holder with an adjustable air gap permits slight adjustments in frequency to be made. It is preferable, however, in applications requiring the oscillator frequency to be definitely fixed, that the holder clamp the crystal securely. For laboratory use in frequency standards, an air gap may be of considerable value. In some circuits the frequency may be more advantageously varied by connecting a suitable reactance element in series or shunt with the crystal.

While the actual construction of crystal holders is beyond the scope of this discussion, it may be pertinent to point out some of the requirements which must be met by the holder.

These may be enumerated briefly as follows:

1. The electrode surfaces must be lapped perfectly flat and must be entirely free of oil and dirt.
2. The electrodes must be made from metal which will not corrode.
3. Where an air gap is employed, means should be provided for clamping the movable electrode after the final adjustment has been made.
4. Some type of construction is generally necessary which will prevent lateral motion of the plate; this may be accomplished by enclosing the plate and electrodes in close-fitting cases of suitable insulating material.
5. The entire assembly should be made dustproof and evacuated if possible.

The methods by means of which the electrodes are plated directly on the quartz are known as the *sputtering* and *evaporation* processes. Mr. H. W. Weinhart of the Bell Telephone Laboratories has prepared the following description of the technique used in these processes. He states:

"Films of metal can be deposited on quartz by sputtering or evaporating on the material. Some metals sputter much more readily than others, for example, gold, silver and platinum films can be deposited at a greater rate than aluminum. Metals that sputter slowly, are therefore, usually plated on by the evaporation process.

"Sputtering is a process involving the releasing of atomic particles of metal by electron and ion bombardment in a gas. The usual method, when plating with air as the gas, is to place the material on which the metal film is to be deposited in a bell jar with a vacuum pump attached. A cathode of the metal to be plated is mounted about 1½ in. above the recipient, and a small leak valve that can be regulated, is attached to the apparatus, together with an aluminum anode located in the tube connection for vacuum pumping.

"The system is pumped out, with the leak adjusted to maintain a pressure of 0.06 mm of mercury. If a potential of about 1,900 volts is applied between the anode and cathode, through a suitable resistance, the gas in the chamber is ionized and the cathode is bombarded. The atomic particles of metal released from the cathode surface diffuses as a gas and a metallic film is deposited on the quartz.

"Evaporation of metal for the deposition of metal films on quartz is a process in which a vacuum chamber is used that can be pumped out to maintain a pressure of 10^{-4} to 10^{-6} mm of mercury. The evaporation unit can be in the form of a wavy wire, and made from tungsten 1/20,000 in. in diameter wound in a close spiral, one eighth inch in diameter, and then stretched to form wide pitch spiral turns.

"Wire, 1/10,000 in. in diameter, of the metal to be plated, is cut into short lengths and formed into hairpin shape. One piece is placed in each depression in the tungsten wire.

"When the proper pressure is attained in the vacuum system, the tungsten wire is slowly heated, until the metal to be plated is melted and flows over the wire or forms globules in the depressions. Slow heating is necessary to maintain a low pressure by pumping out the liberated gases during this part of the process. By increasing the temperature of the tungsten wire the attached metal is evaporated, and deposits by condensation on the quartz surfaces, thus forming a metal plating.

"For some metals such as chromium and beryllium, the preparation of the evaporator unit differs. For plating chromium the usual procedure is to plate the wavy tungsten wire with chromium electrolytically, and then to evaporate it off. Beryllium can be attached to the tungsten wire by spot welding on small pieces along the length of the wire."

9. Magnetostriction Oscillators. Oscillators having their frequency controlled by magnetostriction rods were first described by G. W. Pierce.[1] Magnetostriction in metals is somewhat analogous to the piezoelectric

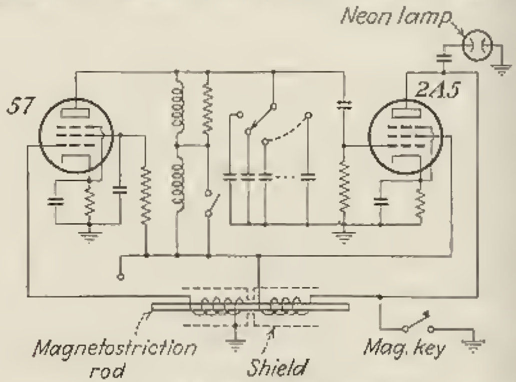


Fig. 7.—Magnetostriction oscillator.

effect in crystals. There is an expansion or contraction of magnetic materials as a result of magnetization and, conversely, a change of magnetic permeability as a result of mechanical stress.

If a rod of magnetostrictive material is placed in an alternating magnetic field, the rod will vibrate longitudinally at a frequency which is twice that of the a.c. producing the field. If, however, the rod is magnetically polarized, the frequency of vibration will be that of the applied a.c. Under this condition the rod may be clamped or pivoted at its exact center, this being a nodal point. For this condition the resonant frequency of the rod (usually in the range from 1,000 cycles to several hundred thousand cycles) is given by

$$f = \frac{v}{2l}$$

where v = the velocity of sound in the rod
l = the length of the rod.

[1] Pierce, G. W., *Proc. Amer. Acad. Arts Sci.*, **63**, April, 1928; reprinted in *Proc. I.R.E.*, **17**, 42, 1929.

The circuit of Fig. 7 shows an improved magnetostriction oscillator.[1] It consists essentially of a two-tube impedance-coupled amplifier having input and output coils shielded from each other except for electromechanical coupling through the vibration of a magnetostrictive rod placed axially in both of them. A neon-glow lamp serves as an indicator of oscillation when connected across the plate coil. Operation of this circuit is dependent upon the correct choice of coupling impedance with regard to the direction of connection of the rod coils and upon the existence of good electromagnetic shielding between the two rod coils. The proper value of the coupling impedance is not at all critical since it requires practically no adjustment over a wide range of frequencies. Magnetostrictive rods for use with this type oscillator have been cut accurately to length to give fundamental frequencies ranging from 5 to 60 kc.

Pierce has given extensive data on oscillators of this type, including such matters as temperature coefficients and values of the function v in the above equation for various magnetostrictive materials.

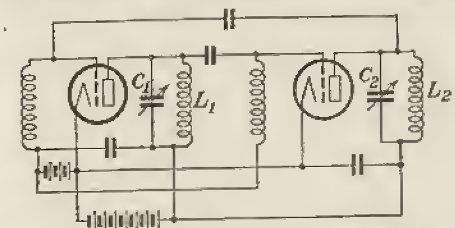


Fig. 8.—Oscillator stabilized by selective filters.

In making magnetostriction rods, nickel, Monel metal, Invar, Nichrome, Stoic metal, and other nickel alloys may be used. Because it is difficult to design magnetostriction rods which have a high natural frequency of oscillation, their use is restricted as cited above. Rods may be designed for very low frequencies by loading them at the ends or by using a tube made of magnetostrictive material which is filled with lead or other material which has a low velocity of propagation of compressional waves. Short-time frequency stabilities of 3 parts in 10^6 have been obtained with oscillators of this type without temperature control. If the temperature of the rod is kept constant, this stability may be increased. By making the rods of special alloys having a low temperature coefficient or making them of a shell of two magnetostrictive materials of opposite temperature coefficient, the change in frequency with temperature may be reduced.

10. Tuned-filter Oscillators. The tuned-filter oscillator is essentially a multistage-feedback oscillator. By feeding back the output of a highly selective multistage amplifier to the input, very good frequency stability may be obtained. Such an oscillator was described by Gunn,[2] and is shown in Fig. 8. Except that the amplification takes place in more than one tube, the principle of operation of this oscillator is the same as that described under Feedback Oscillators. The frequency

[1] Pierce, G. W., and A. Noyes, Jr., *Jour. Acoustic. Sci. Am.*, **9**, 185, 1938.
[2] Gunn, Ross, *Proc. I.R.E.*, **18**, 1560, 1930.

stability is improved by the use of more stages and by the use of more complex filter sections which have a more selective filter action. When oscillators of this type are used at radio frequencies, it is necessary to take particular care that feedback in the individual stages does not occur. Use of tetrodes and pentodes and careful shielding are necessary. Gunn gives the following data as evidence of the excellent stabilizing action. At an a.f. of 1,000 cycles, a 50 per cent change in plate potential of a two-stage system resulted in a frequency shift of less than 1 cycle. At radio frequencies a change in plate potential of 10 per cent results in a frequency shift of 0.0003 per cent of the fundamental frequency. Changing the filament potential 8 per cent changes the frequency less than 0.0003 per cent. The above data was taken with battery-operated filaments. If alternating filament voltage is used, the filament must be of the non-inductive type. The use of a buffer amplifier between oscillator and load will improve the frequency stability.

11. Resistance Stabilization.[1] One of the easiest methods for improving the frequency stability of standard oscillators is by resistance stabilization (see Fig. 9). It was pointed out previously that one of the factors

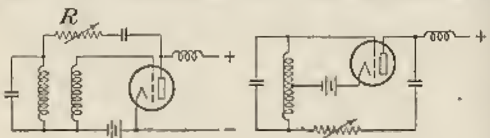


Fig. 9.—Resistance stabilized oscillator.

contributing to frequency drift is change in the plate resistance of the tube. The method of resistance stabilization consists of inserting a high resistance between the plate and resonant circuit of an oscillator so as to make the total effective resistance of the plate circuit so high that variations in the plate resistance of the tube are relatively unimportant. This resistance also performs a second useful function. It makes a convenient means of controlling the amplitude of oscillation by controlling the feedback voltage. Obviously the power consumed by the resistance reduces the efficiency of the system.

Terman has given useful design information for this type of stabilized oscillator.[2] He recommends the following:

1. Amplification factor of tubes should lie between 4.5 and 8.
2. Turns ratio of grid and plate coils should be unity, and coupling should be as close as possible.
3. Feedback resistance should be of the order of from two to five times the plate resistance.
4. Grid bias battery must be used and not grid leak resistance.
5. For audio-frequency oscillators, feedback resistance should not be greater than 500,000 ohms.

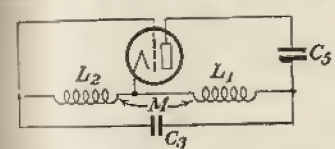
12. Impedance Stabilization. A more general type of stabilization than those previously presented has been worked out by Llewellyn.[3] He has shown that the frequency of oscillation may be made invariant

[1] Horton, J. W., *Bell System Tech. Jour.*, **3**, 508, 1924.
[2] Terman, F. E., *Electronics*, July, 1933, p. 190.
[3] Llewellyn, F. B., *Proc. I.R.E.*, **19**, 2063, 1931; also see Stevenson, G. H., *Bell System Tech. Jour.*, **17**, 458, 1938.

to tube characteristics and hence to polarizing potential by the insertion of capacitance or inductance in series with the grid or plate of the oscillator tube, or both. In his analysis, Llewellyn makes the following assumptions:

1. The resonant circuits of the oscillator have negligible losses.
2. The oscillator tube operates in a linear region of its characteristic.

He then sets up the equivalent circuits for each type of feedback oscillator and the circuit equations applied thereto. From the general solution of these circuit equations he obtains the conditions which make the frequency of oscillation invariant to the tube paremeters. Representative results obtained in this way are shown in Fig. 10. In order for the assumption of negligible losses in the resonant circuits to hold reasonably well, it is necessary that a buffer amplifier be interposed between the oscillator and the load. This buffer stage must be very



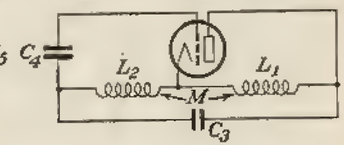
(a)

Fig. 10(a).—Hartley oscillator, plate stabilization.

$$C_5 = C_3\left(\frac{L_0}{L_1 + L_2A^2 - 2MA}\right)$$

where $L_0 = L_1 + L_2 + 2M$,

$$A = \frac{L_1 + M}{L_2 + M}$$



(b)

Fig. 10(b).—Hartley oscillator, grid stabilization.

$$C_4 = C_3A^2\left(\frac{L_0}{L_1 + L_2A^2 - 2MA}\right)$$

where $L_0 = L_1 + L_2 + 2M$

$$A = \frac{L_1 + M}{L_2 + M}$$

loosely coupled to the oscillator so as to draw but a very small fraction of the available power. To meet the second assumption, some form of amplitude control such as described in Art. 14 must be used. Llewellyn further states that with unity coupling between the plate and grid circuits, the frequency of an oscillator depends only upon the inductances and capacitances in the circuit and is independent of plate resistance, grid resistance, and amplification factor, provided (1) that the losses in the external circuit are small and (2) that the harmonic voltages across the tube are small enough to allow the plate and grid impedance to be purely resistive.

The examples of circuit proportions in Fig. 10 will provide impedance stabilization of a Hartley oscillator, provided the assumptions made in text are met. For many more examples see F. E. Terman, "Measurements in Radio Engineering," p. 295, (1936).

When the stabilized circuits that are shown in Fig. 10 are constructed, the stabilizing inductances and capacitances may serve other functions in the circuit. For example, plate stabilizing condensers may serve also as blocking condensers for the plate polarizing potential. Also the grid stabilizing condenser may serve to furnish grid bias when shunted by a high resistance. When the stabilizing condenser is thus shunted, its required value is altered and its effectiveness reduced. The higher the resistance, consistent with the limitation discussed in Art. 14, the smaller its effect on the required value of stabilizing capacity. The correct stabilizing capacity is best determined experimentally. The interelectrode capacitances of the tube are of small impedance in those circuits where these capacitances form a portion of the resonant circuit. In cases where the interelectrode capacitance cannot be combined in this way, variation from predicted performance may be partly explained.

Another factor which may produce variation from the theory is the existence of harmonics as previously mentioned. An effort to provide a low reactance path for the harmonics will reduce their effect (see discussion on Colpitts circuit, Art. 2).

If a variable-frequency oscillator is stabilized in this way, it is necessary to adjust the stabilizing condenser when the frequency is varied. For this type of stabilized oscillator, at 1-Mc operation the frequency varied less than 10 cycles when the plate potential was reduced 50 per cent and practically no change when the filament current was reduced 50 per cent.

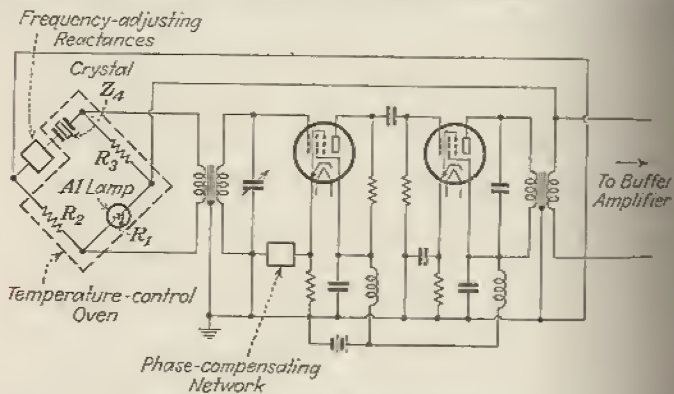


Fig. 11.—Bridge-stabilized oscillator.

13. Bridge Stabilization. The bridge-stabilized oscillator was developed by L. A. Meacham[1] and is a constant-frequency oscillator of extremely high selectivity. Short-time frequency variations no greater than ±6 parts in 10^{10} have been obtained with a single-tube circuit.

This type of oscillator, which consists of an amplifier and a Wheatstone bridge, is shown in Fig. 11.

A crystal Z_4 of high selectivity forms one of the arms of the Wheatstone bridge. Two other arms are made up of the fixed resistances R_2 and R_3. The fourth arm R_1 is a thermally controlled resistance. The output of the amplifier is impressed across one of the diagonals of the bridge, and any unbalanced potential appearing across the conjugate diagonal is supplied to the input terminals of the amplifier. The thermally controlled resistance R_1 is a lamp and is so designed as to keep the bridge out of balance sufficiently to sustain oscillation. Since the temperature of the lamp filament is dependent upon the amplitude of oscillation, any slight variation in this amplitude or in the gain of the amplifier is immediately corrected by a small readjustment of the bridge balance. The frequency of oscillation is stabilized at that value for which the crystal impedance is purely resistive, because only at this frequency can the Wheatstone bridge approach balance. It can be shown by

[1] Meacham, L. A., *Proc. I.R.E.*, **26**, 1278–1294, October, 1938; *Bell System Tech. Jour.*, **17**, 574–591, October, 1938; *Bell Lab. Rec.*, **18**, January, 1940.

means of a vector diagram that a large phase shift introduced in the amplifier results in a very small frequency shift and phase shift in the crystal, owing to the phase magnifying property of a nearly balanced bridge.

When the polarizing potentials are supplied to the amplifier, oscillations build up rapidly since the lamp R_1 is cold and its resistance correspondingly low, resulting in low attenuation of the bridge. When the lamp filament heats up, its resistance increases and approaches the value for which the loss in the bridge equals the gain of the amplifier. If the lamp resistance exceeds its balance value, the unbalance potential becomes too small or even inverted in phase, causing the amplitude to decrease to the equilibrium value. Hence the amplitude of oscillation is also stabilized since the power required to give the lamp a resistance closely approaching that of its balance value is always very nearly the same. Variation in the amplifier gain would cause a readjustment of the tube balance, but resulting variations in the amplifier output or in the value of R_1 would be extremely small.

In place of the crystal in the Z_4 arm of the bridge a coil and condenser connected in series could be substituted. Also a parallel resonance coil and condenser could be used by exchanging its position in the bridge with R_2 or R_3. In Meacham's bridge, Z_4 represents a crystal suitable for operation at its low-impedance or series resonance. This mode of operation minimizes the effects of stray capacitance. He has also found that a small tungsten-filament lamp of low wattage rating is quite suitable. The operating temperature of the lamp is made sufficiently high so that variations in ambient temperature do not appreciably affect balance adjustments. This temperature is found to be low enough to ensure extremely long filament life.

The use of a two-stage amplifier, as shown in Fig. 11, provides high gain and correspondingly high stability. This circuit was designed by Meacham for the Bell System Frequency Standard. Small manual adjustment of frequency is provided by the variable reactances in series with the crystal. Because of the possibility of any tendency of the circuit to break into undesired oscillation as a result of its high gain, the phase-compensating network indicated in the cathode circuit of the first tube is used.

14. Amplitude Control. Control of the amplitude of oscillation is necessary to ensure stable operation. Also amplitude control aids in the reduction of harmonic distortion and in the stabilization of frequency. In the feedback oscillators of Fig. 1, the amplitude of oscillation is usually controlled by the use of the grid-bias resistor and condenser as shown. This aids in making the oscillator self-starting, for initially the bias is zero and the plate current and amplification are large. When any voltage of the frequency of the resonant circuit is set up in the system, caused by thermal agitation or transient conditions, the building up of oscillations will start. This building-up process is accompanied by the flow of grid current, which develops a direct voltage across the grid-resistor-condenser combination, biasing the grid negatively. As the amplitude of oscillation continues to increase, the grid current increases, increasing the grid bias and decreasing the amplification of the tube. This process continues until the amplification is reduced to the point where equilibrium is established. Conversely, any decrease in the amplitude of oscillation causes an increase in the amplification and a reduction in grid bias, tending to produce stable oscillations.

If a fixed bias is used with class C operation of the oscillator, the system will not be self-starting when the plate voltage is applied since the grid bias is greater than the cutoff value.

When the time constant of the grid-resistor-condenser circuit is too large, the bias voltage adjusts itself too slowly with rapid changes in the amplitude of oscillation. This may result in a dying out of oscillations before the bias can change appreciably. When the oscillations have ceased or are about to cease, the condenser charge leaks off through the grid resistance, and oscillations build up again to the equilibrium value. This process may repeat itself, resulting in what are called intermittent oscillations.

Another method which may be used to control the amplitude of oscillation employs a diode rectifier as the limiting device.[1] This type of control is particularly suitable for oscillators operating in class A, in which no grid current flows. Figure 12 shows a Hartley circuit equipped with automatic amplitude control. The action is essentially that of a simple volume-control system employing a diode. By employing a triode, tetrode, or variable-μ pentode, the control system can also be arranged so that it does not start to operate until the amplitude has reached some predetermined level, and in addition the amplification introduced will increase its sensitivity. Equilibrium conditions may be obtained with small amplitudes of oscillation, where the operating conditions are

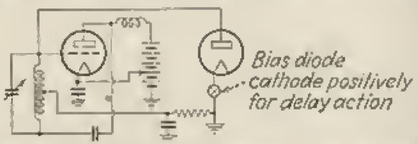


FIG. 12.—Oscillator employing a diode to control the amplitude of oscillation.

substantially those corresponding to class A operation of the oscillator tube. Under these conditions very good frequency stability may be obtained, with good wave form and practically constant amplitude of oscillation as the frequency of oscillation is changed.

Another method for controlling the amplitude of oscillation is described in Art. 13 on the Meacham bridge-stabilized oscillator.

15. Negative-resistance Oscillators. In feedback oscillators it can be shown that a necessary condition for the production of sustained oscillation is that the tube together with the resonant circuit produce an equivalent negative resistance.[2] As distinguished from feedback oscillators, negative-resistance oscillators are those in which the negative resistance of the system does not require the presence of a tuned circuit.

Oscillators of this type are as follows:

1. Dynatron oscillators.
2. Transitron oscillators (negative transconductance).
3. Negative resistance push-pull oscillators.
4. Negative grid-resistance osillators.

16. The Dynatron Oscillator. The dynatron oscillator of Hull[3] (see Fig. 13*a*) depends for its operation on the phenomenon of secondary

[1] ABOUIMHAU, L. B., *Proc. I.R.E.*, **21**, 14, 1933; GROSZKOWSKI, J., *Proc. I.R.E.*, **22**, 145, 1934.

[2] The a-c resistance of a device may be defined as the reciprocal of the slope of its current-voltage characteristic. If this slope is negative for a certain range in voltage the device is said to have a negative resistance throughout this range. Under this condition a positive increment in current through the device results in a negative increment of voltage across its terminals. When the direction of flow of d.c. is opposite to the applied direct voltage, as may be observed in certain devices, such devices are said to have a negative d-c resistance.

[3] HULL, A. W., *Proc. I.R.E.*, **6**, 535, 1918.

emission. He showed that it was possible to use the negative resistance produced by secondary emission for the generation of oscillations. Usually the dynatron oscillator employs a screen grid tube which operates with a plate voltage less than the voltage applied to the screen grid. Under these conditions the characteristic shown in Fig. 13*b* results. It can be seen that there is an appreciable range in which a positive increment in plate voltage causes a negative increment in plate current, *i.e.*, negative resistance. Secondary emission of electrons at the plate causes this negative resistance characteristic and may be explained as follows: The potentials of the control and screen grids determine largely the number of primary[1] electrons which arrive at the plate. The plate potential, however, controls the velocity at which the primary electrons strike the plate. Therefore, the number of secondary electrons[2] produced at the plate increases as the plate voltage is increased. All the secondary electrons produced are drawn to the more positive screen grid, and the effective plate current is the difference between the primary

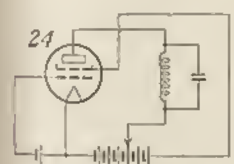


FIG. 13*a*.—The dynatron oscillator.

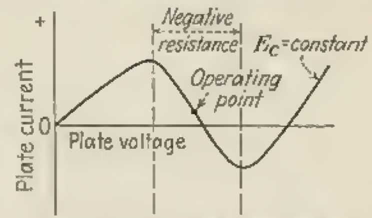


FIG. 13*b*.—Typical plate current-plate voltage characteristic for dynatron operation, showing region of negative resistance.

electrons received at the plate from the cathode and the secondary electrons lost by the plate.

If, as the plate voltage is increased, more electrons leave the plate owing to secondary emission than arrive from the filament, the effective plate current may decrease. This condition results in a negative dynamic resistance, and the characteristic shown in Fig. 13*b* is obtained. Oscillation will be developed if an oscillatory circuit is connected across this negative resistance as shown in Fig. 13*a*, provided the absolute value of the negative resistance is less than, or equal to, the equivalent resistance of the tuned circuit. The amplitude of oscillation may be varied by means of the control-grid voltage, which varies the slope of the current-voltage characteristic in the negative-resistance range.

When designing a dynatron oscillator, the point of operation should be chosen to be in the center of the most linear region of the negative-resistance characteristic,[3] and the amplitude of oscillation should be kept

[1] Primary electrons are those which are emitted from the cathode.

[2] Secondary electrons are those which are obtained from materials as a result of impact of quickly moving electrons which knock electrons out of a solid body when striking with sufficient velocity. One primary electron striking a material at high velocity may produce many secondary electrons.

[3] See F. E. TERMAN, "Measurements in Radio Engineering," p. 289, McGraw-Hill Book Company, Inc., 1935.

small. Under these conditions the curvature in the operating range can be kept small and the harmonic content low.

In addition to excellent wave form the dynatron oscillator possesses good frequency stability and simplicity. The chief disadvantage to this type of oscillator arises from its dependence upon secondary emission, a property which is extremely variable with age and which varies widely in tubes of the same type. With tubes of ordinary size the power output is extremely limited.

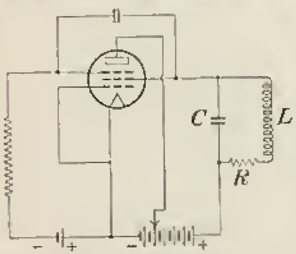


FIG. 14a.—The transitron oscillator.

17. The Transitron Oscillator (Negative Transconductance Oscillator).[1] The name *transitron* has been proposed by Brunetti for the retarding-field negative-transconductance oscillator. This oscillator possesses essentially the same type of negative-resistance characteristic as the dynatron oscillator and has all its advantages without its disadvantages. Its characteristic is independent of secondary emission and remains practically constant throughout the life of the tube. The action of this oscillator, shown in Fig. 14a, is as follows:

The suppressor voltage is chosen so as to make the suppressor grid negative with respect to the cathode. Electrons that have passed through the screen grid are repelled by the suppressor grid and return to the screen because of its high positive voltage. Hence the suppressor grid with its retarding field acts as a virtual cathode. A small negative increment in voltage across the tuned circuit is transmitted to both the screen and suppressor grids, causing the suppressor grid to repel more electrons and the current to the screen grid to increase. Hence the

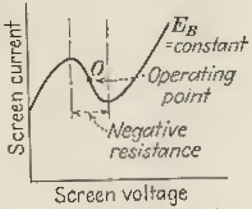


FIG. 14b.

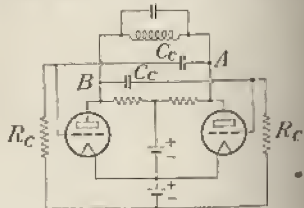


FIG. 14c.

FIG. 14b.—Typical screen current, screen voltage characteristic for transitron operation, showing region of negative resistance.

FIG. 14c.—Push-pull negative-resistance oscillator.

transconductance between the screen and suppressor grids is negative. The characteristic current-voltage curve for this type of oscillator is as shown in Fig. 14b.

This negative transconductance can be employed to produce a negative resistance by the use of the circuit in Fig. 14a. If the equivalent resist-

[1] HEROLD, E. W., *Proc. I.R.E.*, **23**, 1201, 1935. For an excellent treatment on the practical design of the transitron oscillator see C. BRUNETTI, *Proc. I.R.E.*, **27**, 88–94, 1939.

ance of the tuned circuit (which is approximately equal to L/RC) is just equal to the negative reciprocal of the slope of the current-voltage characteristic (Fig. 14b) at the operating point 0, oscillation in the resonant circuit will begin. If L/RC is increased, the amplitude of oscillation increases. As with the dynatron oscillator, it is desirable to keep the amplitude of oscillation small so as to keep the wave form and frequency stability good.

When a small negative bias is applied to the control grid, the total current flowing to the screen grid may be controlled, and the negative slope of the current-voltage characteristic may be varied. Hence a flexible means is available for varying the magnitude of the negative resistance and thus the amplitude of oscillation. By having the oscillation voltage regulate the bias on the control grid, additional amplitude control may be obtained.

Like the dynatron oscillator, this is essentially a low-power oscillator. It will generate sinusoidal oscillations of any frequency from the lower audio to approximately 60 Mc by simply changing the tuned circuit constants. Suitable pentodes for the transitron oscillator are the types 57, 58, 59, 89, 6C6, 6J7, and 6K7. In a properly designed oscillator, Brunetti reports that changes resulting from a 33 per cent change in direct screen-grid voltage may be kept within 10 parts in 10^6 and that, in general, the transition oscillator frequency stability may be compared with that of a crystal oscillator without temperature control.

18. Push-pull Negative-resistance Oscillator. A negative-resistance oscillator of low harmonic content and excellent frequency stability can be designed employing two tubes in push-pull as shown in Fig. 14c.[1] The action of this circuit is as follows. If the two tubes have identical characteristics and if the voltage between A and B is zero, the two plate currents are equal, and there is no current flowing between A and B. When an increment of voltage is applied between A and B, an increment of current will flow which will raise the plate voltage and lower the grid voltage of one of the tubes and lower the plate voltage and increase the grid voltage of the other tube. When this voltage is sufficiently small, the plate resistance and transconductance are substantially constant. If the amplification is large enough, the change in plate current exceeds the current flowing between A and B and is opposite in direction to the applied voltage. This results in a current flowing through the network between A and B which is opposite in direction to the applied voltage, and a negative resistance is obtained. When a parallel resonant circuit of high selectivity is connected between these terminals, sustained oscillations are developed.

The amplitude of oscillation may be readily controlled by means of the grid bias. When the reactance of the coupling condenser C_c is small in comparison with the grid resistance R_c at the lowest frequency of oscillation desired, which condition it is necessary and desirable to meet, the resonant circuit can be connected between either the two plates or the two grids.

A low-frequency oscillator having excellent frequency stability and low harmonic content with approximately uniform output over its a-f range has been designed by Reich.[2] This circuit employs a diode to give automatic

[1] REICH, H. J., *Proc. I.R.E.*, **25**, 1387, 1939; also TURNER, L. B., *Radio Rev.*, **1**, 317, 1920.

[2] *Ibid.*

amplitude control. The use of low amplitude of oscillation and push-pull amplification result in minimized harmonic content. Reich gives the power output of this oscillator as 0.06 watt.

19. Low-frequency Oscillators. At very low frequencies standard circuits become impractical. The condensers, and particularly inductances, required become very bulky and expensive. Accordingly certain rather special methods of obtaining low frequencies have been resorted to. The heterodyne oscillator is one of the best known. Circuits depending upon resistance and capacity in combination to determine the frequency are becoming increasingly important.

20. Beat-frequency Oscillators.[1] By beating together (heterodyning) two r-f voltages of slightly different frequencies, a-f energy may be generated. Oscillators operating on this principle are called *beat-frequency* or *heterodyne* oscillators. A block diagram of such an oscillator is shown in Fig. 15. The outputs of two r-f oscillators of slightly different frequencies are applied at the same time to a detector. In

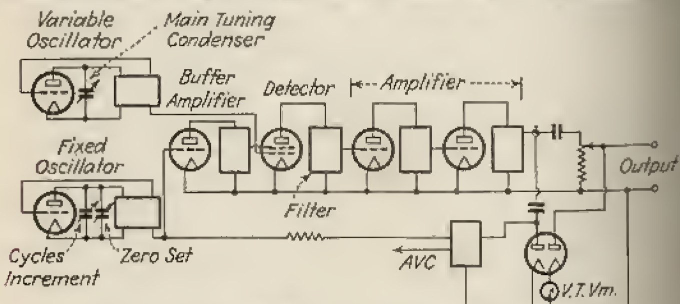


FIG. 15.—Block diagram of beat-frequency oscillator.

addition to the impressed frequencies, the output of the detector contains their sum and difference frequency. The filter, shown connected to the output of the detector, removes the fundamental radio frequencies and their sum and leaves only the difference frequency which may be amplified as desired.

Among the advantages of this type of a-f oscillator is the fact that the whole range of audio frequencies may be obtained by tuning a single dial. Another advantage is that the use of large coils and condensers, such as are employed in other types of a-f oscillators, is avoided. This results in lightness and compactness.

There are a number of special problems that arise in the design and construction of beat-frequency oscillators. One of these problems is to eliminate the tendency of the two oscillators to pull into synchronism when their frequency difference is small, *i.e.*, when low audio frequencies are being produced.

This tendency to interact may be avoided by proper shielding, careful arrangement of the component parts, proper use of decoupling resistors, choke coils, and by-pass condensers and by the use of special methods of

[1] An excellent discussion of beat-frequency oscillators is given by F. E. Terman, "Measurements in Radio Engineering," p. 298, McGraw-Hill Book Company, Inc., 1935.

coupling the oscillators to the detector. The most frequently used methods of coupling are (1) the use of a buffer amplifier between each oscillator and the detector, (2) the use of a balanced modulator circuit, or (3) the use of electron-coupled oscillators.

To avoid beats between harmonics generated by the r-f oscillators, an r-f filter is placed between the r-f oscillators and the detector. The fixed oscillator should have a smaller voltage output than that required by the variable-frequency oscillator in order that distortion of the output be reduced.

Higher order curvature of the detector characteristic produces additional distortion of the output, which may be prevented by the use of the balanced modulator. When square-law detectors are used, this type of distortion may be reduced by correct adjustment of bias and input voltage. Distortion produced by linear detectors may be reduced by making the output amplitude of one of the r-f oscillators small in comparison to that of the other.

The frequency stability of the output of beat-frequency oscillators is generally poor. This is because a very small percentage variation in frequency in the output of one of the r-f oscillators will result in a comparatively large percentage variation in the a-f output.

By making the two r-f oscillators as nearly identical as possible, they may be made to react similarly to variations in temperature, polarizing potential, etc., and thereby the effects of these quantities may be minimized. The r-f oscillators employed are usually stabilized by one of the methods discussed in Art. 3 or by the use of negative-resistance oscillators. To compensate for frequency drift in beat-frequency oscillators, a small trimming condenser is always provided which can be adjusted so that a particular point on the frequency calibration is correct. This point is obtained either by comparison of the output frequency with a standard frequency source or by using the zero-frequency point.

In the output circuit of the detector it is desirable to install a low-pass filter. This filter prevents the overloading of the a-f amplifier due to r-f voltages that may exist in the detector, and hence improves the output wave form.

The frequency at which the r-f oscillators operate is usually between 100 and 500 kc. At these higher frequencies the differences between the design constants of the fixed- and variable-frequency oscillators are less. This allows more nearly identical design, which, as pointed out above, leads to better frequency stability for the a-f output. Also the filter requirements are simplified by the use of the higher frequencies. On the other hand, the a-f stability is decreased as the r.f. is increased, and commercial design usually fixes 500 kc as the upper limit. The General Radio Company has produced an excellent heterodyne oscillator extending to 5 Mc. The fixed frequency in this range is 20 Mc.[1]

21. Special Audio Oscillator. A new type of oscillator particularly suitable for the generation of frequencies in the audio range has been suggested by Scott[2] and is shown in Fig. 16. He has described the use of the inverse feedback principle to obtain sharply selective circuits in which inductances are not necessary and "tuning" may be changed by varying resistances. These circuits may be varied over a wide range of frequencies while maintaining a selectivity curve which is a constant percentage function of the "tuned" frequency.

A low-power oscillator operating on the inverse feedback principle has been designed which has exceptionally pure wave form. By the use of a resistance-capacitance network, all frequencies except the frequency of oscillation are fed from the output of an amplifying system back into the input in such a way as to cancel the gain. Regeneration is introduced

[1] *Gen. Radio Experimenter*, January, 1939.
[2] SCOTT, H. H., *Proc. I.R.E.*, **26**, 226, 1938; *Gen. Radio Experimenter*, Vol. 13, No. 11, 1939.

into the circuit in sufficient amount to cause self-oscillation. This is controlled by the resistance-capacitance network, and hence no inductances or transformers are required in the oscillating circuit.

Figure 17 shows a block diagram which may be helpful in clarifying the action of the system. The circuit includes three separate sections. The section designated A is an amplifier and has substantially flat frequency response and negligible phase shift over the a-f range. The degeneration network, section B, balances to a sharp null at the frequency of oscillation.

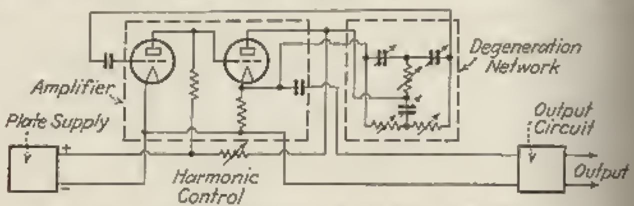


FIG. 16.—Resistance-capacity audio oscillator.

This provides full amplifier gain at this frequency, and gain at all other frequencies is substantially canceled. The regenerative feedback network, section C, is fed through a phase-reversing tube D to provide the proper regenerative action. Section C also has a flat frequency response and is adjusted to provide just sufficient regeneration to produce self-oscillation.

This oscillator covers the frequency range from 20 cycles to 15 kc. Operating under normal conditions approximately 0.25 watt of power may be obtained with less than 1 per cent distortion. With higher outputs the distortion is increased somewhat. This type of oscillator makes possible certain measurements which were previously impractical.

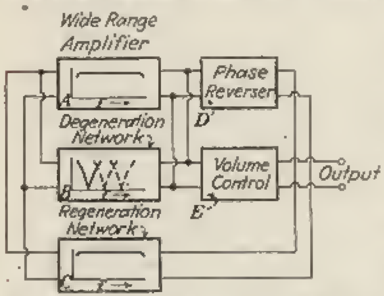


FIG. 17.—Block diagram of resistance-capacity audio oscillator.

Experimental models of this type of oscillator have been constructed which generate sinusoidal frequencies from 1 cycle per second to 200 kc. By the use of class B operation of the oscillator tubes in push-pull circuits, the power output can be extended appreciably.

22. Relaxation Oscillators. Relaxation oscillators are used for the generation of distinctly non-sinusoidal waves. In certain applications this type of oscillator has many advantages over the resonant circuit oscillators. Its output is extremely rich in harmonics, and its frequency, which is not very definitely fixed by the circuit elements, may be easily stabilized by the introduction of small voltages of harmonic or subharmonic frequency into the oscillating system. Relaxation oscillators are also comparatively inexpensive, simple, and compact and can conveniently be designed to cover a wide range of frequency.

[1] See REICH, H. J., "Theory and Application of Electron Tubes," McGraw-Hill Book Company, Inc., for a complete treatment of relaxation oscillators.

The process by which relaxation oscillations are produced involves the building up and breaking down of the energy stored in the electric field of a condenser or the magnetic field of an inductance. Various devices may be used to control this building-up and breaking-down action, such as glow or arc-discharge tubes or high-vacuum tubes.

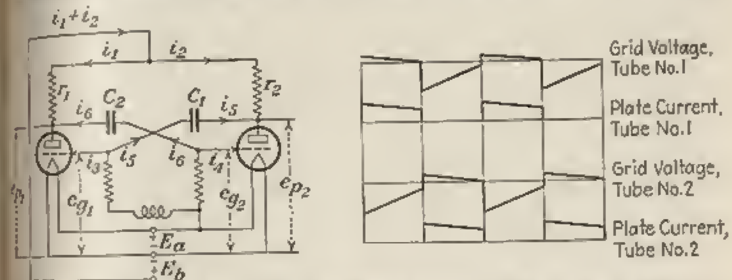


FIG. 18.—Fundamental circuit of the multivibrator and the voltage and current relations of the various branches.

Among the relaxation oscillators employing high-vacuum tubes is the multivibrator.[1] The multivibrator, which is most satisfactory for frequency conversion, was the first relaxation oscillator to be developed. Figure 18 shows the basic circuit with connections for introducing the control voltage. The voltage drop across any of the circuit elements

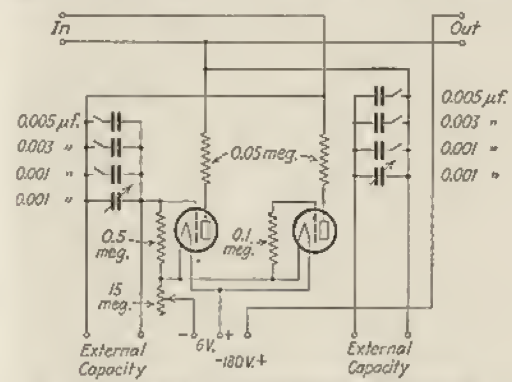


FIG. 19.—Generator of submultiple frequencies.

may be taken as the output voltage, and the frequency of oscillation may be controlled by variation of the resistances and capacitances and is approximately equal to $f = 1/(r_1C_1 + r_2C_2)$. When the circuit is symmetrical, the wave form of the grid and plate voltages of the condenser current is as shown in Fig. 18.

[1] ABRAHAM, H., and E. BLOCK, *Ann. Physik*, 12, 237, 1919.

For the generation of submultiple frequencies, a circuit connected as shown in Fig. 19 may be used. The output of a h-f oscillator is connected to the input terminals. At the output terminals, frequencies which are exact submultiples of the input frequency are obtained. Submultiple frequencies as low as one-fourteenth of the input frequency can easily be had. When the input and output terminals are short-circuited and a small coil connected between the low-potential ends of the grid resistances for coupling to an external circuit, frequencies as low as 1 cycle per 10 sec. and as high as 50,000 cps may be obtained.

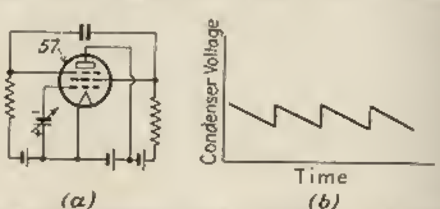


FIG. 20.—(a) Relaxation oscillator of Van der Pol. (b) Typical wave form of condenser voltage.

Another form of relaxation oscillator employing a high-vacuum tube was originally described by Van der Pol.[1] The circuit for this oscillator is shown in Fig. 20a, and the wave form of the condenser voltage is shown in Fig. 20b. This type of wave form, which is known as a *saw-toothed voltage* wave, is used in connection with cathode-ray oscillographs and cathode-ray television tubes.

Relaxation oscillators for generating saw-toothed wave forms are often designed using grid-controlled gas-filled triodes.[2] A property of these tubes that makes them suitable for this purpose is their so-called *trigger* action. If their grid potential is momentarily less than the cutoff value, positive ions are produced in the tube which neutralize the negative space charge of the electrons, as well as the controlling action of the grid. This results in a very rapid change in the plate resistance of the tube from a high value to a very low value. The time required to ionize and deionize the gas in the tube limits the frequency for which oscillators of this type can be built.[3]

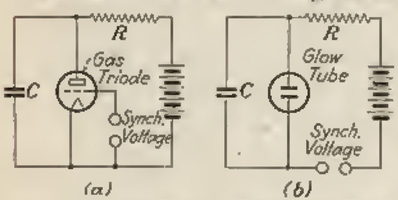


FIG. 21.—Relaxation oscillators using gas-filled tubes.

A basic circuit for a relaxation oscillator using a gas-filled tube is shown in Fig. 21a. The action of this circuit is as follows. The direct plate voltage charges the condenser C through the resistance R until

[1] VAN DER POL, B., *Phil. Mag.*, **2**, 978, 1926.
[2] For an excellent discussion and design data for relaxation oscillators using gas-filled tubes see "Measurements in Radio Engineering" by Terman, pp. 315 to 322, McGraw-Hill, Book Company, Inc., 1935.
[3] Oscillators of this type have been built to operate successfully as high as 20,000 cps

the critical starting potential of the tube is reached. At this potential the positive ions are produced, and the resistance of the tube falls to a very low value, discharging the condenser. When the plate voltage drops to a certain value, the plate resistance returns to its original high value, and the cycle is repeated. The value of the grid polarizing potential controls the critical plate potential at which ionization takes place.

Small alternating voltages may be introduced into the grid circuit for synchronizing purposes as shown in Fig. 21a. If a glow tube, *i.e.*, neon tube, relaxation oscillator is used, the synchronizing voltage may be introduced as shown in Fig. 21b.

A complete circuit diagram of a system suitable for producing saw-toothed wave forms for a cathode-ray oscilloscope is shown in Fig. 22. In this circuit a pentode is used to maintain a constant charging current. By varying its grid bias, the magnitude of the charging current may be controlled.

23. High-frequency Triode Oscillators. Almost all the commercial triodes now available may be depended upon to generate frequencies as high as 30 Mc without a serious loss of power or efficiency. A large number of them may be used at 50 Mc with full ratings and extended to 70 Mc at reduced ratings. Special triodes extend the frequency limit well into the micro-wave[1] region. When the familiar triode oscillator circuits are used for the generation of ultra-short waves, however, certain inherent limitations are brought out. These limitations arise from the interelectrode capacities, the inductance and capacitance of the lead-in wires, and the finite transit time of the electrons which constitute the current in the tube. Also special problems arise in the construction of the component parts of the circuit other than the tube.

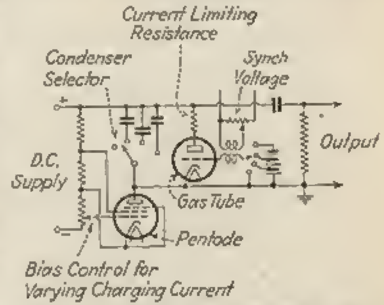


FIG. 22.—Relaxation oscillator employing a pentode to maintain a constant charging current.

In an effort to extend the range of the triode, a number of tubes have been designed to reduce the inherent limitations mentioned above. These new tubes have relatively low interelectrode capacitances, and the inductance and capacitance of the leads have been made very low. At the same time the interelectrode distances have been reduced, extending the range of operation appreciably, before the effect of the transit time enters. Tubes are now commercially available which will operaer at frequencies above 1,500 Mc. The power output at these frequencies is necessarily small due to close spacing and small size of the tube elements.

The resonant circuits employed in oscillators which operate in the u-h-f range[1] usually consist of resonant lines or special metal enclosures instead of the lumped inductance and capacitance used at the lower

[1] The ultra-short wave (u-h-f) range may be taken as lying between 10 meters and 1 cm. The region of the range below 1 meter is often referred to as the micro-wave range.

frequencies. When carefully designed, these resonant circuits may have selectivities ranging from 1,000 to 50,000 when radiation resistance is included.

Because of their high selectivity, resonant circuits of this type may be employed as frequency control elements. The stability of oscillators controlled in this way is comparable to that of crystal-controlled oscillators when the frequency is above 10 Mc and temperature control for the resonant circuit is provided. These oscillators may be used to drive power amplifiers for applications requiring large amounts of power.

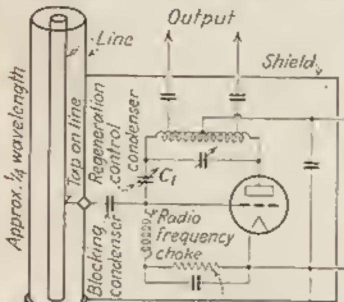


FIG. 23.—Oscillator employing a resonant line as its frequency-control element.

An oscillator circuit which employs a concentric line as a frequency-control element is shown in Fig. 23. This circuit is that of a standard tuned-plate tuned-grid feedback oscillator in which the grid resonant circuit is replaced by the concentric line. For best stability, the line should be loosely coupled to the grid circuit. This is accomplished by making the grid connection at a point comparatively close to the shorted end of the line. By proper adjustment of the regeneration control C_1, the phase of the grid excitation may be advanced so as to compensate for the phase lag of the electron current in the tube. The length of the connections between the tube and the resonant circuits must be small compared with the wave length at which the system is operating. This condition may be very difficult to meet at extremely high frequencies, and special circuits have been designed which help avoid this difficulty.

One of these circuits has been designed by Peterson.[1] His system employs a resonant circuit consisting of an outer containing cylinder with a cylindrical piston-shaped insert. The Q of this oscillating circuit, in the frequency range of 60 to 140 Mc in which they have been built, is approximately 2,000. Frequency stabilities of the order of 5 parts in 10^6 for a 50 per cent change in plate voltage have been obtained. One of the chief advantages, in addition to its excellent stability, is that difficulties arising from tube connections are greatly minimized (see Fig. 24). Also the size of the resonant circuit is only a fraction of its equivalent concentric line. When a continuously variable oscillator is

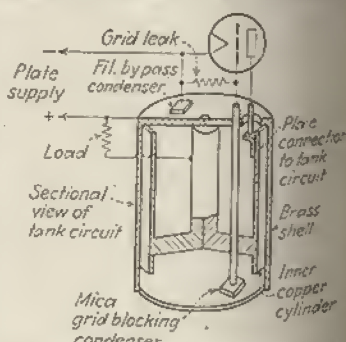


FIG. 24.—Ultra-high-frequency oscillator of Peterson.

[1] PETERSON, *Gen. Radio Experimenter* **12**, October, 1937; *Communications*, **17**, 26–28, 1937.

required, however, other arrangements must be used, for only slight variations in frequency can be obtained with this design.

A continuously adjustable stabilized oscillator has been designed by Barrow[1] for the frequency range from 70 to 700 Mc. The oscillator circuit consists of a coaxial line that is easily and rapidly adjustable over the entire frequency range. Among other things, it affords excellent shielding, mechanical ruggedness, and a coaxial line output connection. Several watts output are obtained over the entire frequency range. Both filament leads are tuned in addition to the tank. The connections are shown in Fig. 25. At frequencies below 300 Mc the stability is roughly 100 parts in 10^6 and decreases with increased frequency, becoming very poor near the limit of oscillation of the tube.

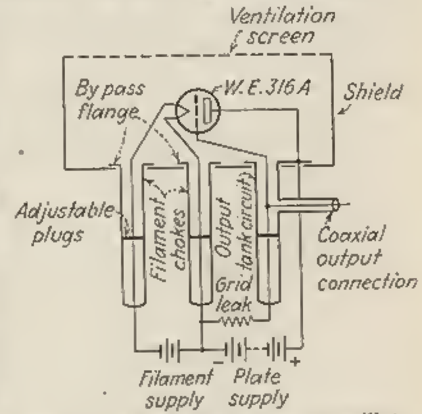


FIG. 25.—Ultra-high-frequency oscillator of Barrow.

Special triodes in which the plate and grid leads provide support for the electrodes and extend through the bulb are especially useful for u-h-f work. A tube of this type having a rated output of 1.5 watts at 1,500 Mc is available commercially. A pair of such tubes has been built into a special oscillator, having a continuous range. This oscillator, designed by King,[2] is particularly suitable for parallel line measurements at ultra-high frequencies (see Fig. 26). It consists essentially of a rectangle of parallel conductors which may be bridged by blocking condensers. The frequency depends on the dimensions of the circuit BAB' in the figure. For the highest frequencies the condenser A must be used, and condensers B and B' must be moved up as close to the triodes as possible. For lower frequencies A need not be used. The oscillator is coupled to parallel lines by placing it below the lines.

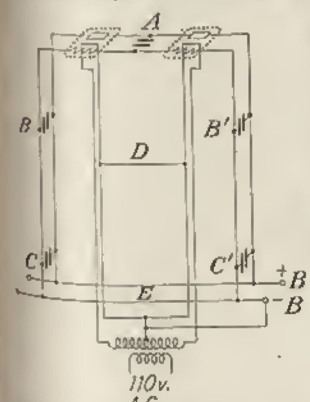


FIG. 26.—Ultra-high-frequency oscillator of King.

Circuits employing parallel lines are often used for the generation of u-h-f waves. Representative circuits are shown in Fig. 27. Although these circuits are comparatively simple to construct, the tube is connected at a high impedance point and the frequency stability is poor.

[1] BARROW, W. L., *Rev. Sci. Inst.*, **9**, 170, 1938.
[2] KING, R., *Rev. Sci. Inst.* (submitted), February, 1940.

The chokes that appear in the filament leads of the circuits of Fig. 27 are made necessary for u-h-f operation because the filament with its leads may often be a considerable portion of a wave length and thereby prevent normal operation of the oscillator. A method which is preferred to the use of choke coils is the provision for tuning of the filament to

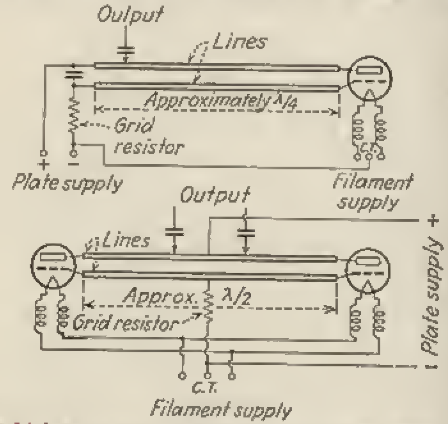


Fig. 27.—Ultra-high-frequency oscillators employing parallel lines as circuit elements.

ground circuit. The use of adjustable concentric lines of approximately one-fourth wave length for each filament lead is probably the most satisfactory method (see Fig. 28). It is desirable at u-h-f operation to avoid the use of dielectric material as much as possible and to confine that which is necessary for mounting the circuit elements to points of low r-f potential.

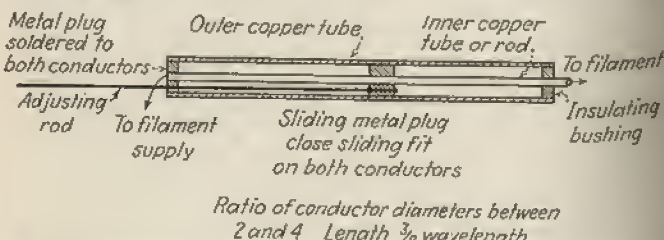


Fig. 28.—Cross section of an adjustable concentric line suitable for use as a filament choke.

24. The Magnetron Oscillator.[1] This type of oscillator is used for the generation of u-h-f waves. The magnetron is essentially a diode with coaxial cylindrical electrodes which is placed in a magnetic field so that lines of electromagnetic force are approximately parallel to the axes

[1] Hull, A. W., *Jour. A.I.E.E.*, **40**, 715, 1921; *Trans. A.I.E.E.*, **42**, 915, 1923; Elder, F. R., *Proc. I.R.E.*, **13**, 159, 1935; Megaw, E. C. S., *Jour. I.E.E.* (*British*), April, 1933; Kilgore, G. R., *Proc. I.R.E.*, **24**, 1140, 1936.

of the diode electrodes. When the intensity of the magnetic field exceeds some critical value, the electrons will travel in orbits within the anode, and very few of them will reach the plate. When the intensity of the field is less than the critical value, all the electrons will reach the plate. Hence the magnetic field can be used to control the anode current in a way similar to the grid in a triode. Originally the magnetron oscillator was employed this way, but its action was restricted to long wave lengths because of the inductance of the coils carrying the alternating field current, and it did not compete successfully with the triode oscillator. As a generator of u-h-f waves, however, the magnetron is superior to the triode. In its simplest form the modern magnetron has

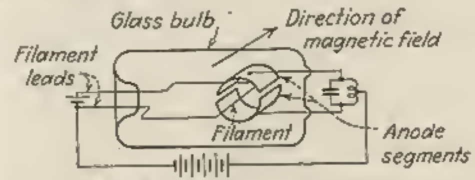


Fig. 29.—Magnetron oscillator.

its cylindrical plate divided into two or more equal segments separated by narrow gaps, as shown in Fig. 29.

There are two distinct methods of producing oscillation with the magnetron tube. These are (1) the negative-resistance method and (2) the electron-resonance method. With either of these a constant magnetic field is used to control the direction rather than the magnitude of the electron current.

With the negative-resistance method, often referred to as the *dynatron method of operation*, a negative resistance is developed which arises from the deflection of the electrons by the magnetic field. This is used to develop oscillations whose frequency is substantially equal to the resonant frequency of the tuned circuit which is connected to the magnetron as shown in Fig. 29. With the electron-resonance method the frequency developed is approximately equal to the rotation frequency of the electrons about the lines of magnetic force, and the oscillations are maintained by the transformation of part of the kinetic energy of the moving electron into potential energy stored in the oscillating circuit. The wave length of oscillations is given approximately by

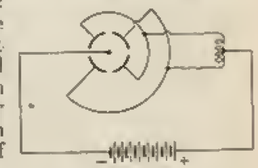


Fig. 30.—The four-segment magnetron oscillator.

$$\lambda \cong \frac{12{,}000}{H}$$

where H = the field strength in gauss.

The highest frequencies are produced by the electronic oscillations of the magnetron oscillator, and wave lengths as low as a fraction of a centimeter have been generated. The efficiency of this method of producing oscillations is quite low, however, being of the order of several per cent. Higher efficiencies and power output are obtained with the negative resistance or dynatron oscillations of the magnetron. With the two-segment anode the efficiency is of the order of 40 to 60 per cent and the wave length between approximately 75 cm and 10 meters. With the four-segment anode, connected as shown in Fig. 30, the wave length is reduced to between 25 cm and 3 meters and the efficiency is also reduced to between 30 to 50 per cent approximately.

25. Barkhausen-Kurtz Oscillator.[1] This oscillator is used for the generation of u-h-f oscillations. Wave lengths as low as a few centimeters have been obtained. The power output is, in general, low, reaching a maximum of approximately 10 watts; the efficiency rarely exceeds a few per cent.

The Barkhausen-Kurtz oscillator operates with the grid at a high positive potential while the plate is usually at a small negative potential. The connections are shown in Fig. 31. The action taking place in this type of oscillator may be explained in terms of a variation in the electric field about the grid due to a periodic motion of electron clouds. The natural frequency of oscillation of the electron clouds is extremely high and is determined by the potential of the grid. The frequency observed at the electrodes of the tube depends upon this natural frequency of the electron clouds and upon the natural frequency of a coupled circuit (which may consist of the tube electrodes only). If the external circuit is in resonance or nearly in resonance, it may greatly affect the observed frequency, as is usual with closely coupled circuits. The electronic oscillation may be considered as constituting the primary circuit. In this case the oscillations are sometimes called Gill-Morell. According to the theory of coupled circuits, as applied by Wundt[1] to the Barkhausen oscillator, several coupling frequencies should be simultaneously possible, depending upon the damping of the coupled circuits. This has recently been verified by King,[2] who observed as many as three coupling frequencies maintained simultaneously.

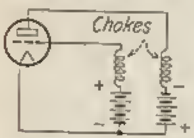


FIG. 31.—Barkhausen-Kurtz oscillator.

It has been found that tubes, in which Barkhausen-Kurtz oscillations may be produced, usually have cylindrical electrodes, and Hollman[3] has found that the ratio of plate to grid radii must be greater than 2 and less than 5. Optimum values for this ratio are between 2.5 and 3. The wave length of the oscillating electron clouds is given approximately by the relation

$$\lambda^2 E_g = K$$

where λ = the wave length
E_g = the grid potential
K = a constant depending upon the geometry of the tube.

26. Klystron Oscillator.[4] The klystron oscillator is at this time of writing still largely in the experimental stage. It has many promising

[1] WUNDT, R., *Hochfrequenztechn.* **36**, 133, 1930.
[2] KING, R., paper submitted for publication, February, 1940.
[3] HOLLMAN, H. E., "Physic und Technik der Ultrakurzen Wellen," Julius Springer, Berlin, 1936.
[4] References on klystron: HANSEN, W. W., and R. D. RECHTMEYER, *J. Applied Phys.*, **10**, 189, 1939; HANSEN, W. W., *J. Applied Phys.*, **9**, 654, 1938; VARIAN, R. H., and S. F. VARIAN, *J. Applied Phys.*, **10**, 321, 1939; WEBSTER, D. L., *J. Applied Phys.*, **10**, 501, 1939; HEIL, A. ARSENJEWA, and O. HEIL, *Zeit. Physik*, **95**, 752–762, 1935; BRUCHE, E., and A. RECKNAGEL, *Zeit. Physik*, **108**, 459–482, 1938; HAHN, W. C., and G. F. METCALF,

features and is being developed in an effort to obtain considerable amounts of power at wave lengths of 5 to 20 cm. It employs a new type of tube known as the *velocity-modulated tube.* In the conventional grid-anode tubes, where the electron stream is controlled by a grid, the time taken for the electrons to go from the cathode through the grid to the plate limits their use at these ultra-high frequencies. The velocity-modulated tube, on the other hand, utilizes this transit time phenomenon in such a way as to obtain the h-f waves by means of new types of structure and modes of operation.

The action of the velocity-modulated tubes is substantially as follows: An electron beam of constant current and speed is passed through a pair of grids which may be in the form of hollow cylinders. Between these grids is applied an oscillating field, parallel to the electron stream and of sufficient strength to change the velocities of the cathode rays by an appreciable fraction of their initial speed. After the electrons in the beam have passed through these grids, the electrons with increased speeds begin to overtake those with decreased speeds that were ahead of them. This produces what has been termed *velocity modulation of the electron beam* and groups the electrons into bunches separated by relatively empty space. These bunches of large density pass into another structure where they induce the output current. After leaving the output structure, the electrons are collected upon an anode held at a fixed positive potential.

By utilizing and properly adjusting the electron transit time, tubes have been produced which will give efficiencies of approximately 30 per cent. When a certain amount of the output power is fed back into the input, this tube acts as an oscillator. Engineers predict that hundreds of watts of u-h-f energy in the 5- to 20-cm region may be produced in this way.

In this type of oscillator, special resonators are used which have extremely high selectivity. These resonators, called rhumbatrons, are metal vessels practically closed.

27. Mechanical-electronic Oscillators. Mechanical-electronic oscillators are those which employ in combination both vacuum-tube circuits and mechanical-rotating members. The electrostatic audio generator developed by Kurtz and Larsen[1] falls into this class. It consists of a number of variable condensers of the rotary type which are driven at a constant speed. When a direct voltage is connected through resistors to each of these condensers, the charging current of the different condensers is varied. If the plates of the condensers are designed so that the charging currents are sinusoidal, a sinusoidal voltage will be developed across each of the series resistors. These voltages may be applied at the same time to the input of an amplifier. This generator was designed to produce a fundamental and 15 harmonic voltages simultaneously. The phase of any sinusoidal voltage with respect to the others is easily adjusted by shifting the position of any one of the stationary plates. To vary the amplitude of any of the voltages, it is necessary only to vary the particular applied direct voltage. A system such as this one is very useful in the study of the effects of changes in amplitude and phase of a complex sound on the ear.

The photo-audio generator developed by Schaffer and Lubszynsky[2] falls also into this class of oscillators. It consists of a system in which

Proc. I.R.E., **27**, 106–116, 1939; HAHN, W. C., *Gen. Elec. Rev.*, **42**, 258, 1939; RAMO, SIMON, *Proc. I.R.E.*, **27**, 757, 1939.
[1] KURTZ, E. B., and M. J. LARSEN, *Elec. Eng.*, **54**, 950, 1935.
[2] SCHAFFER, W., and G. LUBSZYNSKY, *Proc. I.R.E.*, **19**, 1242, 1931; **20**, 363 1932.

a beam of light falling on a photoelectric cell is interrupted by a perforated rotating disk. When a light aperture and the size and shape of the holes in the disk are correctly designed, the voltage developed by the photocell will be very nearly sinusoidal. The output of the cell can be suitably amplified and the frequency of the oscillator read by a tachometer applied to the driving motor.

Many other oscillators of this type have been developed.

28. Tuning-fork Oscillators. For oscillators of low to medium frequency the tuning fork provides an excellent resonator. The range from 100 to 10,000 cps is readily covered. Simplest and least precise are the contact-driven forks which are capable of supplying considerable power output of approximately square wave form. The single-button microphone drive gives a much purer wave and more constant frequency at the expense of power output. The double-button microphone drive gives a still purer wave, better stability, and greater output.

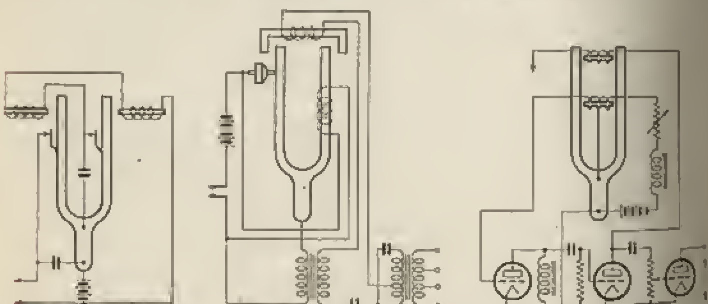


FIG. 32.—Circuits for various types of tuning-fork oscillators (hummers).

The free tuning fork, which is driven by one magnetic coil and excites another, is capable of high frequency stability. With a suitable circuit and a fork of special material a stability of a few parts per million is obtained for relatively long times without benefit of voltage or temperature control.

29. Oscillator Automatic-frequency Control. It may be desirable in practice to control the frequency generated by an oscillator with a small change in voltage. Special circuits which have been designed to do this are shown in Figs. 33*a* and *b*.[1] In Fig. 33*a* the tuned circuit of an oscillator is shunted by a condenser C and vacuum tube V_1. Since the plate resistance of this tube depends upon the magnitude of the control voltage, the effective reactance of the combination is varied by variations in the control voltage. If the plate resistance is adjusted so that it is numerically equal to the reactance of C, the effective resistance does not change with a change in effective reactance. In this way a voltage applied to the control grid will control the frequency of oscillation without affecting the amplitude.

[1] TRAVIS, C., *Proc. I.R.E.*, **23**, 1125, 1935.
For variants of the basic circuit see:
FREEMAN, R. L., *Electronics*, p. 20, November, 1936; FOSTER, D. E., and S. W. SEELEY, *Proc. I.R.E.*, **25**, 289, 1937.

Figure 33*b* shows another type of control circuit for varying the frequency of an oscillator. Here the plate circuit of the control tube V_1 shunts the resonant circuit of the oscillator. The voltage drop in the resistance R is applied to the grid of the control tube. This voltage is approximately 90 deg. out of phase with the resonant circuit voltage, and hence the plate circuit of the control tube affects the tank circuit in the same way as a reactance. Variations in this effective reactance are obtained by changing the bias on the control tube.

30. Electron-coupled Oscillators. It has been pointed out that the frequency of oscillation will be affected by changes in load unless the

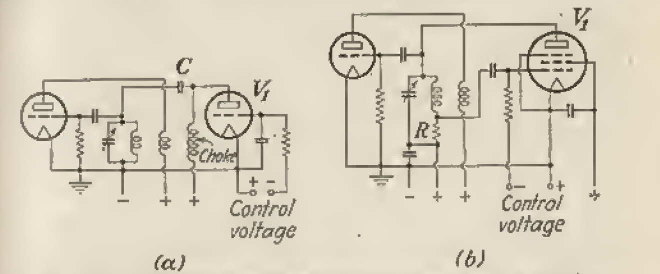


FIG. 33.—Oscillators having automatic-frequency control.

current in the resonant circuit inductance is not changed when the load is varied. The use of buffer amplifiers between the oscillator and the load aids in shielding the former from load variations.

Dow[1] has developed another method making the frequency of oscillation independent of load variations. The method employs electron coupling between the oscillator and the load. This consists of a Hartley oscillator in which the screen grid serves as anode while the plate serves only as an output electrode (see Fig. 34). The screen grid is effectively grounded to alternating currents while the cathode is at an alternating potential above ground. This prevents the load impedance in the plate circuit from reacting back on the oscillator. At the same time the electrons that pass through the screen are attracted by the more positive plate. This results in the plate current having an alternating component which is of the same frequency as the oscillator frequency. Hence energy is delivered to the load through the electron stream, and at the same time the oscillator is effectively shielded from the load.

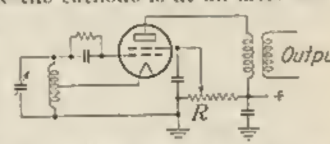


FIG. 34.—Electron-coupled oscillator circuit.

The frequency of oscillation with this type of oscillator can be made independent of supply voltage variations by properly choosing the ratio of screen grid to plate potential. This is possible because there is always some value of this ratio for which the frequency is independent of the applied voltage. By adjusting the position of the tap on R until the frequency becomes independent of applied voltage, a high degree of

[1] DOW, J. B., *Proc. I.R.E.*, **19**, 2095, 1931.

frequency stability may be obtained. It may be necessary to change the position of this tap if the oscillator frequency is varied over wide limits.

When the oscillator is operated normally, the tube is biased beyond cutoff, and the plate current flows in pulses that are very rich in harmonic content. Since the frequency stability is very high and the output extremely rich in harmonics, this oscillator is very satisfactory for use in heterodyne frequency meters. There are many possible ways in which the principle of electron coupling may be used to great advantage.

31. Power-oscillator Design. In applications which require appreciable power output and for which a high degree of frequency stability is not necessary, feedback oscillators may be used. In these oscillators the tube operates as in a class C amplifier, and oscillators adjusted in this way are referred to as class C oscillators.

The design of a class C oscillator may be conducted along the same lines as for a class C amplifier. There are two principal points of difference: (1) the power required to supply the energy for the grid circuit must be obtained from the plate-circuit power supply since this, with the exception of the filament supply, is the only source of power in the circuit; and (2) the oscillating circuit must contain sufficient stored energy to meet the grid-circuit requirements. The first step in design is to determine the correct operating conditions for the tube. These may be determined by various methods, some of which consist of graphical integrations of the current waves that are obtained from the static characteristic curves of the tube.[1] These methods require that the complete static characteristic curves be available far out in the region of positive grid potential. To obtain these curves special experimental techniques have been devised.[2] Other methods consist of analytical integration of (1) simple expressions which are assumed to approximate the wave form of the current pulses or (2) approximate analytical expressions for the static characteristic curves.[3] Among these methods of precalculation, some are extremely laborious but accurate, while others are rapid but inaccurate. The methods due to Chaffee[4] and to Terman and Roake[5] are recommended as sufficiently accurate for engineering purposes, and at the same time they are reasonably rapid. In addition to these methods of precalculation, the operating conditions for the tube may be determined by direct test.

The results obtained by any one of these various methods will be subject to certain limitations which are always imposed on the operation of the tube. These limitations are designed to preserve reasonable life and to prevent sudden failure of the tube. Electrode dissipations are

[1] Prince, D. C., *Proc. I.R.E.*, **11**, 275, 405, 527, 1923; Kilgour, C. E., *Proc. I.R.E.* **19**, 42, 1931; Mouromtseff, I. E., *Proc. I.R.E.*, **20**, 783, 1932; Mouromtseff, I. E., and Kozanowski, H. N., *Proc. I.R.E.*, **22**, 1090, 1934; *Proc. I.R.E.*, **23**, 752, 1935; Wagener, W. G., *Proc. I.R.E.*, **25**, 47, 1937.

[2] Kozanowski, H. N., and I. E. Mouromtseff, *Proc. I.R.E.*, **21**, 1082, 1933; Chaffee, E. L., *Electronics*, June, 1938.

[3] Fay, C. E., *Proc. I.R.E.*, **20**, 548, 1932; Everitt, W. L., *Proc. I.R.E.*, **22**, 152, 1934; Terman, F. E., and J. H. Ferns, *Proc. I.R.E.*, **22**, 359, 1934; Miller, B. F., *Proc. I.R.E.* **23**, 496, 1935; Babits, V. A., *L'Onde electrique*, **14**, 668, 1935; Everitt, W. L., *Proc. I.R.E.*, **24**, 305, 1936; Terman, F. E., and W. C. Roake, *Proc. I.R.E.*, **24**, 620, 1936; Wagener, W. G., *Proc. I.R.E.*, **25**, 47, 1937.

[4] Chaffee, E. L., *Jour. Applied Phys.* **9**, 471, 1938.

[5] Terman, F. E., and W. C. Roake, *Proc. I.R.E.*, **24**, 620, 1936. A method for obtaining the optimum conditions of operation for various applications has been developed which is extremely accurate but not too involved. This method will appear in the revision of E. L. Chaffee's "Thermionic Vacuum Tubes," in preparation.

limited to prevent liberation of gas or injury to the tube structure due to melting or warping. The maximum polarizing potential is limited to prevent flashover due to gas or cold emission, breakdown of insulation, or puncture of the tube due to stray beams of electrons. Also the grid dissipation must be kept small enough to avoid appreciable primary emission from the grid, and the maximum instantaneous current flowing in the tube must be limited to preserve reasonable filament life. When the operating conditions are selected with due regard to these limitations and to the conditions of power output and efficiency desired, they may be used for the design of any one of the oscillator circuits shown in Fig. 1.

These operating conditions are as follows:

E_B The direct plate voltage.
E_c The direct grid voltage.
E_p The r-m-s value of the fundamental component of the plate alternating voltage.
E_g The r-m-s value of the fundamental component of the grid alternating voltage.
I_B The average value of the direct plate current.
I_c The average value of the direct grid current.
I_p The r-m-s value of the fundamental component of the plate a.c.
I_g The r-m-s value of the fundamental component of the grid a.c.

32. Power Relations in Class C Oscillators. The d-c power supplied to the oscillator circuit from the source E_B is

$$P_{\text{input}} = I_B E_B$$

The power output to the tank circuit at fundamental frequency is

$$P_{\text{tank}} = E_p I_p$$

This power must supply the driving power, the load, and the losses in the tank circuit. This power may also be expressed as

$$P_{\text{tank}} = I_L^2 R_L$$

where R_L represents, in addition to the inherent resistance of the total tank inductance, the resistance reflected into the tank circuit by the load requirement of the grid circuit and by the load itself. I_L is the circulating current in the oscillating circuit.

The power lost at the plate is

$$P_{\text{plate}} = I_B E_B - E_p I_p$$

The driving power required by the tube is

$$P_{\text{driving}} = E_g I_g$$

This power supplies the power delivered to the grid resister to maintain the bias voltage

$$P_{\text{grid resister}} = I_c E_c$$

and the power lost at the grid. Therefore the grid loss is

$$P_{\text{grid}} = E_g I_g - I_c E_c$$

The power available for output is then given approximately by

$$P_{\text{output}} = E_p I_p - E_g I_g$$

Actually the useful power which may be obtained from the oscillator will be less than is given by this expression, by the inherent losses in the circuit elements. However, in a well-designed oscillator these may be kept small, and the power output will be very nearly that given above; hence the efficiency of the oscillator may be expressed as

$$\text{Eff.} = \frac{E_p I_p - E_g I_g}{I_B E_B}$$

This is not the over-all efficiency. An expression for the over-all efficiency of an oscillator may include the power required for the filament and the losses incurred in obtaining the high-voltage plate supply.

33. Design of a Hartley Oscillator. The methods illustrated here are applicable to any of the other feedback oscillators. The problem is to determine the values of the circuit elements in such a way that the proper voltages (which have been obtained by any one of the various methods discussed above) will be applied to the tube. It has been found that in general the alternating plate and grid voltages should be 180 deg. out of phase for correct operation. Most of the methods for obtaining the operating conditions assume that this phase condition will be met. Slight variations in phase of a few degrees have only a small effect on the performance of the oscillator.

In the Hartley circuit the voltages developed across the tapped inductance are used as the alternating components of the plate and grid voltage. The total alternating voltage across the tank circuit is the sum of E_p and E_g. The filament tap is adjusted so that the ratio of the plate and grid voltage is E_p/E_g. It has been found experimentally that the effective selectivity Q of the tank should be greater than 12.5, approximately. Higher values of Q increase the stability and lower the harmonic content. The tank-circuit inductance for any frequency may be found by use of the relation

$$L = \frac{(E_{p1} + E_{g1})^2}{2\pi f Q P_{\text{tank}}}$$

and the tank-circuit capacity by

$$C = \frac{QP}{2\pi f(E_p + E_g)^2} \qquad C = \frac{QP_{\text{tank}}}{2\pi f(E_{p1} + E_{g1})^2}$$

where f is the required frequency of oscillation.

The grid-bias voltage required is given by E_c, and with an average grid current of I_c the grid resistance is

$$R_c = \frac{E_c}{I_c}$$

The value of the grid condenser capacity C_c should be large enough to act as an effective short circuit for the grid resistance R_c at the frequency of operation. It should not, however, be so large as to cause intermittent oscillations, as discussed in Art. 14 of this section.

The plate blocking condenser C_B should be large enough so that the reactive voltage developed across it will be small in comparison to E_p. The inductance of the shunt feed choke should be large in comparison with the inductance of the portion of the tank circuit which it effectively shunts. In the Hartley oscillator this blocking condenser and choke can be designed in such a way as to help correct for variations in phase of the alternating plate and grid voltages from the 180-deg. position.

References

Oscillators, General:

Arguimbau, L. B.: *Proc. I.R.E.*, **21**, January, 1933.
Benioff, H.: *Proc. I.R.E.*, **19**, July, 1931.
Brunetti, C.: *Proc. I.R.E.*, **25**, 1595, 1937.
Dow, J. B.: *Proc. I.R.E.*, **15**, May, 1927; **19**, December, 1931.
Eller, K. B.: *Proc. I.R.E.*, **16**, December, 1928.
Groszkowski, Janusz: *Proc. I.R.E.*, **21**, 958, 1933; **22**, February, 1934.

Hamburger, F., Jr.: *Proc. I.R.E.*, **22**, January, 1934.
Herold, E. W.: *Proc. I.R.E.*, **23**, 1201, 1935. (Contains large bibliography on negative-resistance oscillators.)
Horton, J. W.: *Bell System Tech. Jour.*, **3**, 508, July, 1924.
Jimbo, S.: *Proc. I.R.E.*, **17**, November, 1929.
King, R. W.: *Bell System Tech. Jour.*, **2**, October, 1923.
Le Corbeiller, P.: *I.E.E.*, *Wireless Sec.*, **11**, 292, 1936.
Model, S. J.: *Proc. I.R.E.*, **21**, December, 1933.
Roosenstein, H. O.: *Proc. I.R.E.*, **19**, October, 1931.
Stone, J. S.: *Proc. I.R.E.*, **19**, August, 1931.
van der Pol, Balth: *Proc. I.R.E.*, **22**, September, 1934.
Vecchiacchi, F.: *Proc. I.R.E.*, **19**, May, 1931.

Frequency-control Crystals:

Boella, M.: *Proc. I.R.E.*, **19**, July, 1931.
Cady, W. G.: *Proc. I.R.E.*, **10**, April, 1922.
Clayton, J. M.: *QST*, September, 1926.
Colwell, R. C.: *Proc. I.R.E.*, **20**, May, 1932.
Dye, D. W.: *Proc. Phys. Soc. (London)*, **38**, August, 1926.
Harrison, J. R.: *Proc. I.R.E.*, December, 1927; **16**, November, 1928.
Heaton, V. E., and E. G. Lapham: *Proc. I.R.E.*, **20**, February, 1932.
Hight, S. C., and G. W. Willard: *Proc. I.R.E.*, **25**, May, 1937.
Horton, J. W., and W. A. Marrison: *Proc. I.R.E.*, **16**, February, 1928.
Koga, I.: *Proc. I.R.E.*, **19**, June, 1931.
Lapham, E. G.: *Proc. I.R.E.*, **22**, January, 1934.
MacKinnon, K. A.: *Proc. I.R.E.*, **20**, November, 1932.
Mason, W. P.: *Bell System Tech. Jour.*, **19**, January, 1940.
Meahl, H. R.: *Proc. I.R.E.*, **22**, June, 1934.
Mueller, E. L.: *QST*, May, 1927.
Pierce, G. W.: *Proc. Amer. Acad. Arts Sci.*, **59**, October, 1923.
Terry, E. M.: *Proc. I.R.E.*, **16**, November, 1928.
Van Dyke, K. S.: *Proc. I.R.E.*, **16**, June, 1928.
Vigoureux, P.: *Phil. Mag.*, December, 1928.
Watts, W. A.: *QST*, January, 1928.
Wheeler, L. P.: *Proc. I.R.E.*, **19**, April, 1931.
Williams, N. H.: *Proc. I.R.E.*, **21**, July, 1933.
Worrall, R. H., and R. B. Owens: *Proc. I.R.E.*, **16**, June, 1928.
Wright, J. W.: *Proc. I.R.E.*, **17**, January, 1929.

High-frequency Oscillators (including magnetrons, electron coupling, etc.):

Barkhausen and Kurz: *Physik. Zeits.*, **21**, 1920.
Dytrt, L. F.: *Proc. I.R.E.*, **23**, March, 1935.
Gill and Morell: *Phil. Mag.*, **44**, 1922.
Hoag, J. B.: *Proc. I.R.E.*, **21**, August, 1933.
Hollmann, H. E.: *Proc. I.R.E.*, **17**, February, 1929.
Karplus: *Gen. Radio Experimenter*, **5**, May, 1931.
Kilgore, G. R.: *Proc. I.R.E.*, **20**, November, 1932.
Kozanowski, H. N.: *Proc. I.R.E.*, **20**, June, 1932.
Lorber, C. W.: *Electronics*, November, 1930.
McNamara, F. T.: *Proc. I.R.E.*, **22**, August, 1934.
Megaw, E. C. S.: *Proc. I.R.E.*, **21**, December, 1933.
Moore, W. H.: *Proc. I.R.E.*, **22**, August, 1934.
Mouromtseff, I. E., and H. V. Noble: *Proc. I.R.E.*, **20**, August, 1932.
Okabe, Kinjiro: *Proc. I.R.E.*, **21**, November, 1933.
Thompson, B. J., and P. D. Zottu: *Proc. I.R.E.*, **22**, December, 1934.
Wenstrom, W. H.: *Proc. I.R.E.*, **20**, January, 1932.
White, W. C.: *Electronics*, April, 1930.

SECTION 10

MODULATION AND DETECTION

By L. F. Curtis[1]

TYPES OF MODULATION

1. Modulated Waves. A *modulated wave* is a periodic wave of which the amplitude, frequency, or phase is varied in accordance with a signal. *Modulation* is the process by which this variation is accomplished. *Demodulation* or *detection* is the process by which the original signal is recovered from the modulated wave.

The unmodulated component of the original wave is a *carrier wave* or, more broadly, a *carrier*. The frequency of the carrier usually is much greater than the highest frequency component in the original signal.

The *modulating wave* or *signal wave* applied to the carrier may be a direct representation of the original signal, or it may be a wave of a different carrier frequency previously modulated by the original signal. In this case the process is called *double modulation*, and the modulation appears in new frequency groups associated with carrier frequencies equal to the sum and difference of the signal-carrier and fixed-carrier frequencies. Double modulation accompanied by the selection of one of the new carrier frequencies (*intermediate frequency*—i.f.) to which the signal modulation has been transferred is often called *frequency conversion.*

In tubes and circuits the waves are of current in, or of voltage across, circuit elements. A modulated wave of current is expressed by

$$i = I \cos (\omega t + \phi) \tag{1}$$

where I = amplitude
$\omega/2\pi$ = frequency of the carrier
t = time
ϕ = relative phase.

The signal may be imparted to the carrier either by a variation of the amplitude or by a variation of the phase as a function of time. Variations in phase are accompanied by simultaneous variations in frequency, since the frequency is $1/2\pi$ times the derivative of the phase with respect to time. Simultaneous variations of phase and amplitude during modulation usually lead to distortion.

Velocity modulation is a process in which the velocity of electrons in transit in a special tube is controlled by a signal. The wave at the output terminals of the tube is modulated in amplitude. With presently available commercial devices and circuits, this process is usable only at extremely high carrier frequencies.

[1] Hazeltine Service Corporation.

2. Amplitude Modulation (A.M.). A current wave amplitude modulated by a single signal component of frequency $a/2\pi$ may be expressed by rewriting Eq. (1) as

$$i = I_0(1 + m \cos at) \cos (\omega t + \theta) \tag{2}$$

where I_0 = amplitude of the carrier
m = relative variation in amplitude at the signal frequency
θ = constant phase of the carrier.

The value of m is called the *degree of modulation* or *modulation factor.* When multiplied by 100, it is the *percentage of modulation.*

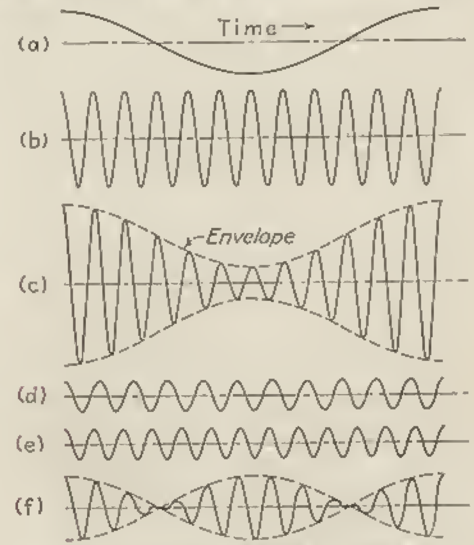


Fig. 1.—Components in a-m waves.

The phase of the carrier may be neglected unless the current is to be combined with other currents of the same frequency. Neglecting θ and making I_0 unity for convenience, Eq. (2) may be expanded to

$$i = \cos \omega t + \frac{m}{2} \cos (\omega + a)t + \frac{m}{2} \cos (\omega - a)t \tag{3}$$

The modulated wave contains, in addition to the carrier, two independent periodic waves spaced therefrom in frequency by the modulating frequency. These are *modulation side frequencies.*

These waves are illustrated in Fig. 1, in which a is the signal, b the unmodulated carrier, c the complete modulated wave, d the lower side frequency, e the upper side frequency, and f the two side frequencies without the carrier. The choice of a sine or cosine function for representing steady-state conditions is a matter of convenience. The *envelope* of the composite wave (shown dotted) has the same shape as the original signal wave. The intercepts of the composite wave with the zero axis are not changed by a.m.

3. Amplitude Modulation with Several Signal Frequencies. The modulation may be expressed as a Fourier series when the original signal wave contains components at several frequencies. The composite signal is

$$i = I_0[1 + m_0 + \Sigma m_n \cos (nat + \alpha_n)] \cos \omega t \tag{4}$$

When Eq. (4) is expanded, independent pairs of side frequencies appear for each signal frequency $(na/2\pi)$. The bands of frequencies above and below the carrier frequency occupied by the side frequencies are *upper side bands* and *lower side bands*, respectively. The *band width* is the frequency spectrum occupied by both side bands and carrier. In a.m. it is two times the highest modulating frequency in the original signal.

The relative phase (α_n) of each of the signal components must be preserved in order to maintain the original form of the signal. This is relatively unimportant in signals for music and speech but is sometimes exceedingly critical, as in video signals in television systems.

The d-c component (m_0) of the modulating signal has the effect of changing the magnitude of the carrier wave and represents a signal variation which is so slow relative to the remaining components that it may be considered constant over the steady-state interval under examination. This is illustrated in Fig. 2, which is representative of the envelope of a modulated television signal wave for two different time intervals. Changes in the d-c component from time to time must be regarded as variations in the original signal.

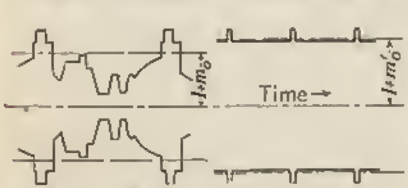


FIG. 2.—Envelopes of a-m waves.

Inward modulation and *outward modulation* are the respective decreases or increases in the envelope of the composite signal wave relative to the carrier. As illustrated in Fig. 2, inward and outward modulation are not necessarily alike. Inward modulation must not reduce the carrier to zero, or the character of the original signal will be lost. The maximum outward modulation determines the maximum power in the modulated wave.

In many cases it is sufficient to express the complete wave as

$$i = I_0(1 + M) \cos \omega t \tag{5}$$

when the instantaneous modulation M varies slowly with respect to the frequency of the carrier.

Carrier suppression is the process of balancing out the carrier component in an a-m wave, leaving only the frequency components in the side bands (see Fig. 1*f*). The transmitted power is then zero in the absence of modulation, resulting in an increase in transmitting efficiency. A carrier must then be suplied locally at the receiver for detection.

Single-side band and *vestigial-side band* signal waves are a-m waves in which all or a portion of the side-frequency components above or below the carrier frequency have been removed by suitable filters. The portion of the frequency spectrum from which the components were removed is then available for other services.

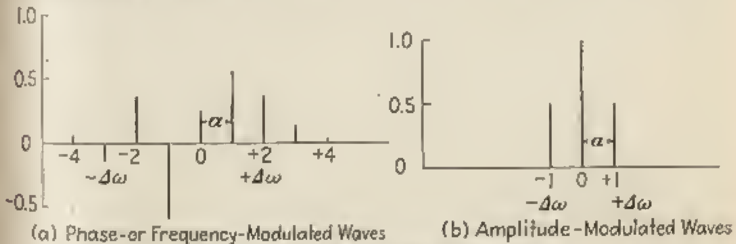
4. Phase Modulation (P.M.) and Frequency Modulation (F.M.). The peak amplitude of the composite signal is constant in p-m and f-m waves. The signal is imparted to the carrier by a variation of the phase as a function of time.

A single signal component of frequency $a/2\pi$ produces a current

$$i = I_0 \cos (\omega_c t + \theta + m \sin at) \tag{6}$$

where I_0 = constant amplitude
$\omega_c/2\pi$ = constant carrier frequency
θ = constant relative carrier phase
m = relative maximum variation in phase or *modulation index* due to the signal.

θ may be neglected unless the current is to be combined with other currents of the same frequency.



(a) Phase-or Frequency-Modulated Waves (b) Amplitude-Modulated Waves
FIG. 3.—Component side frequencies.

Neglecting θ and making I_0 unity for convenience, Eq. (6) may be expanded to

$$\begin{aligned} i = {} & J_0(m) \cos \omega_c t \\ & + J_1(m)[\cos (\omega_c + a)t - \cos (\omega_c - a)t] \\ & + J_2(m)[\cos (\omega_c + 2a)t + \cos (\omega_c - 2a)t] \\ & + \cdots\cdots\cdots\cdots\cdots \\ & + J_n(m)[\cos (\omega_c + na)t + (-1)^n \cos (\omega_c - na)t] \end{aligned} \tag{7}$$

where $J_n(m)$ is the Bessel function of the first kind and nth order for the argument m. The components in Eq. (7) are the carrier and side frequencies in a p-m or f-m wave at a single signal frequency.

An infinite number of side frequencies spaced $a/2\pi$ in frequency is indicated for complete identity, but the Bessel functions $J_n(m)$ are negligible for values of n some 20 to 40 per cent greater than the value of m. The number of necessary side frequencies is somewhat greater than $2m$. The carrier component is less than the unmodulated carrier and may be negative. The intercepts of the composite wave with the zero axis are not equally spaced.

Figure 3*a* illustrates the carrier and side-frequency components for p.m. or f.m. for $m = 2$ for a single modulating frequency. In comparison, Fig. 3*b* illustrates the carrier- and side-frequency components for 100 per cent a.m. for a single modulating frequency.

The instantaneous angular frequency ω of the composite wave is the derivative of the instantaneous phase with respect to time and is

$$\omega = \omega_c + ma \cos at \tag{8}$$

The maximum *frequency deviation* Δf is $a/2\pi$ times the maximum *phase departure* or modulation index m, and is

$$\Delta f = \frac{ma}{2\pi} \tag{9}$$

5. Differences in P.M. and F.M. The original signal usually contains components at several frequencies. The maximum phase departure and the maximum frequency deviation for several unrelated modulating frequencies are then

$$\Delta\phi = m_1 + m_2 + m_3 + \cdots + m_k \tag{10}$$

and

$$\Delta f = \frac{1}{2\pi}(m_1a_1 + m_2a_2 + \cdots + m_ka_k) \tag{11}$$

When the component maximum phase departures m_1 to m_k are made proportional to the amplitudes of the signal components at frequencies $a_1/2\pi$ to $a_k/2\pi$, the composite wave is said to be *phase-modulated*.

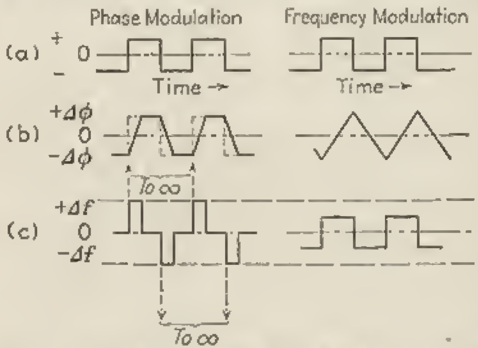


FIG. 4.—Comparison of phase and frequency modulation.

When the component maximum frequency deviations are made proportional to the amplitudes of the signal components, the composite wave is said to be *frequency-modulated*.

These relations for p-m or f-m waves, as indicated by the subscripts p or f are summarized in the following equations:

$$(\Delta\phi)_p = (\Delta\phi)_1 + (\Delta\phi)_2 + \cdots + (\Delta\phi)_k \tag{12}$$

$$(\Delta\phi)_f = \frac{(\Delta\omega)_1}{a_1} + \frac{(\Delta\omega)_2}{a_2} + \cdots + \frac{(\Delta\omega)_k}{a_k} \tag{13}$$

$$(\Delta f)_p = \frac{1}{2\pi}[(\Delta\phi)_1a_1 + (\Delta\phi)_2a_2 + \cdots + (\Delta\phi)_ka_k] \tag{14}$$

$$(\Delta f)_f = \frac{1}{2\pi}[(\Delta\omega)_1 + (\Delta\omega)_2 + \cdots + (\Delta\omega)_k] \tag{15}$$

The expansion of the expression for current when several signal frequencies are present contains side-frequency terms which are spaced from the carrier

frequency by $\pm\frac{1}{2\pi}(n_1a_1 \pm n_2a_2 \pm n_ka_k)$ where the n's are all positive integers to approximately 1.3 times the respective values of m for the corresponding component signal frequencies. This ratio approaches unity as the value of m is increased. The maximum band width required is therefore approximately 2.6 times the maximum frequency deviation for either p-m or f-m waves.

There is no upper limit to the degree of modulation which may be applied without distortion by p.m. or by f.m. except as determined by the capability of the equipment to operate over the required frequency band.

In wide-band f-m systems the maximum allowable frequency deviation is specified by assignment to prevent interference with other services.

Figure 4 illustrates the limiting case for the transmission of a rectangular signal wave by means of p.m. or by f.m. The curves a are the original signal, b the phase departures, and c the frequency deviations, all plotted against time. The required side bands for p.m. with rectangular signals extend to plus or minus infinite frequency for perfect transmission. When the maximum frequency deviation is limited by the equipment as at the dashed lines in c, the best possible operation is as shown by the full lines, whereas ideal operation would be according to the dotted lines.

AMPLITUDE-MODULATED WAVES IN NON-LINEAR CIRCUITS

6. Modulation, Frequency Conversion, and Detection. Essentially the same classes of non-linear devices are used for the modulation, frequency conversion, or detection of a-m waves. In each process the waves to be *combined* or *resolved* are applied to circuit elements which have asymmetrical *E-I* characteristics. These may be series characteristics, as in dry rectifiers or diodes, or may be mutual characteristics, as in multielectrode vacuum tubes.

The output circuit of a modulator is arranged to transmit the carrier and its side bands; that of a frequency converter, the i-f carrier and its side bands; and that of a detector, the components at the frequency of the original signal. The components in the voltage developed by the output current at other frequencies are eliminated by proper filtering.

Intermodulation is the production of new components having frequencies corresponding to undesired sums and differences of the fundamental and harmonic frequencies of the components of the applied waves.

Cross modulation is a type of intermodulation in which the carrier of the desired output signal is modulated by an undesired signal.

Modulation distortion is a change in the character of modulation either in an increase in the percentage of modulation or in the production of harmonics of the modulating signal due to intermodulation.

Spurious modulation components may be predicted by substituting the desired and interfering input signals in the power-series expressions, Eq. (16) or (18), for plate current, expanding, and collecting the terms at the frequencies in question.

7. Input to a Single Grid. The grid-plate characteristic of a vacuum tube in which the plate current is substantially independent of the load may be represented by the power series

$$i = A_0 + A_1e + A_2e^2 + A_3e^3 + \cdots \tag{16}$$

where the A's are coefficients determined by test and e is the instantaneous input voltage. Specifically, the coefficients are

$$A_n = \frac{1}{n!}\frac{\partial^n i}{\partial e^n} \tag{17}$$

at the steady value of e which is maintained by bias voltage.

The power series Eq. (16) is useful quantitatively as well as qualitatively in class A tubes which do not draw grid current and which are not worked to plate-current cutoff. In class B and class C services the power-series equation may require too many terms for the extended range and may converge too slowly to be of use quantitatively, but it is always of value in determining the frequency range of the possible output components. The term A_2e^2 provides the largest part of the useful modulation output, whereas the term A_1e provides the useful output in amplification. When the higher order terms are absent, as at low input levels in class A tubes, the useful output may be calculated accurately. The tube is then said to be operating as a *square-law* device.

In triodes (or when the output impedance is appreciable with respect to the plate impedance), the plate current depends on the plate voltage as well as on the grid voltage. The voltages developed by the plate current in the plate load are reimpressed on the plate, thereby reducing the useful output. To a first approximation, neglecting the higher order terms, the output current from Eq. (16) is reduced by the factor $r_p/(r_p + Z)$ where r_p is the plate resistance of the tube and Z is the load impedance at the frequency of the desired output. If the load impedance is non-uniform over the band, the correction may be applied separately to each of the output-frequency components.

8. Input to Two Electrodes. The plate current of a vacuum tube, in which the plate current is substantially independent of the load and which is controlled by the potentials on more than one electrode, may be expressed by the double-power series

$$i = A_0 + A_1e_1 + A_2e_1^2 + A_3e_1^3 + \cdots + e_2(B_0 + B_1e_1 + B_2e_1^2 + \cdots) + e_2^2(C_0 + C_1e_1 + C_2e_1^2 + \cdots) \tag{18}$$

in which the applied voltages are e_1 and e_2 and in which the coefficients are determined by test.

$$A_1 = \frac{\partial i}{\partial e_1} = g_1$$

$$A_2 = \frac{1}{2}\frac{\partial^2 i}{\partial e_1^2} = \frac{1}{2}\frac{\partial g_1}{\partial e_1}$$

$$B_0 = \frac{\partial i}{\partial e_2}$$

$$C_0 = \frac{1}{2}\frac{\partial^2 i}{\partial e_2^2}$$

$$B_1 = \frac{\partial^2 i}{\partial e_1 \partial e_2} = \frac{\partial g_1}{\partial e_2}, \text{ etc.} \tag{19}$$

at the steady values of e_1 and e_2, which are maintained by bias voltages.

The double-power series Eq. (18) may be used qualitatively for an estimate of the possible frequency components present in the output, even if it converges too slowly for quantitative results.

The term $B_1e_1e_2$ provides useful modulation output. When the higher order terms are small, as at low input levels in class A operation, the useful output may be calculated accurately.

9. Linearity of Output. While the power-series equations are extremely useful in class A calculations and in showing what distortion components may be present under less favorable circumstances, a direct indication of the linearity of the desired output in terms of the variable input is more often used. Tests of vacuum-tube modulators and detectors may be conducted at any convenient frequency, *e.g.*, 60 cycles, if the circuit impedances at the desired operating frequency are duplicated at the test frequency. The linearity of the dynamic characteristic of the controlled current is a direct indication of the linearity of the desired output.

For modulators the output current (or the voltage developed in the load by it) at the test frequency is plotted against steps in the input voltage corresponding to its variation by the modulating voltage. For detectors, the d-c output is plotted against steps in the r-m-s or peak value of the test input voltage, corresponding to the modulated input voltage for different parts of the modulating cycle.

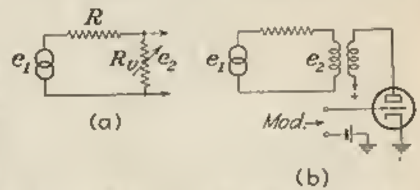


FIG. 5.—Method of absorption modulators.

The output curves which are the most nearly linear over a wide range of the independent variable, taken for different load resistances, bias voltages, etc., indicate the operating condition which will accommodate high percentages of modulation with the least distortion.

AMPLITUDE MODULATORS

10. Absorption Modulators. Absorption modulation is obtained by varying a resistance either in series with or in parallel with the load in accordance with some function of the modulating voltage. In Fig. 5a suppose that R_v is a resistance which includes the load and which is varied linearly with the modulation or

$$R_v = R_0(1 + M) \tag{20}$$

where M = the instantaneous value of the modulation.

The output voltage e_2 is

$$e_2 = \frac{e_1R_v}{R + R_v} \tag{21}$$

when e_1 = applied voltage
R = resistance of the source.

The ratio of e_2/e_1 is plotted against R_v/R in Fig. 6, which indicates that reasonably linear operation is obtained over a small portion of the curve when R_v is small compared to R. The dotted curve is the output voltage across R as a load when R_r is used as a variable series resistor. Physical resistances can be varied over a limited range, say three to one. The usable portion of the curve of Fig. 6 is then limited to the section marked *ab* in Fig. 6. The efficiency and the effective degree of linear modulation in the output are low in any case.

The plate resistance of a vacuum tube as controlled by the modulating voltage applied to its grid may be used as the variable resistor for absorp-

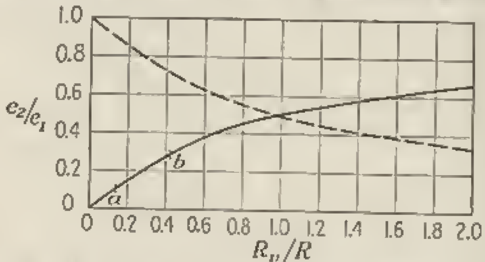


FIG. 6.—Linearity in absorption modulation.

tion modulation in parallel with the output load as shown in Fig. 5b. Since the plate resistance is not a linear function of the grid voltage, the over-all linearity may then be improved somewhat by working on a portion of the tube characteristic which tends to cancel the required curvature indicated in Fig. 6. Tests are then made for linearity of load voltage *versus* grid voltage.

The plate resistance required of the tube is the inverse of the resistance R_r when a quarter-wave transmission line or its filter equivalent are

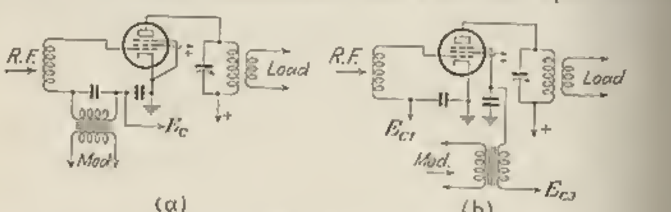


FIG. 7.—Types of grid modulators.

interposed between the tube and the load. The required plate resistance R_p is then

$$R_p = \frac{Z_0^2}{R_r} \tag{22}$$

where Z_0 = image impedance of the line.

Absorption modulation has been used to supplement other modulating methods over portions of the modulating cycle.

11. Grid Modulators. Grid modulators operate with carrier and signal voltages applied to the same or separate grids as illustrated in Fig. 7a or 7b. The plate current may be calculated by Eq. (16) or (18) for

low levels when the plate current is not swung to cutoff. The action is illustrated in Fig. 8 for the connections of Fig. 7a and for a square-law tube. Curve *a* shows the input signal and carrier voltages superimposed, curve *b* the instantaneous plate current, and curve *c* the modulated output voltage with the l-f components filtered out.

In this case, when the applied voltage about the operating point E_c' is

$$e = E \cos \omega t + S_1 \cos a_1 t + S_2 \cos a_2 t + \cdots \tag{23}$$

where S_1 and S_2 are the signal amplitudes at frequencies $a_1/2\pi$ and

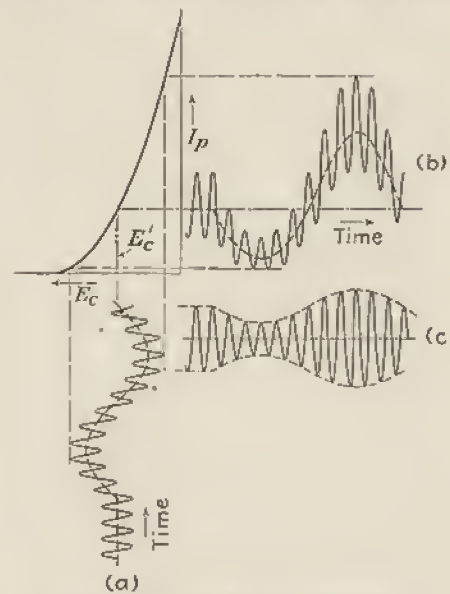


FIG. 8.—Low-level grid modulation.

$a_2/2\pi$, etc., the useful output current is

$$i = E(A_1 + 2A_2S_1 \cos a_1 t + 2A_2S_2 \cos a_2 t + \cdots) \cos \omega t \tag{24}$$

This may be written simply

$$i = E(A_1 + 2A_2M) \cos \omega t \tag{25}$$

where M indicates the instantaneous applied modulating signal. The product $M \cos \omega t$, when expanded, produces all the pairs of side frequencies required for the modulated wave. There are no spurious modulation components. However, this mode of operation does not realize fully the power capability of the tube, and the modulation cannot approach unity.

A grid modulator is operated as a class C carrier-frequency amplifier for higher plate efficiency. A tube with a linear grid-plate characteristic is suitable. The bias is adjusted to about twice the value required for plate-current cutoff, and the carrier input voltage is adjusted until the peaks reach halfway between saturation and plate-current cutoff. The superimposed modulating voltage at signal frequency causes the positive peaks to vary widely in value as shown by curve a in Fig. 9. Curve b shows the instantaneous plate current and curve c the useful modulated

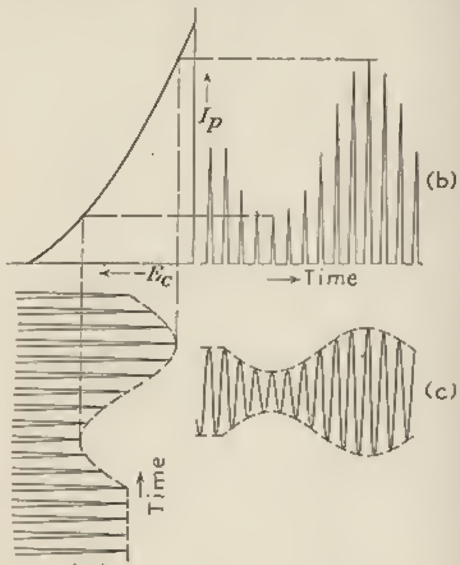


FIG. 9.—Class C grid modulation.

output voltage with the l-f and carrier-harmonic components filtered out.

Linearity may be tested by the method of Art. 9 and observing the output at the test frequency for a range of bias voltages. The exact bias setting is then at the center of the linear portion of the test curve.

Grid modulators have the advantage of requiring small signal input power, particularly when the tubes are not driven to grid current, but have limited ranges of linear modulation and plate efficiencies of only 20 to 30 per cent. They are used ordinarily at low power levels.

Grid modulators are used for television signals since it is difficult to obtain reasonable operation with high-level plate modulation over the required wide band of television modulation frequencies. Triodes may be used in grid modulators if neutralized to prevent h-f feedback. The output is then reduced as explained in Art. 7. The voltage, current, and power in the plate circuit have the following approximate relations:

$$E_{\text{max.}} = E_c(1 + M)$$
$$W_{\text{max.}} = W_c(1 + M)^2$$
$$i_p(\text{max.}) = I_c(1 + M)$$
$$W_{\text{av.}} = W_c\left(1 + \frac{m^2}{2}\right) \tag{26}$$
(for sine-wave modulation)

where the subscript c indicates the conditions for the carrier alone.

12. Balanced Modulators. When carrier voltage is applied in phase and modulating voltage is applied in push-pull to the grids of two modulator tubes, the carrier is balanced out in a push-pull output load. The circuit shown in Fig. 10 with two neutralized triodes is typical.

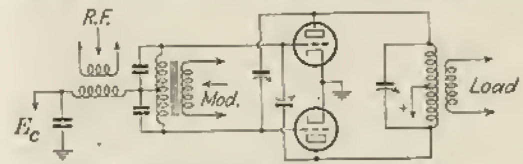


FIG. 10.—Circuit for balanced modulator.

For exact balance of tubes and transformers, and over the range of signals for which the modulation characteristic is linear, the useful modulated output current from the two plates is

$$i_1 = E_c(A_1 + 2A_2e_m) \cos \omega t \tag{27}$$

and

$$i_2 = E_c(A_1 - 2A_2e_m) \cos \omega t \tag{28}$$

where E_c = maximum carrier voltage
and e_m = instantaneous modulating voltage.
The effective input current to the tank circuit is

$$i = 4A_2E_ce_m \cos \omega t \tag{29}$$

which contains only the side bands. For a single modulating frequency $a/2\pi$, this reduces to

$$i = 2A_2E_ce_m[\cos (w + a)t + \cos (w - a)t] \tag{30}$$

The voltage developed in the output circuit is

$$e = 2A_2E_ce_mZ[\cos (w + a)t + \cos (w - a)t] \tag{31}$$

where Z = load transfer impedance. This arrangement is used in suppressed-carrier transmission systems. It has the advantage of balancing out any even-harmonic distortion due to departure of the modulation characteristic from linearity, not considered in the above equations.

When the modulating input voltages to the two grids, the tube coefficients, and the effective load transfer impedances for the two tubes are unequal, the net tank circuit voltage is

$$e = E_c[A_1'Z' - A_1''Z'' + 2(A_2'e_m'Z' + A_2''e_m''Z'') \cos at] \cos \omega t \tag{32}$$

where the ′ and ″ values are for the first and second tubes respectively. Some of the carrier remains when the balance is not perfect.

13. Plate Modulators. The constant-current plate modulator utilizes an a-f choke coil in the circuit which supplies plate power to r-f and a-f amplifier tubes, as shown in Fig. 11. The total plate current remains constant by virtue of the inductance of the choke. The instantaneous audio-plate voltage is added to the plate supply voltage and over the audio cycle changes the latter to $E_b(1 + M)$. The r-f inductance L prevents the loss of h-f power in the audio tube, and the condenser C prevents the short circuit of the audio-plate voltage.

The a-f tube, frequently called the *modulator*, supplies modulating power, but the actual modulation occurs in the plate circuit of the r-f tube.

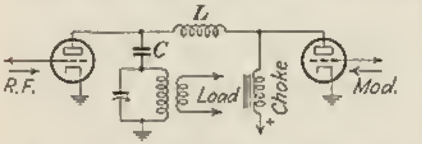


Fig. 11.—Circuit for plate modulation.

The plate current of the a-f tube cannot be reduced to zero during the modulation cycle without introducing audio distortion. It is necessary, therefore, to operate the audio tube with a higher zero-signal plate current than the radio tube in order to reach unity modulation without audio distortion. This is done by applying a higher plate voltage to the audio tube, either by using a by-passed resistance in the plate circuit of the r-f tube or by supplying the audio tube (or tubes) through transformer coupling as illustrated in Fig. 12. A further improvement is indicated in Fig. 12, since saturation of the transformer core is prevented by eliminating the d-c magnetizing component and since even-order audio harmonics are canceled by the push-pull arrangement. The efficiency of the system is increased by operating the audio tubes in push-pull class B. In transformer coupling the total d-c is no longer constant but varies with the modulation.

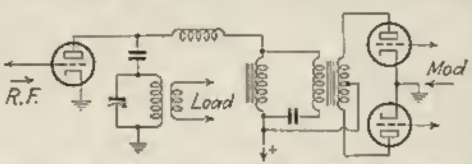


Fig. 12.—Transformer-coupled plate modulator.

The voltage, current, and power in the plate circuit have the following relations for sine-wave modulation:

Tube and circuit voltage,

$$E_{\max.} = (1 + m)E_b$$

R-f input power,

$$W_{\max.} = (1 + m)^2 E_b I_b$$

Average r-f input power,

$$W_{\text{av.}} = \left(1 + \frac{m^2}{2}\right) E_b I_b$$

Average output power,

$$W_0 = \eta\left(1 + \frac{m^2}{2}\right) E_b I_b$$

Average audio input power,

$$W_a = \frac{m^2}{2} E_b I_b$$

R-f plate loss,

$$W_p = (1 - \eta)\left(1 + \frac{m^2}{2}\right) E_b I_b \tag{33}$$

where E_b and I_b = the d-c supply voltage and current, respectively
m = the degree of modulation
η = the plate efficiency of the radio tube.

The plate efficiency is high (often 0.7 to 0.8), and the chief disadvantage of plate modulation is the large amount of audio power which must be supplied.

The radio tube is operated as a class C amplifier with bias at approximately two times cutoff. Since the plate losses are 50 per cent higher with full modulation than for unmodulated output, the tube must be used at two-thirds its rated power. Low-mu triodes are suitable and ensure low plate and grid voltages.

Grid-leak bias helps in obtaining linearity up to complete modulation. Linearity may be checked by direct adjustment or by test at 60 cycles with a proper plate load for a range of plate voltage from 0 to $2E_b$.

14. Modulated Oscillator. Plate modulation was originally applied directly to the oscillator tube and circuit. Practically full modulation may be obtained with excellent linearity, but the arrangement has the disadvantage of introducing f.m. The frequency of the oscillator varies with the plate voltage, and, since in plate modulation this varies between $2E_b$ and 0 during full modulation, the oscillator frequency deviates from its mean value with the modulating signal.

The same circuits are used between the two tubes, and the same voltage, current, and power relations hold as with a plate-modulated amplifier. Linearity is obtained by adjusting the value of the grid leak.

Modulated oscillators are now considered suitable only for test equipment in which the f.m. is not objectionable.

15. Copper Oxide Modulators. Copper oxide rectifiers are applicable in bridge modulators and are used widely in carrier-current telephony. They function as carrier-operated switches for opening, shorting, or reversing the elements carrying the modulating currents.

Copper oxide rectifiers are not suitable for use at frequencies much above 1 Mc, except at low impedance levels on account of inherently large shunt capacitance. They are compact in size and eliminate the heater connections necessary in similar circuits using vacuum tubes. They maintain a satisfactory balance in carrier-suppression circuits using balanced modulators.

The power-series current equation for a copper oxide unit converges slowly, and its characteristics are expressed more easily quantitatively in terms of resistance for different applied voltages. For voltages in the reverse direction and for less than 0.02 volt in the forward direction, the resistance is high and substantially constant. For forward voltages between 0.02 and 0.6 volt the resistance is approximately

$$r = r_0 \epsilon^{-ke} \tag{34}$$

where k is a constant which may be as great as 18. For forward voltages larger than 0.6 volt the resistance is low and nearly constant.

Representative circuits using copper oxide units are shown in Figs. 13 and 14. In these figures, f_c and f_s indicate voltage sources at the carrier and signal frequencies, respectively. The impedances of the carrier source, signal source, and load are Z_c, Z_s, and Z_L, respectively. The forward conducting direction of the units is shown by the arrows.

Since the carrier and signal voltages are applied to conjugate terminals of the rectifier bridges, the resistance effects are balanced. In Fig. 13 the output signal is short-circuited for one polarity of the carrier cycle. In Fig. 14 the effective connections between the signal source and the

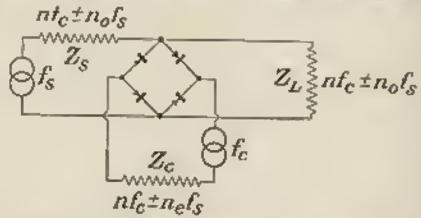


FIG. 13.—Type of bridge modulator using non-tube rectifiers.

load are reversed as the polarity of the carrier changes. This arrangement is called a *double-balanced* or *ring modulator.*

The frequencies of the current components produced in the individual rectifier units are determined qualitatively by an expansion of each of the terms of a power-series equation for the current. Current components of frequencies equal to the sums and differences of the integral multiples of the carrier and signal frequencies appear in each unit in the forward direction. These combine additively or differentially in the connected circuits depending on the polarity. In flowing through the circuit impedance these current components produce voltages of the same frequency which are reimpressed upon the rectifier units. The final result may be obtained quantitatively only by a series of approxima-

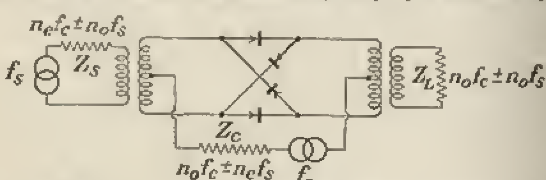


FIG. 14.—Double-balanced or ring modulator.

tions. The frequencies of the components appearing in the circuit impedances are indicated in Figs. 13 and 14, where n is any whole number or zero, n_o is any odd number, and n_e is any even number or zero.

The output impedances are designed as filters to eliminate voltages at frequencies involving undesired multiples of f_c. The useful output is at a frequency $f_c \pm f_s$, and, when double modulation is used, the carrier frequency f_c is eliminated. The ring or double-balanced modulator differs from the simple-bridge modulator in having no term of frequency f_s in its output.

By making the carrier voltage large in comparison with the signal voltage, the terms involving multiples of f_s may be reduced satisfactorily in magnitude. The units are operated with about 0.5 volt carrier across each disk in the forward direction. The optimum impedance in the signal and output circuits is

$$Z = \sqrt{R_f R_r} \tag{35}$$

where R_f and R_r are the resistances in the forward and reverse directions, respectively. The loss in conversion is then only 6 to 8 db.

Copper oxide bridge modulators differ from van der Bijl vacuum-tube balanced modulators in that they transmit in either direction. They function equally well as modulators or demodulators.

MODULATORS FOR P-M AND F-M WAVES

16. Phase Modulators. The usual method of producing p-m waves is to combine the output of a suppressed-carrier balanced amplitude modulator with an unmodulated carrier which differs in phase by 90 deg from the original carrier. A vector diagram of the carrier and the net side-frequency components plotted relative to the carrier is shown in Fig. 15 for a single modulating frequency. The net side-frequency voltage E_s is in the direction shown but varies in magnitude according to cos at.

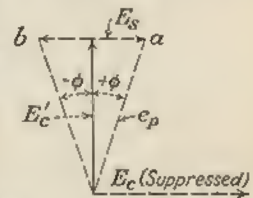


FIG. 15.—Vector relations in phase modulation.

The resultant phase-modulated voltage e_p varies in phase from the new carrier E_c' by the angle,

$$\phi = \tan^{-1} \frac{E_s \cos at}{E_c'} \tag{36}$$

which, for angles less than about 25 deg., is approximately

$$\phi \cong \frac{E_s \cos at}{E_c'} = m_p \cos at \tag{37}$$

The resultant voltage varies only slightly in magnitude and is

$$e_p = E_c' \cos (\omega_c t + m_p \cos at) \tag{38}$$

When modulating signals at more than one frequency are present, the coefficients m_1, m_2, etc., are proportional to the original a.m., and p-m waves are produced.

The small phase departure of less than 25 deg. may be increased by frequency multiplication of the instantaneous frequency. (See Art. 37.) The new voltage is then

$$e_p' = E_c' \cos (n\omega_c t + nm \cos at) \tag{39}$$

17. Frequency Modulators. Frequency-modulated waves are obtained by the method described in the section above when the modulating signal is passed through a filter whose response is inversely proportional to the signal frequency. The instantaneous frequency is multiplied several hundred times before the output is applied to the antenna.

A more direct method consists in controlling the reactance of the oscillator tuned circuit by a reactance control tube in accordance with the signal. A typical circuit is shown in Fig. 16 as one of many possible arrangements. Radio-frequency voltage, shifted in phase 90 deg. from that appearing across the tank circuit, by means of the resistance R in series with the capacitance C, is applied to the grid 1 of the control tube. The plate current of the control tube is 90 deg. out of phase with the tank voltage and provides an effective reactance which is controlled in magnitude by the modulating voltage on grid 3.

The change in oscillator frequency is proportional to the instantaneous modulating voltage when the control tube operates on a linear part of its characteristic and when the total change in effective reactance is small compared to the net average reactance.

Any component of control-tube plate current not at 90 deg. with the tank voltage will introduce a.m. This is eliminated by adjusting the phase shift to grid 1 of the control tube.

The curve of frequency deviation from the carrier frequency should be linear with respect to the modulating voltage. It may be checked by

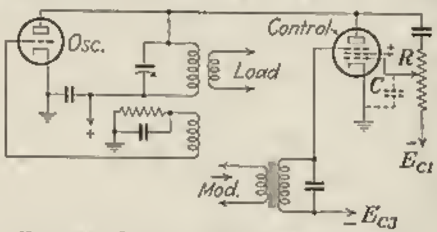


FIG. 16.—Circuit of frequency modulator.

applying direct voltages, over the operating range, to the modulating grid and observing the oscillator frequency.

The circuit illustrated in Fig. 16 provides f-m waves when the modulating voltages are proportional to the amplitude of the signal. If the signal is passed through a filter whose output is proportional to the signal frequency before application to the control tube, p-m waves are produced.

FREQUENCY CONVERTERS AND MIXERS

18. Class A Frequency Converters. A power-series expression for plate current accurately determines the output of a tube used in double-modulation service when the plate current is not swung to cutoff.

The signal and local oscillator voltages e_s and e_0 may be applied to the same or separate grids. The third term of the series A_2e^2 of Eq. (16) when expanded yields the i-f plate current,

$$i_{if.} = \frac{A_2E_0E_cr_p(1 + m\cos at)\cos(\omega_0 \pm \omega_c)t}{r_p + Z} \tag{40}$$

when the signal and oscillator voltages are

$$e_s = E_c(1 + m\cos at)\cos\omega_c t$$

and

$$e_0 = E_0\cos\omega_0 t$$

and r_p and Z are the plate and output impedances, respectively.

The impedance of the output circuit at the i-f frequency is usually too large in comparison with the plate resistance to be neglected. It is desirable to use a low-impedance primary winding in the output circuit of triodes in order to obtain the best conversion gain.

The output of a class A frequency converter is low since the capability of the tube is not realized fully. It is a van der Bijl modulator with a change in frequency between the input and output carriers. When terms of higher order than A_2e^2 are negligible, the operation is according to a square-law characteristic, and there are no spurious intermodulation responses. The output voltage may then be expressed as

$$e_{if.} = \frac{Zr_pA_2E_0E_c}{r_p + Z}(1 + M)\cos\omega_{if.}t \tag{41}$$

$$= \frac{Zr_p}{r_p + Z}S_cE_c(1 + M)\cos\omega_{if.}t \tag{42}$$

where M represents the instantaneous modulation. S_c is the *conversion transconductance* which is the i-f plate current per volt of applied signal for the selected oscillator and bias adjustments.

Conversion gain is the ratio of the voltage developed in the i-f circuit (usually measured at the grid of the following tube) to the signal voltage applied to the converter tube.

19. Superheterodyne Frequency Converters. Frequency converters, sometimes called *first detectors,* in superheterodyne receivers are operated with as large a local oscillator voltage as possible without endangering non-linear i-f response. This ensures the highest conversion transconductance and conversion gain.

When the signal and oscillator voltages are applied to the same grid, the tube is biased nearly to cutoff in the absence of oscillator volage, and the oscillator voltage is made a volt or so less than that which would cause grid current. Plate current then flows for the positive peaks of oscillator voltage and is cut off for a large part of the cycle. The plate current is modulated by a relatively small signal voltage, and the sum or difference frequency components (usually the latter) are selected and tuned in the plate circuit as the i-f output.

The conversion transconductance under the most favorable conditions does not exceed about 0.3 of the transconductance of the same tube as an amplifier. Limited a-v-c bias may be applied to the signal grid for control of the conversion gain of variable-mu tubes.

20. Special Converter and Mixer Tubes. Interaction between the signal and oscillator circuits of frequency converters produces undesirable oscillator detuning. This may be reduced somewhat by coupling the oscillator voltage to the suppressor grid or to the cathode of a pentode converter tube, but even these expedients are ineffective when the percentage difference between the signal and oscillator frequencies is small, as in the h-f bands of so-called *all-wave broadcast receivers.*

Special pentagrid mixer tubes, such as the 6L7 tube, have been designed for frequency-converter service which give superior performance due to better shielding between the signal and oscillator grids, high plate resistance, high conversion transconductance, and suitability for a.v.c.

Specially designed multigrid converter tubes of several types are also available in which two of the electrodes serve as the grid and plate of the oscillator circuit. The electron stream as initially controlled by the oscillator is modulated by the signal applied to a screened control grid.

The best stability and conversion gain for u-h-f signals are obtained with a high i.f., with a separate oscillator tube, and with the signal and oscillator voltages applied to the same grid of the converter tube.

21. Spurious Responses. The desired signal input applied to converter tubes must not be sufficient to draw signal grid current or to drive the plate current to complete cutoff for the positive peaks of oscillator voltage during maximum modulation. Too large a signal produces harmonic distortion of the envelope of the i-f carrier.

The linearity may be checked by measuring the i-f output for a range of unmodulated signal voltages for the oscillator and bias conditions selected. For full modulation the linearity is satisfactory to half the signal level indicated by the test. Such a test is analogous to that described in Art. 9 for modulator tubes, but in this case it is difficult to develop a sufficiently high plate-circuit impedance at a low test frequency to simulate the value at i.f.

Strong interfering signals of a number of definite frequencies produce spurious responses in the output. A signal at the converter input, either higher or lower in frequency than the oscillator by the i.f., produces equal response in the tuned output circuit. One of these having been selected as the desired signal, the other is known as the *image response* and must be attenuated in preselector circuits ahead of the converter tube.

Other spurious responses are due to the terms A_3e^3, A_4e^4, etc., in the expression for plate current. These terms are absent in a tube with a square-law characteristic but are present in a normally operated converter tube for which the series converges slowly.

With insufficient preselection at high levels of interfering signal, an interfering program response may be heard, *e.g.*, when the second harmonic of the interfering signal carrier frequency differs from the fundamental of the oscillator frequency by the i.f. When a desired signal is also present, a beat between the intermediate frequencies from the two signals is rectified in the second detector.

This spurious response is predicted from the expansion of the term A_3e^3 where

$$e = E_s \cos \omega_s t + E_i \cos \omega_i t + E_0 \cos \omega_0 t \tag{43}$$

where E_s and E_i are the amplitude coefficients at the signal and interfering frequencies. The desired output is

$$A_2 E_s E_0 \cos (\omega_0 - \omega_s)t \tag{44}$$

and the particular interfering output is

$$\tfrac{3}{4} A_3 E_i^2 E_0 \cos (2\omega_i - \omega_0)t \tag{45}$$

Other interferences may be predicted from the other terms of the expansion of A_3e^3, A_4e^4, etc.

DETECTORS

22. Two-terminal Rectifiers. Units having asymmetrical *I-E* characteristics, and therefore having inherent rectifying properties, are suitable as detectors or demodulators. Many crystals possess these properties but are rather unstable and will carry only small currents. Copper oxide rectifiers are satisfactory at low carrier frequencies but have high self-capacitances which prevent their efficient use at high carrier frequencies. Diodes have low capacitances and are suitable at high carrier frequencies at any level of impedance which can be developed in their input circuits. Diodes have high voltage-handling capability and give substantially linear demodulation when used with proper load circuits.

The current-voltage characteristic of a diode changes from an exponential curve for negative voltages to a 3/2-power curve for positive voltages. For negative voltages

$$i = i_0 \epsilon^{\frac{e}{e_1}} \tag{46}$$

For positive voltages

$$i = h(e + e_0)^{3/2} \tag{47}$$

where i_0, e_1, h, and e_0 are constants. i_0 and e_0 increase with cathode temperature; e_1 is nearly independent of operating conditions; and h increases with cathode area.

A power-series static characteristic converges too slowly for practical analysis, and actual experimental curves are used in circuit design.

23. Diode Peak Detectors. A diode used with a load impedance that is high at zero and modulation frequencies in comparison with its forward resistance is an excellent peak or envelope detector. Satisfactory performance without audio distortion at high signal levels depends on the design of the associated circuits. Rectified load voltage, either positive or negative with respect to ground, is developed in the load resistance which is by-passed at carrier frequency, as shown in Fig. 17.

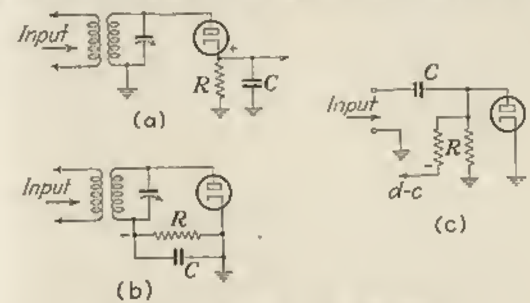


Fig. 17.—Typical diode detectors.

Diode current flows only for an instant at the peak of the carrier voltage in the forward direction. The pulses of current charge the load by-pass condenser to nearly the same voltage as the carrier envelope and bias the diode beyond cutoff except during the short pulses. Neglecting the slight h-f variation between pulses, the voltage across the load resistance is proportional to the carrier envelope.

The charge, which is replenished at each pulse, must leak off sufficiently before the following pulse for the bias voltage to follow the carrier envelope at its maximum slope. The critical relation is

$$\frac{1}{aRC} \geq \frac{m}{\sqrt{1-m^2}} \tag{48}$$

where $a/2\pi$ = modulation frequency
R = load resistance
C = by-pass capacitance
m = degree of modulation.

The capacitance C is made large enough to by-pass the carrier but small enough to reproduce the modulation. Since full modulation is seldom used at high frequencies and since the harmonics of the higher frequencies cannot be heard, it is sufficient to follow to about 0.8 modulation at 5,000 cycles in detection for sound reproduction.

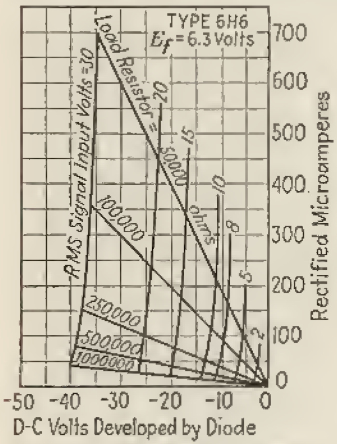


FIG. 18.—Rectification characteristic of diode.

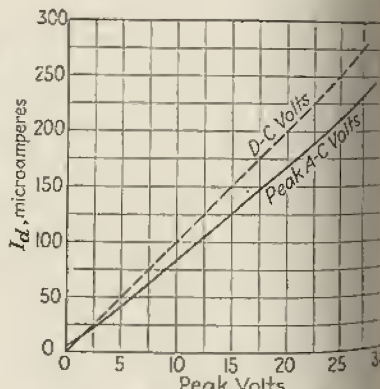


FIG. 19.—Load rectification diagram of diode.

Figure 18 shows the rectification characteristic of a typical diode. Rectified current is given for several steps of r-m-s input voltage in terms of d-c load voltage. These characteristics may be determined by test at 60 cycles according to the method of Art. 9.

The ratio of the d-c voltage E_d to the peak value of the applied voltage E_a is the voltage efficiency of the diode.

$$\eta = \frac{E_d}{E_a} \tag{49}$$

Curves for constant efficiency plotted on Fig. 18 would have the same general characteristic slope as the load-resistance lines shown.

Figure 19 shows a load rectification diagram for a 100,000-ohm load resistance as taken from Fig. 18 or plotted directly from test data. The linearity of Fig. 19 indicates that a diode is suitable for signals as large as can be supplied by the previous amplifier without overloading. The slight departure from linearity at the origin shows a small amount of inherent distortion for small input signals.

The difference between the peak a-c and the d-c voltages is almost proportional to the rectified current I_d and, except with small input signals, is equivalent to a resistance drop in series with the output. This effective internal diode resistance R_d is

$$R_d = \frac{E_a - E_d}{I_d} \tag{50}$$

A Fourier series expansion of the pulses of current shows that the peak value of its fundamental component at carrier frequency is twice the d.c. in the load with less than 1 per cent error. Thus the effective input resistance of the diode is

$$R_i = \frac{E_a}{2I_a} = \frac{E_d}{2\eta I_d} = \frac{R}{2\eta} \tag{51}$$

The impedance of the source Z_0 produces a drop in voltage

$$Z_0 I_a = 2Z_0 I_d \tag{52}$$

and the equivalent generator peak voltage E_0 is

$$E_0 = I_d(R + R_d + 2Z_0) \tag{53}$$

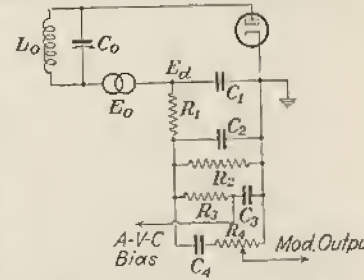


FIG. 20.—Typical diode circuit.

The above equations apply to an unmodulated carrier. In a practical circuit the d-c load may be shunted by other impedances at the modulation frequency. Figure 20 shows a circuit in which R_3 is a decoupling resistor for a-v-c supply and R_4 is an output resistor blocked for d.c. The resistor R_1, with the by-pass capacitors C_1 and C_2, filters the h-f components from the output. The d-c load resistance R is

$$R = R_1 + R_2 \tag{54}$$

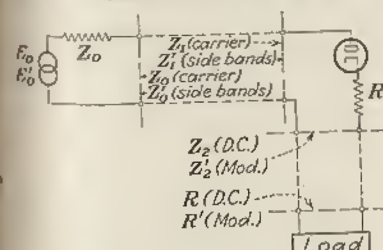


FIG. 21.—Equivalent diode circuit.

and the impedance R' at low modulation frequencies is

$$R' = R_1 + \frac{R_2R_3R_4}{R_2R_3 + R_2R_4 + R_3R_4} \tag{55}$$

The ratio R'/R is called the a-c/d-c ratio and is the most important single circuit relation in the operation of the diode.

The modulation-frequency voltage E_m developed in the load by the modulation-frequency component of the load current I_m is

$$E_m = R'I_m \tag{56}$$

The corresponding generator side-band voltage E_s relative to the carrier voltage is

$$E_s = I_m(R' + R_d + 2Z_0') \tag{57}$$

where Z_0' is the impedance of the source at the side-band frequency.

The equivalent output impedance Z_2 of the diode is

$$Z_2 = 2Z_0\eta \tag{58}$$

for the d-c component and

$$Z_2' = 2Z_0'_n$$

for the modulation frequencies.

The diode acts like a motor generator with input and output impedances which depend on the connected output and source impedances. An equivalent diagram is shown in Fig. 21.

24. Diode Performance. The relations between the rectified current and the peak voltages with resistive load during modulation are shown in Fig. 22. The impedance of the source to the side-band frequencies Z_0' is here assumed to be the same as for the carrier Z_0. The ordinates of the upper part of the diagram are for d-c or modulating-frequency load current. The abscissas are for d-c or a-c peak voltages. The load voltages are the current multiplied by R for d-c values and by R' for modulating-frequency values. The drop in the diode and in the source is the current multiplied by $(2Z_0 + R_d)$. The envelopes of the input, source and voltage are shown for one cycle of full modulation and for one cycle of modulation which just reaches diode cutoff during inward modulation. Specific dimensions are indicated for one point in the envelope in Fig. 22. The figure is exaggerated for purposes of illustration. The output voltage across resistor R_4 in Fig. 20 is reduced further by the voltage drop in resistor R_1 (not shown).

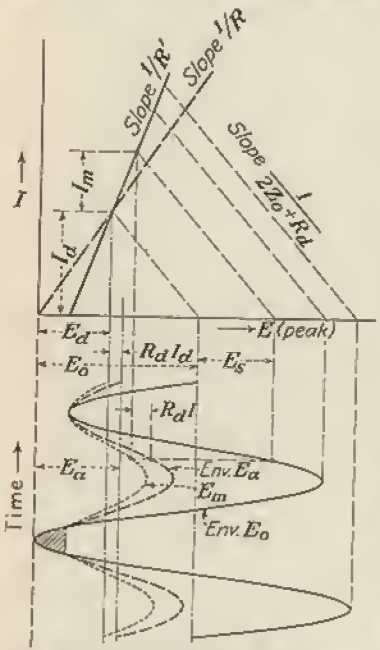


Fig. 22.—Diode detection with resistive load.

The peak of output voltage is clipped and the envelope of the input voltage rises when the degree of instantaneous inward modulation of the source exceeds the value

$$M = \frac{2Z_0' + R_d + R'}{2Z_0 + R_d + R} \tag{60}$$

as shown by the shaded area in Fig. 22. This is the reciprocal of the factor by which the degree of modulation of the source is increased in terms of the current.

The degree of modulation which is subject to linear detection may be increased by making R' nearly equal to R or by increasing Z_0.

A Fourier analysis of a sine wave of which one peak per cycle is clipped by a fraction V indicates that the r-m-s distortion d is approximately

$$d = 0.7V^{3/2} \tag{61}$$

The equivalent degree of modulation which can be produced without clipping for a permissible distortion D is

$$m = 1 - 1.33D^{2/3} \tag{62}$$

The unidirectional output voltage E_d at the terminals of the diode in terms of the carrier voltage E_0 is

$$E_d = \frac{E_0R}{2Z_0 + R_d + R} \tag{63}$$

and the demodulated output voltage E_m in terms of the envelope of the source voltage E_s is

$$E_m = \frac{E_sR'}{2Z_0' + R_d + R'} \tag{64}$$

Further reductions in output voltage not included in the above analysis are present under certain conditions. The output voltage is reduced by the ratio

$$\frac{C}{C + C_d} \tag{65}$$

when the diode capacitance C_d is appreciable in comparison with the capacitance C.

When the capacitances C_0 and C have appreciable reactance at the carrier frequency, the charge leaks off rapidly between current pulses, and the voltage is reduced by an amount which is equivalent to a voltage drop in series with the output load due to an equivalent resistance R_e, which is

$$R_e = \frac{\pi}{\omega C_0} + \frac{\pi}{\omega C} \tag{66}$$

where $(\omega/2\pi)$ is the carrier frequency.

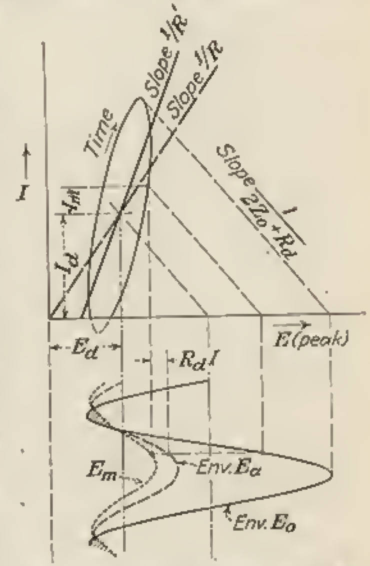


Fig. 23.—Diode detection with reactive load.

The capacitances C_0 and C or C_1 should be approximately equal for best performance with respect to peak clipping. On account of the effect indicated by Eq. (66), it is undesirable to feed a diode from an untuned winding of a transformer.

The portion of the voltage appearing across resistors R_3 and/or R_4 is calculated readily in terms of the total output voltage.

When the susceptance of condenser C (or C_1 and C_2) is appreciable in comparison with the conductance $1/R'$, the dynamic load line follows an ellipse such as is shown in Fig. 23 instead of the slope $1/R'$ as illustrated in Fig. 22. The modulation output voltage is then

$$E_m = \frac{E_sR'}{|2Z_0' + R_d + R' + jR'C\omega(2Z_0' + R_d)|} \tag{67}$$

instead of that indicated by Eq. (64).

This departure from a resistive load slightly increases the tendency to peak clipping at high modulation frequencies as shown by the shaded portion in Fig. 23.

The highest d-c output is required when the circuits supply a-v-c bias and it may then be desirable to use an input transformer with an untuned primary designed for optimum energy transfer.

The reduction of modulation-frequency voltages indicated by Eq. (67) is the least when the impedance of the source to the side-band frequencies is the conjugate of the diode input impedance. This is most nearly realized in practice by the use of an input transformer with tuned primary and secondary, which is recommended when it delivers a sufficient output voltage.

In general, the input impedance R_i is matched approximately to the source impedance Z_0 for maximum power transfer.

$$Z_0 = R_i = \frac{R}{2\eta} \qquad (68)$$

The over-all voltage efficiency of the diode and its associated circuits is ordinarily of the order of 0.2 for the modulation components and 0.3 for the direct voltage.

25. Biased Diodes. Fixed negative bias applied to a diode to prevent its operation with very weak signals shifts the load line of Fig. 22 parallel to its original position along the voltage axis. This results in peak clipping at lower degrees of modulation of the source. A separate biased diode for delayed a.v.c. should not be fed from the same circuits as a signal diode, since during peak clipping the voltage of the source rises as shown by the shaded areas in Fig. 22 or 23 and distorts the envelope delivered to the signal diode.

Peak clipping in a biased diode used for a.v.c purposes develops a lower rectified output during prolonged periods of deep modulation and may cause fluctuation of the receiver gain.

26. Push-pull Diodes. Diodes in push-pull require a minimum of load by-pass capacitance since only carrier harmonics are by-passed to ground. Such circuits are useful where high modulation frequencies must be reproduced, as in video detection in television receivers.

The input transformer for push-pull diodes must be carefully balanced since the bias developed across the load resistance is applied to both tubes. If the voltage peaks for one diode are less than the bias developed by the other, only one will function.

This condition is most critical when the diode efficiency is high, as with high d-c load resistances. Lack of balance is not so serious at the low-impedance levels used in detection for television.

27. Grid Detectors. The circuit elements connected between the grid and the cathode of a grid detector act substantially the same and are determined by the same considerations as are the corresponding elements in a diode. As with a diode the load on the previous circuit is equivalent to a shunt load of approximately half the resistance of the grid leak.

Figures 24*a* and *b* show typical triode and pentode power-grid detector circuits. The tube must be operated with low plate (or screen) voltage since in the absence of signal there is no bias voltage. This limits the useful plate swing. The range may be extended by operating the tube from a high-voltage *B* supply with resistance coupling, or with a resistor in series with the load, by-passed for audio, with transformer coupling.

The carrier-frequency components are filtered from the load circuit by *LC* or *RC* networks.

The operation is illustrated in Fig. 25 which shows the dynamic characteristic including the plate load selected. This may be obtained at any convenient frequency by the method of Art. 9.

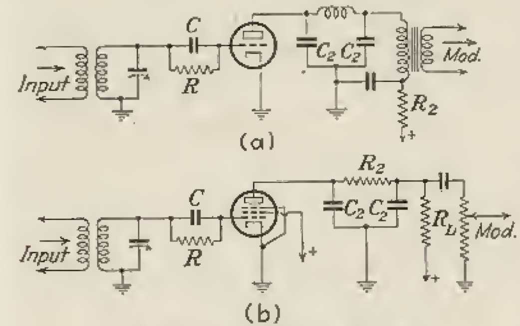


FIG. 24.—Power grid detectors.

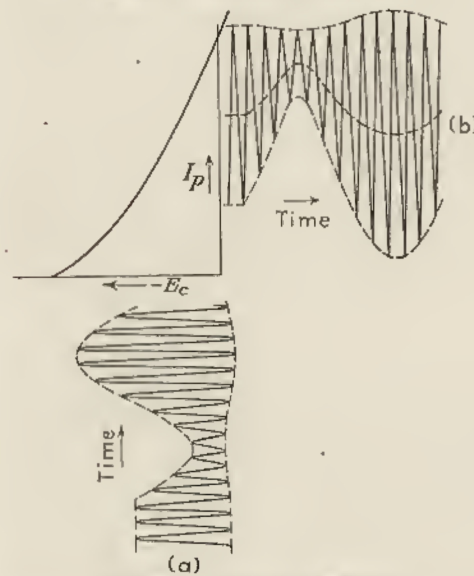


FIG. 25.—Power grid detection.

A pulse of grid current charges the grid condenser at each positive peak of grid voltage and establishes a negative grid bias with the same losses relative to the envelope as in a diode. The net applied-grid voltage, after the loss (not shown) in the source impedance is deducted, is indicated by curve a in Fig. 25. The instantaneous grid voltage due to the modulated wave is superimposed on the bias and produces instantaneous values of plate current as shown in curve b of Fig. 25. The amplified useful output is proportional to the average of the plate current upon which the individual cycles are shown superimposed.

If the negative peaks of instantaneous grid voltage swing over the curved lower portion of the characteristic, the audio is reduced and distorted by partial plate detection. This effect is exaggerated in Fig. 25, which shows more than normal curvature of the grid-plate characteristic to illustrate even-harmonic distortion of the modulation frequency in the output.

The audio output voltage which may be obtained satisfactorily is about 0.3 to 0.4 of the corresponding value for the same tube when used as an amplifier.

Power-grid detectors operate over a limited range of voltages which is insufficient for a.v.c.

28. Square-law Detectors. The sensitivity of a grid-leak detector for very weak signals may be increased by using a high-resistance grid leak. The grid then operates on a portion of the grid-cathode characteristic which is substantially square law and over which the grid current is never cut off. This method introduces harmonic distortion with high degrees of modulation and produces an output proportional to the square of the carrier voltage.

When the applied voltage is

$$e = E_c(1 + m \cos at) \cos \omega t \tag{69}$$

the power-series expression for grid current i_g yields from the term Ae^2 for square-law detection

$$i_g = \frac{A_2E_c^2}{2}\left(1 + 2m \cos at + \frac{m^2}{2} + \frac{m^2}{2} \cos 2\,at\right) \tag{70}$$

The audio component i_a of the grid current is

$$i_a = A_2E_c^2\left(m \cos at + \frac{m^2}{4} \cos 2\,at\right) \tag{71}$$

The same audio components appear in the output. For full modulation the second harmonic distortion is then 25 per cent.

A tube biased to a curved portion of its grid-plate characteristic operates approximately as a square-law detector for low input voltages. The output and distortion vary in the same way as with the grid-leak detector for weak signals. However, the weak-signal plate detector does not load the previous circuit and is therefore suitable for use as a vacuum-tube voltmeter. It is seldom used for demodulation.

29. High-level Plate Detectors. A plate-circuit detector for large signals is biased nearly to cutoff in the absence of a signal, and a high plate voltage is used to extend the range of operation. A typical circuit is shown in Fig. 26.

The operation is illustrated in Fig. 27 in which plate current is plotted against bias voltage for the load impedance selected. The positive excursions of the instantaneous grid voltage produce substantially half waves of plate current, the average values of each pulse producing the

audio voltage in the load, while the h-f components are by-passed to ground. Curve a of Fig. 27 shows the instantaneous values of input signal applied to the grid which is biased by the voltage $-E_c'$. Curve b shows the instantaneous values of plate current superimposed on the demodulated output. The effect of the curvature of the grid-plate characteristic is exaggerated in this figure to illustrate the even harmonic distortion from this cause in the output.

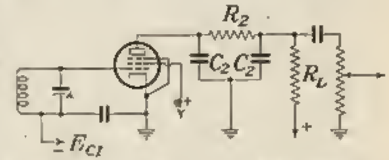


FIG. 26.—Circuit for plate detection.

The power-series expression for plate current converges too slowly for analytical purposes, and the performance is determined by test. The linearity of the output *versus* the input is shown in Fig. 28 which is a load-rectification diagram for a pentode plate detector. This diagram may be obtained by test at 60 cycles when the impedances are made the same as in actual use. Load current is plotted against r-m-s values of input voltage for the selected load and bias conditions.

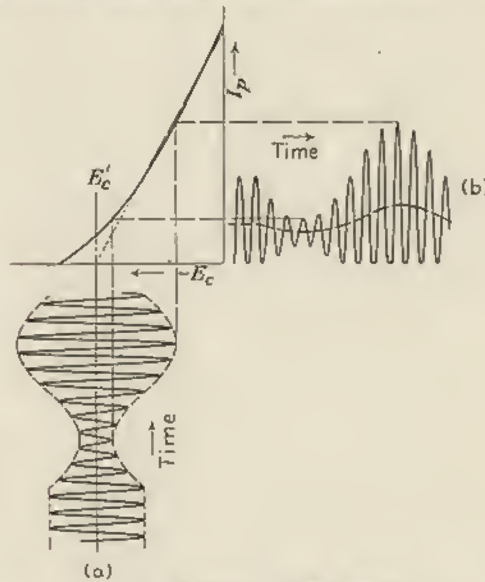


FIG. 27.—Analysis of plate detection.

The intercept of the extension of the linear portion of the grid-plate characteristic with the E_c-axis indicates the approximate value of bias voltage for maximum output and minimum distortion, as shown by the dotted line in Fig. 27.

When the tube is driven hard enough to draw grid current, the source impedance must be low or grid rectification will reduce and distort the output. Pentodes may be biased to draw no grid current over the working range.

The degree of modulation which may be handled without distortion is limited and may be calculated from the linear portion of the selected curve of Fig. 28. A slight improvement in the performance over a range of input voltages may be obtained by increasing the bias for large signals.

While the performance of a plate detector is indicated by the curves of Fig. 28, the following detector quantities are often used:
The detection plate resistance R_d is

$$R_d = \frac{\partial E_p}{\partial I_p}\bigg]_{E = E_c} \tag{72}$$

evaluated under operating conditions with a carrier E_c (or its equivalent at 60 cycles) applied to the grid. It replaces R_p in detector equations.

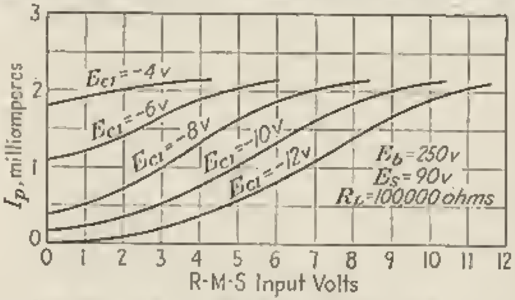


FIG. 28.—Load-rectification diagram of pentode.

The conversion transconductance S_c is

$$S_c = \frac{\partial I_p}{\partial E}\bigg]_{E = E_c} \tag{73}$$

The efficiency of detection D is

$$D = \frac{S_c R_d}{\mu} \tag{74}$$

where μ is the amplification factor and is also evaluated under operating conditions.

The change in plate current due to modulation is

$$\Delta I_p = \frac{\mu D m E_c}{Z_L + R_d} = \frac{S_c m E_c R_d}{Z_L + R_d} \tag{75}$$

where m = the degree of modulation
Z_L = the load impedance.

30. Infinite-impedance Detector. A triode self-biased nearly to plate-current cutoff by a large cathode resistance (by-passed for the carrier

frequency) passes pulses of plate current at the positive peaks of grid voltage. These pulses act in the RC circuit of Fig. 29 much like the pulses in the RC circuit of a diode with the exception that the energy is obtained from the plate circuit.

The grid does not draw current and therefore does not load the preceding circuit. The bias increases with carrier voltage and follows the modulation up to the limit where the degree of modulation is

$$m = \frac{I_0 + I_s}{I_s\sqrt{R^2a^2C^2 + 1}} \tag{76}$$

where I_0 = plate current for zero signal
I_s = increase in plate current with signal.

A proper choice of R and C therefore permits full modulation without peak clipping.

The shunting effect of any impedance in the grid circuit of the following tube can be made negligible, since R is smaller than in diode circuits.

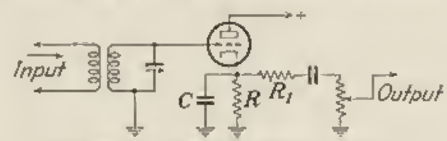


FIG. 29.—Infinite-impedance detector.

The operation is linear up to an output level limited only by the plate-supply voltage. A disadvantage of this circuit is its inability to supply voltage for conventional a-v-c circuits.

DETECTORS FOR P-M AND F-M WAVES

31. Conversion to A.M. Phase-modulated or frequency-modulated waves are detected after being converted to a-m waves. In general, a.m. is produced when an f-m wave is applied to a circuit of which the amplitude characteristic or the phase characteristic is non-uniform over the range of applied component side frequencies. In particular, when the amplitude and phase characteristics are linearly variable with frequency, a.m. proportional to the original f.m. is obtained.

A current of the form indicated in Eq. (5) applied to such a circuit produces a voltage

$$e = Z_0 I_0[1 + Sma \cos(at - Pa)] \cos[\omega_c t + m \sin(at - Pa)] \tag{77}$$

where S = slope of the impedance characteristic $\Delta Z/\Delta\omega$
P = slope of the phase characteristic $\Delta\phi/\Delta\omega$
Z_0 = impedance at the carrier frequency.

Such an impedance is obtained approximately, over a limited frequency range, on the side of the resonance curve of a parallel-tuned circuit or near the resonant frequency of a series-tuned circuit. The resultant voltage is amplitude-modulated at the modulation frequency $a/2\pi$ to a degree Sma. The phase shift of the modulation envelope and of the remaining f.m. or p.m. by the angle Pa is usually of no interest. The actual demodulation is made ordinarily in a conventional linear amplitude detector.

For f-m waves the modulation index m_f is inversely proportional to the modulating frequency $a/2\pi$ and directly proportional to the depth of modulation $\Delta\omega$.

$$m_f = \frac{\Delta\omega}{a} \tag{78}$$

The depth of amplitude modulation m_a obtained is

$$m_a = Sm_f a = S(\Delta\omega) \tag{79}$$

For p-m waves the modulation index (m_p) is independent of the modulating frequency. The output is then distorted unless the demodulated signal is passed subsequently through a circuit whose response is inversely proportional to the modulation frequency.

The linear slope filter is a true frequency-amplitude converter. It may be used as a phase-amplitude converter when followed by a corrective network.

32. Frequency Detectors. A single frequency-amplitude converter of any type responds to spurious a.m. present in the original signal.

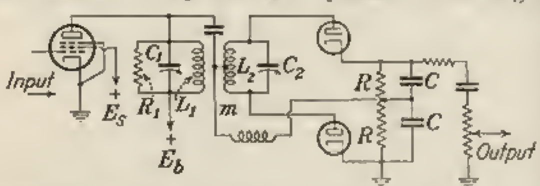


FIG. 30.—Discriminator or frequency detector.

The differentially combined outputs of two detectors, operated from converters with opposite slopes, produce a net output which is linear with respect to frequency but zero at the carrier frequency. This arrangement gives no output from a pure a-m wave when the circuit is carefully tuned to the carrier. During the reception of desired f-m waves, spurious a.m. is also detected, but with a lower output than with a single detector.

The response to spurious a.m. may be further reduced by an amplitude limiter ahead of the frequency-amplitude converter.

Figure 30 shows a typical discriminator for the detection of wide-band f.m., similar to the type used for a.f.c., but designed to be linear over the required frequency-deviation range. It combines two opposite slope converters in one device and may be operated from a single i-f amplifier or limiter. The primary voltage plus half the secondary voltage is applied to one diode, and the primary voltage plus half the secondary voltage, in reversed polarity, to the other diode. The difference between the two rectified outputs is obtained by the series connection shown.

Typical response curves for discriminators are shown in Fig. 31. These curves may be obtained by observing the d-c output voltage over the required range of frequency deviation at a definite level of unmodulated carrier voltage applied to the grid of the previous tube under selected operating conditions. The separation of the peaks in the characteristic is determined by the mutual inductive reactance X_m. The linearity is controlled by loading the tuned circuits with a resistor

R_1 or by adjusting the diode resistors R-R. Curve a shows the output with too little loading, and curve b, the output with excessive loading.

33. Phase Detectors. It is necessary to use a frequency detector for p-m waves, followed by a corrective network, when the modulation index is greater than about 0.5 radian, in order to avoid distortion of the demodulated output.

When the modulation index is small, a p-m wave may be combined with an auxiliary carrier synchronized 90 deg. out of phase with the modulated carrier. The composite wave is then amplitude-modulated by the reverse of the process indicated in Fig. 15.

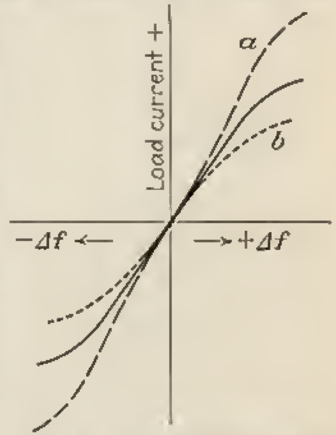


FIG. 31.—Discriminator characteristic.

MISCELLANEOUS APPLICATIONS

34. Grid-bias Amplitude Limiters. An overloaded class C amplifier with grid-leak bias may be used as an amplitude limiter at low and intermediate frequencies. The connections and representative input-output curves are shown in Figs. 32 and 33, respectively. The tube is operated at low screen and plate voltages to prevent excessive current in the absence of a signal and to provide a low overload point.

The resistance R is selected to provide bias, due to pulses of grid current, at a rate which forces the tube to plate-current cutoff over greater portions of the i-f cycle as the input voltage is increased. The exact value of resistance required to give a uniform output over a range of input depends on the impedance of the plate load and on the supply voltages. Curves a and b (Fig. 33) illustrate the output with too small and too large resistances, respectively.

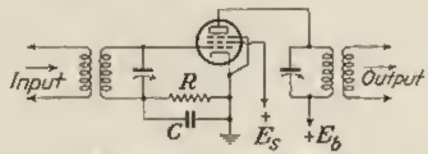


FIG. 32.—Grid-bias limiter.

The grid current loads the input circuit, and the curves are obtained by applying the input voltage to the grid of the preceding tube. The capacitor C is made as small as is consistent with over-all gain so that the grid bias may follow rapid changes in the amplitude of the input.

35. Diode Limiters. Diodes may be used as limiters either in series or in parallel with the load. Figures 34a and b show two typical examples of many possible arrangements.

In Fig. 34a, when the resistances of the source and load are equal, the first diode passes current when the input voltage is more positive than

$-E/2$, and the second diode when the input voltage is less than $+E$. No current reaches the load outside of these limits except through capacity coupling between the diode elements. It is desirable that the resistances be large compared to the forward resistances of the diodes.

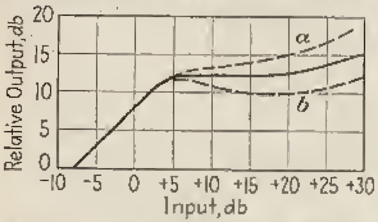


FIG. 33.—Grid-bias limiter characteristic.

The circuit is suitable at low frequencies of input signal where the time constant of the circuit capacitances with the resistances is small compared to the period of the wave.

In Fig. 34*b* each diode is nonconducting as long as the output voltage is less than $\pm E$. For larger values of voltage ($\pm$), one diode or the other shunts the load with its low forward resistance. For effective limiting, the resistances of the source and load must be considerably larger than the diode resistance.

Limiting stages may be used in cascade, alternated with amplifying stages, to obtain nearly rectangular i-f wave forms.

In both arrangements shown in Fig. 34 a small bias voltage may be needed for symmetrical limiting on both half waves to neutralize the zero-current diode voltage. When limiting in one polarity only is needed, one diode may be omitted.

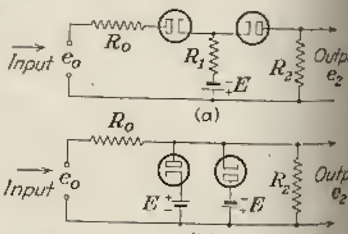


FIG. 34.—Types of diode limiters.

36. Threshold Limiters. A typical example of a threshold limiter for quiet automatic volume control is illustrated in Fig. 35. In the absence of a signal the first diode is biased beyond cutoff by the d-c drawn through the second diode. When the signal is sufficient to develop a voltage drop in R equal to $-E$, the second diode is biased beyond cutoff while the first functions as a normal detector without bias. At signal levels at which each diode carries some current, there is peak clipping when the signal is fully modulated.

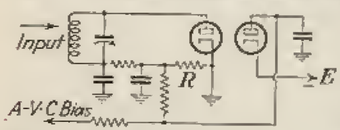


FIG. 35.—Threshold limiter for Q.A.V.C.

37. Frequency Multipliers. Any tube with a non-linear grid-plate characteristic contains harmonics of the input frequency in its plate circuit. The output circuit may be tuned to the desired multiple of the original frequency and the other components by-passed to ground. Frequency multipliers are used only for constant-amplitude or telegraph waves.

When an a-m voltage wave

$$e = E(1 + m \cos at) \cos wt$$

is applied to a square-law tube, the third term of Eq. (16) yields a modulated current wave at twice the original carrier frequency

$$i = \frac{A_2E^2}{2}\left(1 + 2m \cos at + \frac{m^2}{2} + \frac{m^2}{2} \cos 2at\right) \cos 2wt \qquad (80)$$

The new degree of modulation m' for the fundamental is

$$m' = \frac{4m}{2 + m^2} \qquad (81)$$

In addition there is a second harmonic modulation

$$m'' = \frac{m^2}{2 + m^2} \qquad (82)$$

When a p-m or an f-m voltage wave

$$e = E \cos (w_c t + m \sin at)$$

is applied to a square-law tube, the third term in Eq. (10) yields a modulated current wave at twice the original carrier frequency

$$i = \frac{A_2E^2}{2} \cos (2\omega_c t + 2m \sin at) \qquad (83)$$

An expansion of the power-series expression for plate current shows that the degree of modulation of an a-m wave is increased by each multiplication and intolerable distortion results. The modulation index m and the carrier frequency are multiplied by the same figure in the successive frequency multiplications of p-m or f-m waves.

Class A multipliers, in which the plate current is never cut off, are inefficient. More economical use of the power capability of the tube is realized in class C service where the grid-bias and input voltage are adjusted for maximum output. The expansion of the power series Eq. (16) then contains many frequency terms of more than twice the original carrier frequency.

The proper operating conditions are best determined by test. The bias is approximately that for plate-current cutoff in the absence of signal in doublers and somewhat greater than this for triplers. A slight amount of feedback at the harmonic frequency increases the output and the plate efficiency.

Higher harmonics than the third may be selected for laboratory work although the available power is limited.

Two tubes may be used in push-pull with increased efficiency. The outputs are connected in parallel for doublers and in push-pull for triplers.

References

Detection:

AIKEN, C. B.: The Detection of Two Modulated Waves Which Differ Slightly in Carrier Frequency, *Proc. I.R.E.*, **19**, 120–137, January, 1931.

——: Further Notes on the Detection of Two Modulated Waves Which Differ Slightly in Carrier Frequency, *Proc. I.R.E.*, **20**, 569–578, March, 1932.

——: Theory of the Detection of Two Modulated Waves by a Linear Rectifier, *Proc. I.R.E.*, **21**, 601–629, April, 1933.

BALLANTINE, S.: Detection by Grid Rectification with the High-vacuum Triode, *Proc. I.R.E.*, **16**, 593–613, May, 1928.
———: Detection at High Signal Voltages—Plate Rectification with the High-vacuum Triode, *Proc. I.R.E.*, **17**, 1153–1177, July, 1929.
BENSON, S.: Notes on Large Signal Diode Detection, *Proc. I.R.E.*, **25**, 1565–1572, December, 1937.
CHAFFEE, J. G.: The Detection of Frequency Modulated Waves, *Proc. I.R.E.*, **23**, 517–540, May, 1935.
———, and G. H. BROWNING: A Theoretical and Experimental Investigation of Detection for Small Signals, *Proc. I.R.E.*, **15**, 113–153, February, 1927.
DAY, J. R.: A Receiver for Frequency Modulation, *Electronics*, **12**, No. 6, 32–35, June, 1939.
FYLER, G. W., and J. A. WORCESTER, JR.: A Noise-free Radio Receiver, *Gen. Elec. Rev.*, **42**, 307–310, July, 1939.
GRONDAHL, L. O., and W. P. PLACE: Copper-oxide Rectifier Used for Radio Detection and Automatic Volume Control *Proc. I.R.E.*, **20**, 1599–1614, October 1932.
JARVIS, K. W.: Linear Detector Distortion, *Electronics*, **7**, 386–387, December, 1931.
KILGOUR, C. E., and J. M. GLESSNER: Diode Detection Analysis, *Proc., I.R.E.*, **21**, 930–943, July, 1933.
NELSON, J. R.: Grid Circuit Power Rectification, *Proc. I.R.E.*, **19**, 489–500, March, 1931.
———: Some Notes on Grid Circuit and Diode Rectification, *Proc. I.R.E.*, **20**, 989–1003, June, 1932.
ROBINSON, G. D.: Test Procedure for Detectors with Resistance Coupled Output, *Proc. I.R.E.*, **19**, 806–811, May, 1931.
RODER, H.: Some Notes on Demodulation, *Proc. I.R.E.*, **20**, 1946–1961, December, 1932.
———: Effects of Tuned Circuits upon a Frequency Modulated Signal, *Proc. I.R.E.*, **25**, 1617–1647, December, 1937.
———: Theory of the Discriminator Circuit for Automatic Frequency Control, *Proc. I.R.E.*, **26**, 590–611, May, 1938.
STRUTT, M. J. O.: Anode Bend Detection, *Proc. I.R.E.*, **23**, 945–957, August, 1935.
TERMAN, F. E., and N. R. MORGAN: Some Properties of Grid Leak Power Detection, *Proc. I.R.E.*, **18**, 2160–2175, December, 1930.
WEEDEN, W. N.: New Detector Circuit, *Wireless World*, **40**, 6–8, January, 1937.
WHEELER, H. A.: Design Formulas for Diode Detectors, *Proc. I.R.E.*, **26**, 745–780, June, 1938.
WOODYARD, J. R.: Application of the Autosynchronized Oscillator to Frequency Demodulation, *Proc. I.R.E.*, **25**, 612–619, May, 1937.

General Modulation Topics:

BARROW, W. L.: Contribution to the Theory of Nonlinear Circuits with Large Applied Voltages, *Proc. I.R.E.*, **22**, 964–980, August, 1934.
ESPLEY, D. C.: Harmonic Production and Cross Modulation in Thermionic Valves with Resistive Loads, *Proc. I.R.E.*, **22**, 781–790, June, 1934.
FERRIS, W. R.: Graphical Harmonic Analysis for Determining Modulation Distortion in Amplifier Tubes, *Proc. I.R.E.*, **23**, 510–516, May, 1935.
HARRIS, W. A.: The Application of Superheterodyne Frequency Conversion Systems to Multirange Receivers, *Proc. I.R.E.*, **23**, 279–294, April, 1935.
KLIPSCH, P. W.: Suppression of Interlocking in First Detector Circuits, *Proc. I.R.E.*, **22**, 699–708, June, 1934.
MCILWAIN and BRAINERD: "High Frequency Alternating Currents," John Wiley & Sons, Inc. 1931.
MORECROFT, J. H.: "Principles of Radio Communication," John Wiley & Sons, Inc., 1933.
NESSLAGE, C. F., E. W. HEROLD, and W. A. HARRIS: A New Tube for Use in Superheterodyne Frequency Conversion Systems, *Proc. I.R.E.*, **24**, 207–218, February, 1936.
SHEAFFER, C. F.: A Volume Limiter, *Electronics*, **10**, No. 12, 20–21, December, 1937.
SMITH, C. E.: Frequency Doubling in a Triode Vacuum Tube Circuit, *Proc. I.R.E.*, **21**, 37–50, January, 1933.
STERKY, H.: Frequency Multiplication and Division, *Proc. I.R.E.*, **25**, 1153–1173, September, 1932.
STRUTT, M. J. O.: On Conversion Detectors, *Proc. I.R.E.*, **22**, 981–1008, August, 1934.
TERMAN, F. E.: "Radio Engineering," McGraw-Hill Book Company, Inc., 1937.
———, and J. H. FERNS: The Calculation of Class C Amplifier and Harmonic Generator Performance of Screen-grid and Similar Tubes, *Proc. I.R.E.*, **22**, 359–373, March, 1934.
TITTLE, H. C.: Internal Cross Modulation in Radio Receivers, *R.M.A. Eng.*, 9–14, May, 1938

VAN DER BIJL, H. J.: "The Thermionic Vacuum Tube and Its Applications," McGraw-Hill Book Company, Inc., 1920.
WAGENER, W. G.: Simplified Methods for Computing Performance of Transmitting Tubes, *Proc. I.R.E.*, **25**, 47–77, January, 1937.
WHEELER, H. A.: The Emission Valve Modulator for Superheterodynes, *Electronics*, **6**, 76–77, March, 1933.

Modulation:

ARMSTRONG, E. H.: A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation, *Proc. I.R.E.*, **24**, 689–740, May, 1936.
BERANEK, L. L.: Applications of Copper-oxide Rectifiers, *Electronics*, **12**, No. 7, 15–18, 68–71, July, 1939.
BYRNES, I. F.: Recent Developments in Low Power and Broadcasting Transmitters, *Proc. I.R.E.*, **16**, 614–651, May, 1928.
CARUTHERS, R. S.: Copper-oxide Modulators in Carrier Telephone Systems, *Trans. A.I.E.E.*, **58**, 253–260, June, 1939.
CHAFFEE, J. G.: The Application of Negative Feedback to Frequency-modulation Systems, *Proc. I.R.E.*, **27**, 317–331, May, 1939.
CHAMBERS, J. A., *et al.*: The WLW 500 Kilowatt Broadcast Transmitter, *Proc. I.R.E.*, **22**, 1151–1180, October, 1934.
CHIREIX, H.: High Power Outphasing Modulation, *Proc. I.R.E.*, **23**, 1370–1392, November, 1935.
CROSBY, M. G.: Frequency Modulation Noise Characteristics, *Proc. I.R.E.*, **25**, 472–514, April, 1937.
———: Communication by Phase Modulation, *Proc. I.R.E.*, **27**, 126–136, February, 1939.
CULVER, C. A.: Series Modulation, *Proc. I.R.E.*, **23**, 481–495, May, 1935.
DINGLEY, E. N., JR.: An Analysis of Efficient Modulation, *Electronics*, **7**, 78–81, March, 1934.
DOME, R. B.: High-efficiency Modulation System, *Proc. I.R.E.*, **26**, 963–982, August, 1938.
EASTMAN, A. V., and E. D. SCOTT: Transmission Lines as Frequency Modulators, *Proc. I.R.E.*, **22**, 878–885, July, 1934.
FAY, C. E.: The Operation of Vacuum Tubes as Class B and Class C Amplifiers, *Proc. I.R.E.*, **20**, 548–568, March, 1932.
GAUDERNACK, L. F.: Some Notes on the Practical Measurement of the Degree of Amplitude Modulation, *Proc. I.R.E.*, **22**, 819–845, July, 1934.
———: A Phase Opposition System of Amplitude Modulation, *Proc. I.R.E.*, **26**, 983–1008, August, 1938.
HAHN, W. C.: Small Signal Theory of Velocity-modulated Electron Beams, *Gen. Elec. Rev.*, **42**, 258–270, June, 1939.
———: Wave Energy and Transconductance of Velocity-modulated Electron Beams, *Gen. Elec. Rev.*, **42**, 497–502, November, 1939.
———, and G. F. METCALF: Velocity-modulated Tubes, *Proc. I.R.E.*, **27**, 106–116, February, 1939.
HARTLEY, R. V. L.: Relations of Carrier and Side-bands in Radio Transmission, *Proc. I.R.E.*, **11**, 34–55, February, 1923.
———: Transmission of Information, *Bell System Tech. Jour.*, **7**, 535–563, July, 1928.
HELLMANN, R. K.: The Modulator Bridge, *Electronics*, **11**, No. 3, 28–30, March, 1938.
HUTCHESON, J. A.: Application of Transformer Coupled Modulators, *Proc. I.R.E.*, **21**, 944–957, July, 1933.
JAFFEE, D. L.: Armstrong's Frequency Modulator, *Proc. I.R.E.*, 26, 475–481, April, 1938.
MOUROMTSEFF, I. E., and H. N. KOZANOWSKI: A "Short-cut" Method for Calculation of Harmonic Distortion in Wave Modulation, *Proc. I.R.E.*, **22**, 1090–1101, September, 1934.
PARKER, W. N.: A Unique Method of Modulation for High-fidelity Television Transmitters, *Proc. I.R.E.*, **26**, 946–962, August, 1938.
PETERSON, E., and H. P. EVANS: Modulation in Vacuum Tubes, *Bell System Tech. Jour.*, **6**, 442–460, July, 1927.
———, and L. W. HUSSEY: Equivalent Modulator Circuits, *Bell System Tech. Jour.*, **18**, 32–48, January, 1939.
———, and C. R. KEITH: Grid Current Modulation, *Bell System Tech. Jour.*, **7**, 106–139, January, 1928.
———, and F. B. LLEWELLYN: The Operation of Modulators from a Physical Viewpoint, *Proc. I.R.E.*, **18**, 38–48, January, 1930.
RAMO, S. The Electron-wave Theory of Velocity-modulation Tubes, *Proc. I.R.E.*, **27**, 757–763, December, 1939.
RODER, H.: Amplitude, Phase and Frequency Modulation, *Proc. I.R.E.*, **19**, 2145–2176, December, 1931.

———: Superposition of Two Modulated Radio Frequencies, *Proc. I.R.E.*, **20**, 1962–1970, December, 1932.
———: Analysis of Load-impedance Modulation, *Proc. I.R.E.*, **27**, 386–395, June, 1939.
Television Transmitters, *Electronics*, **12**, No. 3, 26–29, March, 1939.
TERMAN, F. E., and J. R. WOODYARD: A High-efficiency Grid-modulated Amplifier, *Proc. I.R.E.*, **26**, 929–945, August, 1938.
VANCE, A. W.: A High-efficiency Modulating System, *Proc. I.R.E.*, **27**, 506–511, August, 1939.
VAN DER POL, B.: Frequency Modulation, *Proc. I.R.E.*, **18**, 1194–1205, July, 1930.
WEIR, I. R.: Field Tests of Frequency- and Amplitude-modulation with Ultrahigh-frequency Waves, *Gen. Elec. Rev.*, **42**, Part I, 188–191, May, 1939; Part II, 270–273, June, 1939.

SECTION 11

AUDIO-FREQUENCY AMPLIFIERS

BY GLENN KOEHLER[1]

1. Classification of A-f Amplifiers. An a-f amplifier is usually defined as one which is to work in the range of frequencies from 20 to 20,000 cps. Amplifiers for this purpose may be either selective or non-selective; *i.e.*, they may be made to amplify substantially a single frequency or a range of frequencies. Ordinarily the terminology implies that the amplifier will work over a range of frequencies.

There are four general classifications of vacuum-tube amplifiers. These classifications relate to the manner in which the tube is operated with respect to its I_p-E_g characteristics. They are class A, class AB, class B, and class C.

A class A amplifier operates in such a manner that the output wave form for a single tube and any kind of output impedance is substantially the same as the input wave form. In a class A amplifier, operation must take place such that the dynamic characteristic is nearly a straight line over the complete cycle of the input e.m.f. Ordinarily the grid in a class A amplifier is not driven positive.

A class AB amplifier is operated with more grid bias than a class A, and the grid is driven positive with respect to the cathode. In this class of amplifiers the a-c plate current for each tube flows for less than the full 360 electrical degrees of the input cycle. This type usually requires some driving power. It requires two tubes in push-pull to give an output wave form that is nearly like the input wave form.

A class B amplifier is operated with sufficient grid bias to reduce the plate current almost to zero when no input voltage is applied. For a single tube the a-c plate current flows for only 180 electrical degrees of the input cycle. It requires two tubes in push-pull to produce an output wave form that is nearly like the input wave form.

A class C amplifier is operated with more than sufficient grid bias to reduce the plate current to zero when no input voltage is applied. Plate current flows for less than 180 electrical degrees of the input cycle. It requires the use of a selective circuit in the plate circuit in order to give an output wave form that is comparatively free from distortion.

2. General Requirements of an A-f Amplifier. An a-f amplifier must satisfy the following general requirements:

1. The gain of the amplifier must conform to a certain amplification-frequency characteristic.
2. The output wave form must not contain more than a certain amount of distortion that is generated in the amplifier itself.

[1] Department of Electrical Engineering, University of Wisconsin.

3. The gain of the amplifier must be such that a certain output power is obtained from a given input voltage.

4. The noise and "hum" level of the amplifier should be within a preassigned limit.

5. The gain should not vary much with the usual variations in d-c operating voltage, temperature of filaments, etc.

6. The input and output conditions should be such as to work the amplifier out of a certain source impedance into a certain load impedance.

3. Elements of an A-f Amplifier. The a-f amplifier tube acts as a power converter taking continuous power from the battery or d-c source in the plate circuit and converting this power into a-c power. The converted power is used to set up a voltage across an impedance in the plate circuit for the case of a voltage amplifier, or to supply power to a load for the case of a power amplifier. For carrying out this function, each stage of an amplifier must be furnished with an input coupling device, an output coupling device, and the necessary sources of power to actuate the tube. For the case of a multistage amplifier the input coupling device of one tube may be the output coupling device of the tube ahead of it.

CLASS A AMPLIFIERS

4. Voltage Amplification per Stage. *a. Simple Theory.* A single triode amplifier is shown in Fig. 1. The voltage-amplification theory given below applies to a tube of three elements or more when operated as a class A amplifier without external impedances in any of the elements other than the anode or plate circuit. In the simple theory the interelectrode capacitances of the tube and socket are neglected.

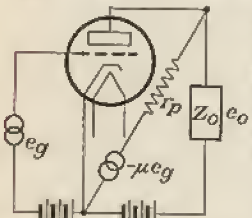


FIG. 1.—Triode amplifier.

The two important constants of the amplifier tube are the amplification constant μ and the plate resistance r_p. The tube acts as a source of alternating e.m.f. which is controlled by the input voltage e_g. This equivalent source which has a voltage $-\mu e_g$ and an internal impedance r_p sets up a.c. in the external impedance $\dot{Z}_o$. The a.c. through $\dot{Z}_o$ produces an alternating voltage across $\dot{Z}_o$ which is the output voltage e_o. The voltage amplification, or voltage gain of the amplifier is

$$\dot{A}_v = \frac{\dot{E}_o}{\dot{E}_g} = \frac{-\mu \dot{Z}_o}{r_p + \dot{Z}_o} \tag{1}$$

In this expression $\dot{Z}_o = R_o + jX_o$ and $\dot{E}_o$ and $\dot{E}_g$ are the vector values of e_o and e_g. Voltage amplification is also a vector quantity. The voltage $\dot{E}_g$ is used as the reference vector. Figure 2 shows the voltage amplification in per cent of μ plotted against ratios of output impedance to plate resistance: for cases where, $\dot{Z}_o$ is a resistance R_o, or a reactance X_o.

Because of the approximate way in which the ear responds to sound, *i.e.*, logarithmically, it is convenient to express the gain of an amplifier logarithmically. The unit is the decibel, which is equal to 20 times the common logarithm of the absolute value of the voltage ratio. Hence the gain in decibels is $20 \log_{10} |\dot{A}_v|$. The power gain in decibels can be determined from the voltage in decibels, only when the input and output impedances are known. Strictly speaking the power gain in decibels is the more fundamental quantity.

b. Effects of the Interelectrode Capacitance. The location of the interelectrode capacitances for a triode are shown in Fig. 3. These capacitances should include the tube itself and the socket. The capacitances given in the tube handbooks and manuals are usually for the tube alone. In many cases the socket interelectrode capacitances are as large as for the tube alone. When the socket capacitances are not known it is good practice to add about 4 $\mu\mu$f for adjacent electrodes and 3 $\mu\mu$f for all others except in the case where the grid comes out the top which requires no change from that given in the handbook. Multigrid tubes used as class A triode amplifiers are treated similar to the triode when there are no impedances in any of the other grid circuits.

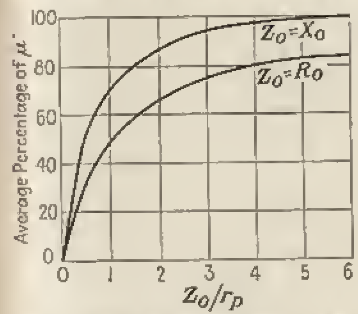


FIG. 2.—Voltage amplification of a triode.

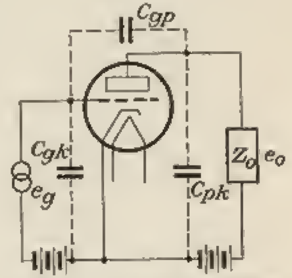


FIG. 3.—Triode amplifier showing interelectrode capacitances.

The voltage amplification $\dot{A}_v$ for the circuit of Fig. 3 is

$$\dot{A}_v = \frac{\dot{E}_o}{\dot{E}_g} = \frac{j\omega C_{gp} - G_{pg}}{G_p + \dot{Y}_o + j\omega(C_{gp} + C_{pk})} \tag{2}$$

in which $\dot{Y}_o = 1/\dot{Z}_o$, $G_{pg} = \mu/r_p$, and $G_p = 1/r_p$.

Usually the interelectrode capacitances are not very effective upon $\dot{A}_v$ over the a-f range and the susceptances $j\omega C_{gp}$ and $j\omega C_{pk}$ are negligible. Under these conditions Eqs. (2) and (1) are identical.

5. The Input Impedance. The input impedance of the tube shown in Fig. 3 is the voltage $\dot{E}_g$ divided by the current $\dot{I}_g$ that would flow in the external grid circuit. For a high vacuum tube, when operated so that the grid never goes positive, the current $\dot{I}_g$ would be the vector sum of the currents through the capacitances C_{gp} and C_{gk}. Since these two branches are effectively in parallel, it is better to consider input admittances. The expression for the input admittance is

$$\dot{Y}_i = j\omega C_{gk} + j\omega C_{gp}(1 - \dot{A}_v) \tag{3}$$

The impedance $\dot{Z}_i$ is the reciprocal of $\dot{Y}_i$. The voltage amplification $\dot{A}_r$ is a vector quantity and is obtained from Eq. (2) or (1) when the interelectrode capacitances are negligible in their effects upon $\dot{A}_r$. When the output impedance is a resistance, the value of $\dot{A}_r$ is usually a negative real quantity, and the capacitance C_{gp} is multiplied by $(1 + |\dot{A}_r|)$. Under certain conditions when the impedance $\dot{Z}_o$ has an inductive reactance, the input impedance $\dot{Z}_i$ is made up of a capacitive reactance and negative resistance. This is an important consideration in an a-f amplifier because it may cause sustained oscillations which in turn may cause very bad distortion.

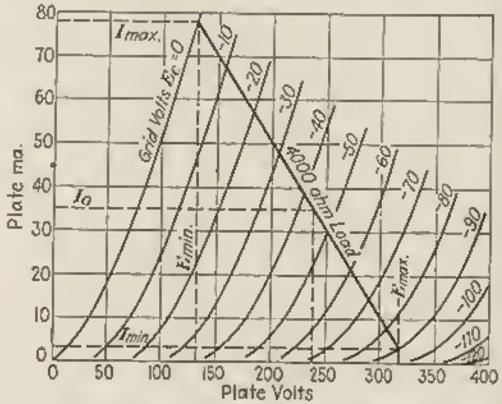


FIG. 4.—Load characteristics of a triode.

The input impedance of an amplifier tube is an important consideration when designing multistage amplifiers. As a general rule this impedance plays a part in the performance of a voltage amplifier for all frequencies above about 3,000 cps.

6. The Power Amplifier. The tube that is used to deliver power to a utilization device such as a loud-speaker is generally called a *power amplifier*. For this tube the voltage amplification is not a consideration, but the power sensitivity and the amount of power that can be converted without appreciable distortion are important. The power sensitivity is the power output in watts for a unit volt impressed on the grid.

The power sensitivity is given by the expression,

$$\text{Power sensitivity} = \frac{\mu^2 R_o}{(R_o + r_p)^2} \tag{4}$$

when the output impedance is a pure resistance R_o. The power sensitivity is a maximum and equal to $\mu^2/4r_p$ when $R_o = r_p$. However, this is not the best value of R_o for maximum undistorted power output. From theoretical considerations maximum undistorted power output is obtained when $R_o = 2r_p$ and when the peak a-c input voltage is equal to the grid-bias voltage. When $R_o = 2r_p$,

$$P_o = \frac{\mu^2 E_g{}^2}{9r_p}$$

where E_g is the r-m-s value of the a-c input voltage. For maximum undistorted power output $E_g\sqrt{2}$ is equal to the grid-bias voltage. Because the current-voltage characteristics of a tube are not straight lines, the output resistance R_o should usually be greater than $2r_p$ to limit the second-harmonic current to 5 per cent of the fundamental.

The maximum power output and second-harmonic distortion[1] can be calculated approximately for assumed values of load resistance by applying the following relations and referring to Fig. 4:

$$\text{Power output} \quad \frac{(I_{\max.} - I_{\min.}) \times (E_{\max} - E_{\min.})}{8} \tag{5}$$

$$\text{Per cent second-harmonic distortion} = \frac{\dfrac{I_{\max.} + I_{\min.}}{2} - I_o}{I_{\max.} - I_{\min.}} \times 100 \tag{6}$$

CLASS A MULTISTAGE AMPLIFIER THEORY AND DESIGN

7. Methods of Coupling. Multistage class A voltage amplifiers are usually divided into three classes as follows:

1. Resistance-capacitance coupled amplifier, illustrated in Fig. 5.

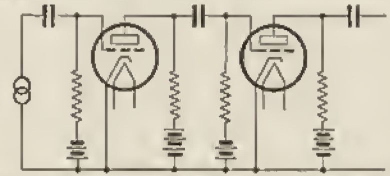


FIG. 5.—Resistance-capacitance coupled amplifier.

2. Impedance-capacitance coupled amplifier, illustrated in Fig. 6.

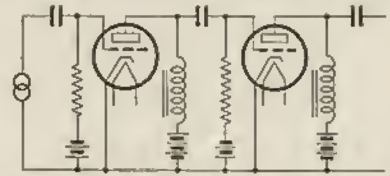


FIG. 6.—Impedance-capacitance coupled amplifier.

3. Transformer-coupled amplifier, illustrated in Fig. 7.

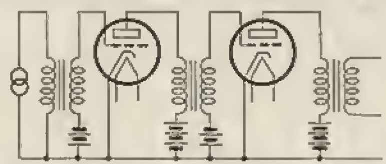


FIG. 7.—Transformer-coupled amplifier.

There are several variations of the class 2 type. The resistances in the grid circuits may be replaced by inductive impedances. In general

[1] See also Art. 52, Sec. 8.

the elements in both the plate and grid may be any type of impedances as long as they pass d-c. The more common types are the one shown and the one with simple inductive impedances in both the plate and grid. A single multistage amplifier may be a combination of these different fundamental types.

8. The Resistance-capacitance Coupled Amplifier. This class of multistage amplifiers is illustrated in Fig. 8, with the interelectrode capacitances of the tubes shown in dotted line. Consider the voltage amplification of stage 1, *i.e.*, E_{g2}/E_{g1}. Over a middle range of frequencies the voltage amplification is substantially independent of the frequency; neither the coupling condenser nor the interelectrode capacitances have

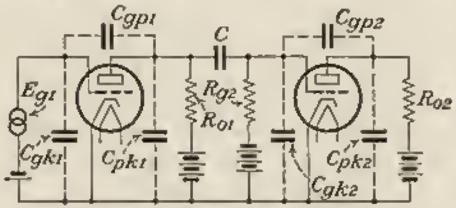


FIG. 8.—Resistance-capacitance coupled amplifier showing tube capacitances.

any effect. At the low frequencies the coupling condenser C causes the amplification to decrease with decrease in frequency, because there is a voltage drop, in C, from the plate of tube 1 to the grid of tube 2 which increases with decrease in frequency. At the high frequencies the interelectrode capacitances cause the amplification to decrease with increase in frequency because these capacitances lower the impedance in the external plate circuit of tube 1.

Frequency Characteristic. The medium-frequency gain A_M of stage 1 is

$$\frac{E_{g2}}{E_{g1}} = A_M = \frac{G_{pg1}}{G_{g2} + G_o + G_{p1}} \tag{7}$$

in which

$$G_{pg1} = \frac{\mu_1}{r_{p1}},\ G_{g2} = \frac{1}{R_{g2}},\ G_o = \frac{1}{R_{o1}},\ \text{and}\ G_{p1} = \frac{1}{r_{p1}}$$

20 $\log_{10} A_M$ will be used as the reference level, or zero level, to show what happens at low and high frequencies. The low-frequency gain, A_L, in terms of medium-frequency gain is

$$A_L = \frac{A_M}{\sqrt{1 + (G_e/\omega C)^2}} \tag{8}$$

in which $G_e = \dfrac{G_{g2}(G_o + G_{p1})}{G_o + G_{p1} + G_{g2}}$ and C is the capacitance of the coupling condenser between stages 1 and 2. The loss at low frequencies, due to C, is equal to $20 \log_{10}\sqrt{1 + (G_e/\omega C)^2}$. The curves of Fig. 9 show the relation between C and G_e for particular decibel losses at a frequency of 50 cps. The curves may be used to predict the decibel loss due to C at any other frequency f_x by multiplying the ordinates by $50/f_x$ and locating the known value of C on the new scale. Both scales may be changed simultaneously by multiplying by a

factor x in order to provide a more convenient range for G_e. To illustrate the use of the curves, suppose $r_{p1} = 100{,}000$, $R_{o1} = 200{,}000$, and $R_{g2} = 500{,}000$,

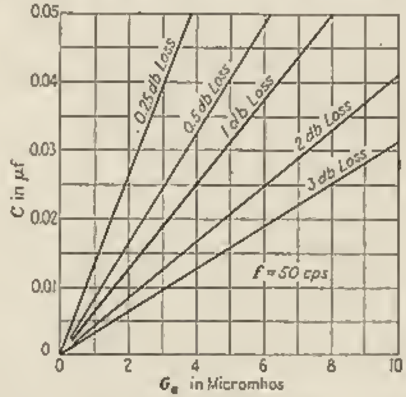


FIG. 9.—Loss in low-frequency amplification due to coupling condenser.

then $G_e = 1.76 \times 10^{-6}$. For 0.5 db loss at 50 cps. it requires a coupling condenser C equal to 0.0125 μf.

The high-frequency gain, A_H, is

$$A_H = \frac{A_M}{\sqrt{1 + (\omega C_e/G_e')^2}} \tag{9}$$

in which $C_e \cong C_{gp1} + C_{pk1} + C_{gk2} + C_{gp2}(1 + |A_{v2}|)$ (see Fig. 8), and $G_e' = G_{p1} + G_o + G_{g2}$.

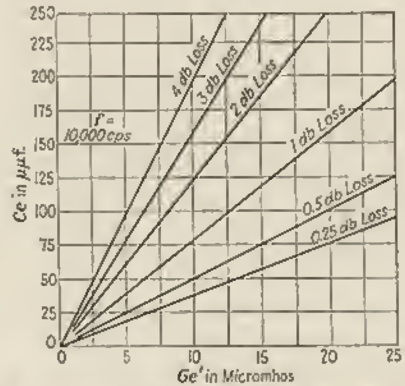


FIG. 10.—Loss in high-frequency amplification due to interelectrode capacitances.

The loss due to the shunting action of the effective capacitance C_e at the high frequencies is $20 \log_{10}\sqrt{1 + (\omega C_e/G_e')^2}$. The curves of Fig. 10 show the relation between C_e and G_e' for various decibel losses at a frequency of 10,000 cps. For a frequency f_x, multiply the present ordinates by $10{,}000/f_x$ and locate the capacitance C_e on the new scale. Suppose C_e is equal to 84 $\mu\mu f$, then for the values given in the example above $G_e' = 17 \times 10^{-6}$ and the loss at 10,000 cps is about 0.5 db.

In an amplifier of this type there is some phase distortion at both the highest and the lowest frequencies which the amplifier will pass without appreciable loss. The change in the phase angle of the voltage amplification with the frequency is, at low frequencies $\theta_L = \tan^{-1} G_e/\omega C$, and at high frequencies it is $\theta_H = \tan^{-1} \omega C_e/G_e'$. The phase shift in the amplifier, *i.e.*, the angle of departure of A_v from 180°, is illustrated by Fig. 10a. This figure also shows how the decibel loss below the gain at the medium frequencies depends upon $\omega C_e/G_e'$ at the high frequencies and $G_e/\omega C$ at the low frequencies.

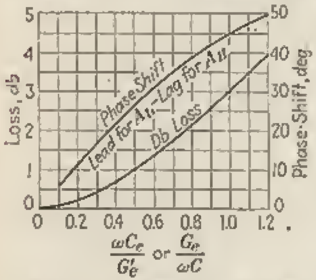


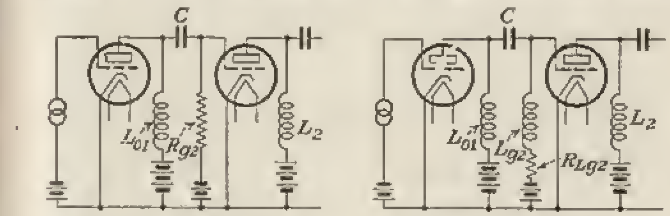
FIG. 10a.—Decibel loss and phase shift for resistance-capacitance coupled amplifier.

9. Design of a Resistance-capacitance Coupled Amplifier. When considering the proposition of using a certain tube in stage 1 (Fig. 8), to drive tube 2 and also give a preassigned amount of gain for stage 1, the first question is what will be the response at the highest frequency to be amplified. This question is settled by determining the effective capacitance C_e (it is assumed that A_{v2} is known) and using the curves of Fig. 10 to find the value of G_e' for the allowable loss at the highest frequency. This value of G_e' will determine the medium-frequency gain of stage 1 [see Eq. (7)]. In calculating C_e the interelectrode capacitances given in the tube handbooks and manuals must be increased by 3 to 5 $\mu\mu f$ to include the interelectrode socket and other stray capacitances except for the electrode that comes out the top of the tube.

To determine the size of the coupling condenser C for a preassigned response at the lowest frequency, it is necessary at this point in the design to fix the size of G_{g2}, or R_{g2}, and G_o or R_o. The following considerations are pertinent to fixing the sizes of these resistors. It is always well to use as small a coupling condenser as possible. Hence, because of the way that C depends upon G_{g2}, R_{g2} should be as large as possible but should not exceed the maximum recommended value given in the tube tables. In any event the product of $R_{g2}C$ should not exceed approximately 0.05 because of the tendency of C to become charged from a very small grid current and thereby cause the grid bias to become shifted. For a given value of $G_o + G_{g2}$, and this is fixed when G_e' is fixed for a given tube, it is well to make R_o somewhat higher than the plate resistance r_{p1} to reduce distortion if the tube is worked very hard. On the other hand, R_o consumes d-c voltage which must be supplied by the plate-voltage source.

10. Impedance-capacitance Coupled Amplifier. Under this classification of multistage amplifiers would fall almost any type of coupling except transformer coupling. Resistance-capacitance coupling has special characteristics and is therefore treated under Art. 8. The usual accepted types of the classification herein discussed are the two shown in Figs. 11a and b. The type shown in Fig. 11b is sometimes called

double-impedance coupled. These types have frequency characteristics inferior to the resistance-capacitance coupled amplifier but possess some other advantages. For example, it requires less B supply voltage to give the same plate voltage because of the much lower d-c voltage drop in the plate circuit. By a double-impedance scheme the gain at the low frequencies can be made higher than the gain at intermediate frequencies. This is sometimes useful in frequency-response equalization.



(a)-Impedance-Resistance Capacitance Coupling

(b)-Double Impedance Capacitance Coupling

FIG. 11.—Impedance-capacitance coupled amplifier.

For the type shown in Fig. 11a the voltage amplification for stage 1 at medium frequencies is

$$A_M = \frac{E_{g2}}{E_{g1}} = \frac{G_{pg1}}{G_{p1} + G_{g2}} = \frac{\mu_1 R_{g2}}{r_{p1} + R_{g2}} \tag{10}$$

in which $G_{pg1} = \mu_1/r_{p1}$ and $G_{p1} = 1/r_{p1}$ for the tube of stage 1 and $G_{g2} = 1/R_{g2}$. In some cases it may be necessary to add the core-loss conductance for L_{o1} to G_{g2}. The voltage amplification at low frequencies in terms of A_M is rather involved. It is

$$A_L = \frac{A_M}{\sqrt{1 + \left(\frac{r_{p1}R_{g2}}{r_{p1} + R_{g2}}\right)^2\left[\frac{1}{\omega^2 L_{o1}^2}\left(1 + \frac{1}{R_{g2}^2\omega^2 C^2}\right) - \frac{2}{R_{g2}^2\omega^2 L_{o1} C} + \frac{1}{r_{p1}^2 R_{g2}^2 \omega^2 C^2}\right]}} \tag{11}$$

When $C \gtreqless 0.05\ \mu f$ and $R_{g2} \gtreqless 0.5$ megohm and $f \gtreqless 50$ cps, this equation reduces to

$$A_L \cong \frac{A_M}{\sqrt{1 + \frac{1}{\omega^2 L_{o1}^2}\left(\frac{r_{p1}R_{g2}}{r_{p1} + R_{g2}}\right)^2}} \tag{12}$$

From Eq. (12) it is seen that there is a loss in amplification at the low frequencies. The loss in amplification in decibels due to insufficient reactance in choke L_{o1} is equal to

$$20 \log_{10} \sqrt{1 + \frac{1}{\omega^2 L_{o1}^2}\left(\frac{r_{p1}R_{g2}}{r_{p1} + R_{g2}}\right)^2}$$

The curves in Fig. 21 in Art. 16 may be used to get the relation between L_{o1} and $r_{p1}R_{g2}/(r_{p1} + R_o)$ for a given decibel loss at 50 cps by substituting

L_{e1} for L_m and $r_{p1}R_{g2}/(r_{p1} + R_{g2})$ for R. This holds true as long as the loss is not less than 0.5 db.

At the high frequencies the voltage amplification, A_H, is

$$A_H = \frac{A_M}{\sqrt{1 + \left(\frac{\omega C_e}{G_e}\right)^2}} \tag{13}$$

in which C_e is the effective capacitance due to the tubes (see Art. 8), plus the distributed capacitance of the choke, and G_e equals $G_{p1} + G_{g2}$ plus a conductance $1/R_e$ due to the core loss of the choke. The relation between C_e and G_e at 10,000 cps is the same as that given by the curves of Fig. 10. (See explanation in Art. 8 for extending the range of the curves.)

The type of amplifier illustrated in Fig. 11*b* has some interesting characteristics. The medium-frequency amplification is $A_M = \mu_1$, neglecting the core losses of the two coils. For the case in which ωL_{e1} is several times R_{lg2} and is at least three times r_{p1}, the amplification per stage at low frequencies in terms of that at medium frequency is

$$A_L = \frac{A_M}{\sqrt{\left(\frac{f_r}{f}\frac{1}{Q}\right)^2 + \left(1 - \frac{f_r^2}{f^2}\right)^2}} \tag{14}$$

where $f_r = \dfrac{1}{2\pi\sqrt{L_{g2}C}}$

$$Q = \frac{\omega_r L_{g2}}{r_{p1} + R_{lg2}}$$

Using the medium-frequency gain as the reference and plotting

$$20 \log_{10} \sqrt{\left(\frac{f_r}{f}\frac{1}{Q}\right)^2 + \left(1 - \frac{f_r^2}{f^2}\right)^2}$$

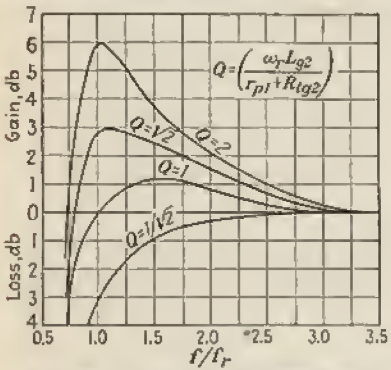


FIG. 12.—Low-frequency characteristic of a double impedance-capacitance coupled amplifier.

as ordinates and f/f_r as abscissas for various values of Q, the curves of Fig. 12 result. These curves explain the characteristics of this type of coupling and furnish quantitative information on how to fix the values of L_{g2} and C for a particular performance at the low frequencies. At the frequency f_r the gain, or loss, in decibels is equal to 20 $\log_{10} Q$. The curves also show how the gain, or loss, varies with the frequency f for a particular case. The phase distortion at low frequencies would be very bad for an amplifier of this type.

At the high frequencies the amplification per stage, A_H, is

$$A_H = \frac{A_M}{\sqrt{1 + (\omega C_e'/G_e')^2}} \tag{15}$$

where $G_e' = 1/r_{p1}$ plus the conductances due to the core losses in the two chokes

$C_e = C_{pk1} + C_{gp1} + C_{gk2} + C_{gp2}(1 + |A_{r2}|)$ plus the effective distributed capacitances of the two chokes.

The quantitative relation between C_e and G_e for different decibel losses at 10,000 cps can be obtained from the curves of Fig. 10. (See explanation in Art. 8 for extending the ranges or finding values at another frequency.)

RESISTANCE-COUPLED AMPLIFIER CHART
RCA Receiving Tube Manual

C = blocking condenser in μf
Cc = cathode by-pass condenser in μf
Cd = screen by-pass condenser in μf
Ebb = plate-supply voltage in volts
Eo = voltage output in peak volts
Rc = cathode resistor in ohms
Rd = screen resistor in megohms
Rg = grid resistor in megohms
RL = plate resistor in megohms
V.G. = voltage gain

2A6, 2B7: See 6SQ7 and 6B8, respectively.
6A6*, 6B6-G, 6B7: See 6N7, 6SQ7, and 6B8, respectively.
6B8, 6B8-G, 12C8, 6B7, 2B7:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rd	0.5	1.1	2.8	0.5	1.18	1.2	1.5	2.8	0.55	1.2	2.9
Rc	2,200	3,500	6,000	1,200	1,900	2,100	2,200	3,500	1,100	1,600	2,500
Cd	0.07	0.04	0.04	0.08	0.05	0.06	0.05	0.04	0.09	0.06	0.05
Cc	3	2.1	1.55	4.4	2.7	3.2	3	2	5	3.5	2.3
C	0.01	0.007	0.003	0.015	0.01	0.007	0.003	0.003	0.015	0.008	0.003
Eo[2]	28	33	29	52	39	55	53	55	89	100	120
V.G.[3]	33	55	85	41	55	69	83	115	47	79	150

6C5, 6C5-G, (6C6, 6J7, 6J7-G, 6J7-GT, 6W7-G, 12J7-GT, 57 as triodes):

Ebb	90			180					300		
RL	0.05	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	3,400	6,400	14,500	2,700	3,900	5,300	6,200	12,300	2,600	5,300	12,300
Cc	1.62	0.84	0.4	2.1	1.7	1.25	1.2	0.55	2.3	1.3	0.59
C	0.025	0.01	0.006	0.03	0.035	0.015	0.008	0.008	0.04	0.015	0.008
Eo[2]	17	22	23	45	41	51	55	52	70	84	85
V.G.[3]	9	11	12	11	12	12	13	13	11	13	14

6C6: As pentode, see 6J7; as triode, see 6C5.
6C8-G (one triode unit)†:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc	3,700	7,870	15,000	3,080	5,170	6,560	7,550	12,500	2,840	6,100	11,500
Cc	1.48	0.81	0.43	1.84	1.25	0.95	0.85	0.5	2.01	0.96	0.48
C	0.0115	0.0065	0.0035	0.012	0.012	0.007	0.0035	0.004	0.013	0.0065	0.004
Eo[2]	17	19	20	40	35	45	50	44	73	80	83
V.G.[3]	20	23	24	22	24	25	26	26	23	26	27

[1] For following stage.
[2] Voltage across Rg at grid-current point.
[3] Voltage gain at 5 volts r.m.s. output.
* Cathodes of the two units have a common terminal.
† Cathodes of the two units have separate terminals.

6F5, 6F5-G, 6F5-GT: See 6SF5.
6F8-G (one triode unit)†, 6J5, 6J5-G, 6J5-GT, 12J5-GT:

Ebb	90			180					300		
RL	0.05	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	2,070	3,940	9,760	1,490	2,330	2,830	3,230	7,000	1,270	2,440	5,770
Cc	2.66	1.29	0.55	2.86	2.19	1.35	1.15	0.62	2.96	1.42	0.64
C	0.029	0.012	0.007	0.032	0.038	0.012	0.006	0.007	0.031	0.0125	0.0075
Eo[2]	14	17	18	30	26	31	38	36	51	56	57
V.G.[3]	12	13	13	13	14	14	14	14	14	14	14

6J5, 6J5-G, 6J5-GT: See 6F8-G.
6J7, 6J7-G, 6J7-GT, 6W7-G, 12J7-GT, 6C6, 57: As triodes, see 6C5:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rd	0.44	1.18	2.6	0.5	1.1	1.18	1.4	2.9	0.5	1.18	2.9
Rc	1,100	2,600	5,500	750	1,200	1,600	2,000	3,100	450	1,200	2,200
Cd	0.05	0.03	0.05	0.05	0.04	0.04	0.04	0.025	0.07	0.04	0.04
Cc	5.3	3.2	2	6.7	5.2	4.3	3.8	2.5	8.3	5.4	4.1
C	0.01	0.005	0.0025	0.01	0.008	0.005	0.0035	0.0025	0.01	0.005	0.003
Eo[2]	22	32	29	52	41	60	60	56	81	104	97
V.G.[3]	55	85	120	69	93	118	140	165	82	140	350

6L5-G:

Ebb	90			180					300		
RL	0.05	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	2,500	4,620	10,300	2,210	3,180	4,200	4,790	9,290	2,160	4,130	9,100
Cc	1.89	1.08	0.49	2.2	1.46	1.1	1	0.54	2.18	1.1	[illegible]
C	0.03	0.015	0.0085	0.03	0.03	0.0115	0.009	0.009	0.032	0.014	0.0075
Eo[2]	18	22	22	41	36	46	50	46	68	79	80
V.G.[3]	10[c]	12[c]	12[c]	11[c]	12[c]	12[c]	12[c]	12[c]	12[c]	13[c]	13

6N7*, 6N7-G*, 6A6, 53:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc‡	2,250	4,950	8,500	1,700	2,950	3,800	4,300	6,600	1,500	3,400	6,100
C	0.01	0.006	0.003	0.015	0.015	0.007	0.0035	0.0035	0.015	0.0055	0.003
Eo[2]	19	20	23	46	40	50	57	54	83	87	94
V.G.[3]	19	22	23	21	23	24	24	25	22	24	24

[c] At 4 volts r.m.s. output. For other marks see p. 369.
‡ Values for phase-inverter service.

6P5-G, 76, 56:

Ebb	90			180					300		
RL	0.25	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	3,200	6,500	15,100	3,000	4,500	6,500	7,600	14,700	3,100	6,400	15,200
Cc	1.6	0.82	0.36	1.9	1.45	0.97	0.8	0.45	2.2	1.2	0.5
C	0.03	0.015	0.007	0.035	0.035	0.015	0.008	0.007	0.045	0.02	0.009
Eo[2]	21	23	24	48	45	55	57	59	80	95	96
V.G.[3]	7.7	8.9	9.7	8.2	9.3	9.5	9.8	10	8.9	10	10

6Q7, 6Q7-G, 6Q7-GT, 12Q7-GT:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc	4,200	7,600	12,300	1,900	3,400	4,000	4,500	7,100	1,500	3,000	5,500
Cc	1.7	1.2	0.6	2.5	1.6	1.3	1.05	0.76	3.6	1.66	0.9
C	0.01	0.006	0.003	0.01	0.01	0.005	0.003	0.003	0.015	0.007	0.004
Eo[2]	8	11	13	26	25	31	37	36	52	52	60
V.G.[3]	28[b]	32	33	33	36	38	40	40	39	45	46

6R7, 6R7-G:

Ebb	90			180					300		
RL	0.05	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	2,600	4,400	9,800	2,100	3,000	4,100	4,600	8,880	2,000	3,800	8,400
Cc	1.7	0.9	0.42	1.9	1.3	0.9	0.8	0.4	2	1.1	0.5
C	0.03	0.01	0.007	0.03	0.03	0.01	0.006	0.006	0.03	0.015	0.007
Eo[2]	18	19	18	40	35	43	46	40	62	68	62
V.G.[3]	9	10	11	9	10	10	10	10	9	10	11

6S7, 6S7-G:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rd	0.65	1.6	3.5	0.68	1.6	1.8	1.9	3.6	0.67	1.95	3.9
Rc	900	1,520	2,800	540	850	890	950	1,520	440	650	1,080
Cd	0.061	0.014	0.03	0.07	0.05	0.044	0.046	0.037	0.071	0.057	0.041
Cc	5	3.23	1.95	6.9	4.6	4.7	4.4	3	8	5.8	3.9
C	0.01	0.0055	0.0026	0.01	0.0071	0.005	0.0037	0.003	0.01	0.005	0.0029
Eo[2]	21	18	15	43	33	40	44	38	75	66	66
V.G.[3]	47[c]	66[c]	84[c]	66[c]	79[c]	104[c]	118[c]	134[c]	78[c]	122[c]	162[c]

[b] At 3 volts r.m.s. output.
[c] At 4 volts r.m.s. output. For other marks see p. 369.

6SC7*, 12SC7*:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc‡	1,960	3,750	6,300	1,070	1,850	2,150	2,400	3,420	930	1,680	2,9[illegible]
C	0.012	0.006	0.003	0.012	0.011	0.006	0.003	0.003	0.014	0.006	0.003
Eo[2]	5.9	8.6	10	24	21	28	32	32	50	55	62
V.G.[3]	23[b]	33	33	29	35	39	41	43	34	42	48

6SF5, 12SF5, 6F5, 6F5-G, 6F5-GT, 12F5-GT:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc	4,800	8,800	13,500	2,000	3,500	4,100	4,500	6,900	1,600	3,200	5,400
Cc	2.1	1.18	0.67	3.3	2.3	1.8	1.7	0.9	3.7	2.1	1.2
C	0.01	0.005	0.003	0.015	0.01	0.006	0.004	0.003	0.01	0.007	0.004
Eo[2]	5	7	10	23	21	26	32	33	43	54	62
V.G.[3]	31[b]	43[c]	46	44	48	53	57	63	49	63	70

6SJ7, 12SJ7:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rd	0.29	0.92	1.7	0.31	0.83	0.94	0.94	2.2	0.37	1.10	2.2
Rc	880	1,700	3,800	800	1,050	1,060	1,100	2,180	530	860	1,410
Cd	0.085	0.045	0.03	0.09	0.06	0.06	0.07	0.04	0.09	0.06	0.05
Cc	7.4	4.5	2.4	8	6.8	6.6	6.1	3.8	10.9	7.4	5.8
C	0.016	0.005	0.002	0.015	0.001	0.004	0.003	0.002	0.016	0.004	0.002
Eo[2]	23	18	22	60	38	47	54	44	96	88	79
V.G.[3]	68	93	119	82	109	131	161	192	98	167	238

6SQ7, 12SQ7, 2A6, 6B6-G, 75:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc	6,600	11,000	16,600	2,900	4,300	4,800	5,300	8,000	2,200	3,900	6,100
Cc	1.7	1.07	0.7	2.9	2.1	1.8	1.5	1.1	3.5	2	1.3
C	0.01	0.006	0.003	0.015	0.015	0.007	0.001	0.004	0.015	0.007	0.004
Eo[2]	5	7	10	22	21	28	33	33	41	51	62
V.G.[3]	29[b]	40[c]	44	36	43	50	53	57	39	53	60

[b] At 3 volts r.m.s. output.
[c] At 4 volts r.m.s. output. For other marks see p. 369.

6T7-G:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc	1,750	8,300	14,200	2,830	4,410	5,220	5,920	9,440	2,400	4,580	8,200
Cc	1.5	1	0.6	2.25	1.5	1.25	1.11	0.74	2.55	1.35	0.82
C	0.012	0.0075	0.0045	0.0135	0.012	0.008	0.005	0.0045	0.0135	0.0075	0.0055
Eo[2]	7.8	10	12	29	27	34	39	39	58	69	77
V.G.[3]	21[b]	30[c]	33[c]	28[c]	34[c]	30[c]	38[c]	41[c]	32[c]	40[c]	43[c]

6W7-G: See 6J7 and 6C5.

6Z7-G*:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc*	1,760	3,390	6,050	1,100	1,820	2,110	2,400	3,890	950	1,680	3,110
Cc	2.02	1.1	0.61	2.6	1.71	1.38	1.1	0.703	2.63	1.46	0.72
C	0.0115	0.006	0.003	0.0115	0.012	0.007	0.0035	0.0035	0.012	0.006	0.0035
Eo[2]	11	15	18	28	28	34	41	38	52	59	70
V.G.[3]	25	30	33	31	35	38	39	40	34	40	44

12C3, 12F5-GT, 12J5-GT: See 6B8, 6SF5, and 6F8-G, respectively.
12J7-GT, 12Q7-GT: See 6J7 and 6C5, and 6Q7, respectively.
12SC7, 12SF5, 12SQ7: See 6SC7, 6SF5, 6SJ7, and 2SQ7, respectively.
53, 55, 56: See 6N7, 85, and 6P5-G, respectively.
57, 75, 76: See 6J7 and 6C5, 6SQ7, and 6P5-G, respectively
79*:

Ebb	90			180					300		
RL	0.1	0.25	0.5	0.1	0.25			0.5	0.1	0.25	0.5
Rg[1]	0.25	0.5	1	0.25	0.25	0.5	1	1	0.25	0.5	1
Rc*	2,200	4,250	6,850	1,250	2,050	2,450	2,750	4,100	1,000	2,050	3,600
C	0.015	0.006	0.004	0.02	0.02	0.01	0.005	0.0035	0.01	0.0055	0.003
Eo[2]	8.4	9.7	12	27	26	34	40	39	57	66	75
V.G.[3]	29[c]	33	38	31	37	41	42	44	34	42	46

85, 55:

Ebb	90			180					300		
RL	0.05	0.1	0.25	0.05	0.1			0.25	0.05	0.1	0.25
Rg[1]	0.1	0.25	0.5	0.1	0.1	0.25	0.5	0.5	0.1	0.25	0.5
Rc	4,600	9,000	20,500	4,100	6,200	8,700	10,000	20,000	4,100	8,300	19,400
Cc	1.1	0.55	0.25	1.6	0.9	0.7	0.57	0.20	1.5	0.54	0.22
C	0.03	0.015	0.007	0.045	0.04	0.015	0.008	0.008	0.015	0.015	0.006
Eo[2]	19	22	23	44	37	47	50	48	74	82	84
V.G.[3]	4.9	5.4	5.5	5.2	5.3	5.5	5.5	5.7	5.5	5.7	5.7

[b] At 3 volts r.m.s. output.
[c] At 4 volts r.m.s. output. For other marks see p. 369.

11. Design of Impedance-capacitance Coupled Amplifiers. The application of the type of coupling shown in Fig. 11a to tubes of high plate resistance is limited principally by the amount of inductance that can be obtained in choke L_{o1} without a large amount of distributed capacitance. The distributed capacitance of the choke adds to the tube capacitance and therefore helps to lower the amplification at the high frequencies. Chokes for this purpose are sometimes wound in pie sections in order to reduce the distributed capacitance. Of course, for tubes having high plate resistance some of the maximum possible gain can be sacrificed by lowering R_{g2} to have a small variation in gain over the frequency range. This will make it easier to satisfy the requirements at both the highest and lowest frequency.

For tubes that have low plate resistance, the design procedure is to fix the value of R_{g2} so that it will not be greater than the maximum recommended value or the value which will keep the highest frequency response within the desired limit. The curves of Fig. 10 are useful for determining the limit to R_{g2} so far as frequency response is concerned. In this figure for this purpose G_e is equal to $G_{p1} + G_{g2}$ plus a conductance allowed for the core loss of L_{o1}. After R_{g2} is fixed, the value of L_{o1} is determined tentatively by the use of the curves in Fig. 20. For this purpose R_4 on the graph becomes $r_{p1}R_{g2}/(r_{p1} + R_{g2})$. The last step is to determine C such that the loss due to it is not more than 0.25 db. In some cases it may be necessary to check the results by applying Eq. (11).

For tubes that have high plate resistance, the design procedure is about the same as the above except it may be necessary to work back and forth from h-f consideration to l-f consideration in order to obtain the desired characteristics.

In designing an amplifier of the type shown in Fig. 11b the general procedure is the same as above. In some cases the medium-frequency amplification may be less than μ_1 because of the core losses of the two chokes. These core losses are equivalent to two resistances in parallel from the grid to the cathode of tube 2 and their effect is similar to R_{g2} in Fig. 11a.

The following example will illustrate how to apply Eq. (14) and the curves of Fig. 12. The plate resistance r_{p1} of the tube is 10,000 ohms, the allowed resistance for R_{Lg2} is 1,000 ohms, and the desired gain at 50 cps is 3 db over the gain at medium frequencies. From the curves of Fig. 12, Q must be $\sqrt{2}$ to give the desired gain. From the expression for Q,

$$L_{g2} = Q\frac{(r_{p1} + R_{Lg2})}{\omega_r}$$

L_{g2} is equal to $11{,}000/2\pi 50$ which gives 35 henrys. The size of the coupling condenser is given by $C = 1/\omega_r^2 L_{g2}$ and is equal to 0.29 μf.

12. The Equivalent Circuit of a Transformer-coupled Amplifier. The complete equivalent circuit of one stage of a transformer-coupled amplifier comprises the plate resistance of the tube ahead of the transformer, the input capacitance of the tube after the transformer, and the equivalent circuit of the transformer itself. Figure 13 illustrates the complete equivalent circuit for one stage. This circuit does not apply to all types but represents the condition quite accurately for a great many.

In this diagram the symbols shown represent the following: μE_{g1} is the voltage generated in the tube source and r_{p1} is the plate resistance of the tube source. R_p and R_s are the primary and secondary winding resistances.

L_p and L_s are the primary and secondary leakage inductances. These inductances are due to the magnetic fluxes that link with each coil and not the other, i.e., the fluxes that are not mutual to the two coils.

C_p and C_s are the effective distributed capacitances of the primary and secondary windings. C_m is the effective mutual capacitance between the windings. C_m may not be present in certain transformers. Sometimes C_m is of a complicated nature and difficult to estimate. C_L is the input capacitance of the tube load.

L_m and R_c are the magnetizing inductance and core-loss resistance of the transformer. The magnetizing current and the equivalent core-loss current of a transformer are nearly proportional to the induced voltage.

L_1 and L_2 are fictitious inductances necessary to transfer the current and voltage to the load and to provide the proper phase change from primary to secondary. The phase of the secondary voltage with respect to the primary is important when the mutual capacitance C_m is equal to, or greater than, 25 per cent of C_s and C_L. The ratio of the primary turns to the secondary turns is equal to $\sqrt{L_1/L_2}$. This ratio is called N, the ratio of transformation.

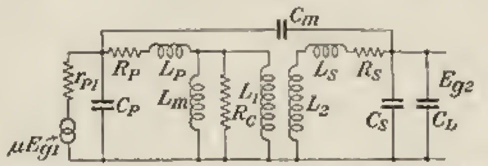


FIG. 13.—Equivalent circuit of a transformer-coupled amplifier.

In Arts. 14 and 15 it is shown how the equivalent circuit is modified in order to simplify matters. This simplification is possible for a transformer which is intended to cover a range of frequencies like 50 to 5,000 cps and when the variation in amplification over the range is not more than 6 db.

13. Calculation of Transformer Constants. The material under this article applies to both interstage transformers and impedance-matching transformers. The most important constants required in a given design are the magnetizing inductance L_m and leakage inductances L_p and L_s.

The magnetizing inductance L_m is given by the expression

$$L_m \text{ in henrys} = \frac{4\pi 10^{-9} N_p^2 \mu_r A}{l} \tag{16}$$

where N_p is the number of turns on the primary; μ_r is the relative permeability; if the primary carries d.c., μ_r is the apparent incremental permeability; A is the net area of the core in square centimeters and l is the mean length of path in centimeters. When A is not the same for the entire length of the path, the total reluctance must be calculated from the sum of the reluctances of the paths over which the net area is constant. To evaluate L_m when the winding carries d-c current, there must be available curves of μ_r plotted against the d-c magnetizing ampere-turns per centimeter for various flux a-c densities on the particular magnetic material.[1]

[1] See also Sec. 2, article on Magnetic Circuit.

The leakage inductances L_p and L_s depend upon the configuration of the windings. These inductances are due to the fluxes that link with one coil and not the other. For the type illustrated by Fig. 14,

$$L_t = L_p + N^2L_s = \frac{16\pi N_p^2}{10^9 W}\Big\{(D_1 + D_2 + 2D_o)\frac{D_t}{3} + \frac{1}{2}(D_t^2 - D_i^2) + \left[\frac{2}{3}(D_1 + D_2) + 2D_i + D_b\right]D_b\Big\} \quad (17)$$

where $D_t = D_i + D_b + D_o$.

For an interspaced winding of this type, *i.e.*, one in which one coil is placed between the two halves of the other coil, $L_p + N^2L_s$ is approximately one-fourth of that given by Eq. (17). All dimensions are in centimeters and are indicated in the figure.

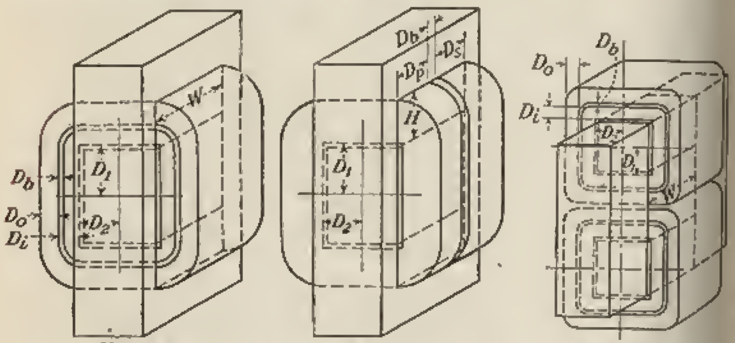


Fig. 14. Fig. 15. Fig. 16.

Fig. 14.—Simple winding scheme for a transformer.
Fig. 15.—Winding scheme for low effective capacitance.
Fig. 16.—Simple winding scheme for a core-type transformer.

The leakage inductance for a winding of the type shown in Fig. 15 is approximately

$$L_t = L_p + N^2L_s = \frac{16\pi N_p^2}{10^9 H}\left[(D_1 + D_2 + H)\left(\frac{D_p}{3} + \frac{D_s}{3} + D_b\right)\right] \quad (18)$$

For an interspaced winding of this type the total leakage inductance is approximately one-fourth of the value given by Eq. (18).

For a core-type transformer as shown in Fig. 16, in which half of each primary and secondary is wound on two opposite legs of the core, the approximate expression for the leakage inductance is

$$L_t = L_p + N^2L_s = \frac{8\pi N_p^2}{W10^9}\Big\{(D_1 + D_2 + D_o)\frac{D_t}{3} + \frac{1}{2}(D_i^2 - D_o^2) + \left[\frac{2}{3}(D_1 + D_2) + 2D_i + D_b\right]D_b\Big\} \quad (18a)$$

For an interspaced winding of the core type, *i.e.*, one in which one coil of each leg is placed between the two halves of the other coil of the same leg, the leakage inductance is approximately one-fourth the value given by Eq. (18a).

For interstage and impedance-matching transformers the core losses under most ordinary circumstances are usually small compared to the copper losses, but for the sake of completeness the expression for the core-loss resistance R_c is given. It is

$$R_c = \frac{2\pi^2 10^{-16} f^2 N_p^2 A}{K_c l} \quad (19)$$

where $K_c = \dfrac{\text{total core loss per cc}}{B^2}$ at the operating conditions

B = flux density in gausses.

It is assumed that the hysteresis losses as well as the eddy-current losses are proportional to B^2. It has been found by the author that the hysteresis losses at low flux densities are nearly proportional to B^2, but sometimes the exponent of B is even greater than 2.

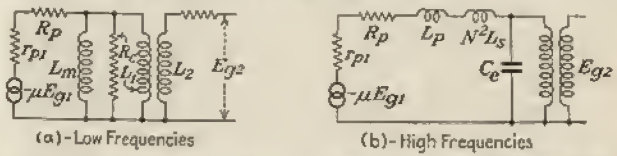


Fig. 17.—Equivalent circuits of a transformer-coupled amplifier.

The distributed capacitance of transformer windings is due mainly to the layer-to-layer capacitances. The effective capacitance of a winding is approximately equal to the capacitance between the two mean layers divided by the number of layers. In most cases the layers may be treated as parallel plates having a dielectric equal to thickness of paper between layers plus 2 times the thickness of the insulation on the wire. If the dielectric constants of the paper and insulation are much different, they must be treated accordingly.

14. Theory of Transformer-coupled Amplifiers. The characteristics of this type of amplifier are best explained by dividing the frequency range into the low frequencies, the medium frequencies, and the high frequencies. The equivalent circuits of Figs. 17*a* and 17*b* apply to the low and the high frequencies. At the medium frequencies the core-loss resistance R_c is usually so large compared to $r_{p1} + R_p + N^2R_s$ that the voltage amplification per stage, *i.e.*, $E_{g2}/E_{g1} = A_M$ is practically equal to μ_1/N. Hence 20 $\log_{10}(\mu_1/N)$ will be used as the reference level in decibels, and the performance at the low and high frequencies will be termed a loss, or gain, in decibels measured from this reference level.

At the low frequencies the magnetizing inductance is effective and the l-f amplification A_L, in terms of A_M, is

$$A_L = \frac{A_M}{\sqrt{1 + \left[\dfrac{1}{\omega L_m}\dfrac{(R_p + r_{p1})R_c}{R_c + R_p + r_{p1}}\right]^2}} \quad (20)$$

The loss at the low frequencies due to L_m is

$$20 \log_{10} \sqrt{1 + \left[\frac{1}{\omega L_m} \frac{(R_p + r_{p1})R_e}{R_e + R_p + r_{p1}}\right]^2}$$

This case is so similar to the one illustrated by the equivalent circuit of Fig. 19a for an impedance-matching transformer that the curves given in Fig. 20 may be used to see the relation between L_m and $\frac{(R_p + r_{p1})R_e}{R_e + r_{p1} + R_p}$ for various decibel losses at 50 cps. In many cases R_e is so large compared with r_{p1} and R_p that the quantity

$$\frac{(R_p + r_{p1})R_e}{R_e + R_p + r_{p1}} \cong R_p + r_{p1}$$

Hence in most cases $R_p + r_{p1}$ can be substituted for R_4 when using the curves in Fig. 20 to determine L_m for a given loss in decibels. The curves of Fig. 20 can be used for any other frequency f_x by multiplying the ordinates by $50/f_x$ and locating L_m on the new scale.

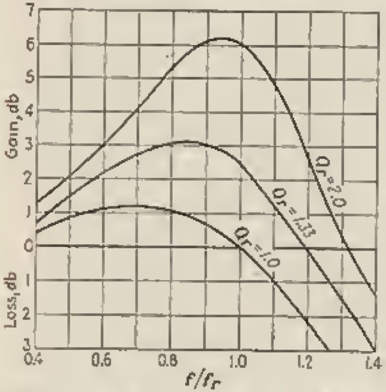


Fig. 18.—High-frequency characteristics of a transformer-coupled amplifier.

At the high frequencies the leakage inductances and the tube and distributed capacitances affect the voltage amplification. For cases in which C_m is small compared to $C_s + C_L$, the amplification at the high frequencies in terms of A_M is

$$A_H = \frac{A_M}{\sqrt{\left(1 - \frac{f^2}{f_r^2}\right) + \frac{f^2}{f_r^2}\frac{1}{Q_r^2}}} \tag{21}$$

The gain, or loss, equals

$$20 \log_{10} \sqrt{\left(1 - \frac{f^2}{f_r}\right)^2 + \frac{f^2}{f_r^2}\frac{1}{Q_r^2}},$$

where $Q_r = \frac{\omega_r L_t}{R_e}$

$$\begin{aligned}
\omega_r &= 1/\sqrt{L_t C_e} \text{ and } f_r = 1/(2\pi\sqrt{L_t C_e}) \\
C_e &= (C_m + C_s + C_L)/N^2 \\
R_e &= r_{p1} + R_p + N^2 R_s \\
L_t &= L_p + N^2 L_s \\
C_L &= C_{gk2} \mp C_{gp2}(1 + |A_{v2}|)
\end{aligned}$$

N = the ratio of primary turns to secondary turns.

The curves of Fig. 18 show how the loss, or gain, varies around the frequency f_r for different values of $\omega_r L_t/R_e$. The best results are obtained

when $\omega_r L_t/R_e$ is approximately equal to 1. This can be accomplished to some extent by controlling L_t and C_e in the design.

When C_m is not small compared to $C_s + C_L$, the voltage amplification is approximately the value given by Eq. (21) times

$$1 + \frac{NC_m}{C_m + C_s + C_L}\frac{f^2}{f_r^2}$$

where N may be either positive or negative in numerical value. N is positive if the two coils form a single winding in one direction about the common core when connected together at the cathode ends, and negative when the windings are in opposite directions. The mutual capacitance may be avoided by the use of static shields.

15. Design of Transformer-coupled Amplifiers. Usually transformer coupling is used with voltage amplifier tubes that have a comparatively low plate resistance. This is necessary to obtain the desirable characteristics at the low frequencies because the magnetizing inductance for a given l-f response is almost directly proportional to the plate resistance of the tube. It is essential also that the d-c plate current be as small as possible so that it will not saturate the core of the transformer. The magnetizing inductance L_m is the first consideration in the design of an interstage transformer. The curves of Fig. 20 can be used for determining the value of L_m for a given decibel loss at the lowest frequency. In the preliminary procedure the core loss can be neglected and $R_p + r_{p1}$ can be substituted for R_4 in Fig. 20. An allowance of 8 to 10 per cent of r_{p1} is made for the primary winding resistance.

The amount of voltage amplification per stage required at the medium frequencies is nearly equal to the amplification constant μ_1 times the ratio of secondary turns to primary turns; in the theory this is μ_1/N. Practical values for this ratio are 2 to 4. If higher, difficulty is experienced at the high frequencies because of the tube load and distributed capacitance of the secondary windings, even though the leakage inductance is very small.

The performance of the transformer at the high frequencies depends largely upon the leakage inductance and the capacitance of the secondary winding and tube load. This is illustrated in Fig. 18. For practically constant gain up to any frequency f_h either the frequency f_r must be at least two times f_h or else the winding must be so designed that $f_r = f_h$ and the quantity $\omega_r L_t/R_e = Q_r$ is approximately equal to 1.

Interspacing the windings of a transformer, placing one winding between the two halves of the other, lowers the total leakage inductance by a factor of one-fourth but generally results in a much higher effective capacitance. Therefore the net result of interspacing is not to raise the frequency f_r by a factor of 2. Even if f_r were raised by a factor of 2, the quantity Q_r might be reduced below 1 at f_r and the gain of the amplifier would not be constant up to f_r.

Winding the transformer like Fig. 15 except with interspaced coils is very effective in reducing the capacitance of the windings, but this is very uneconomical as to space.

The theory and design given here apply to input transformers as well as interstage transformers. The input transformer must be designed for a particular source impedance and a particular tube load.

16. Impedance-matching Transformers. When a given load resistance R_l is not of the proper magnitude to result in maximum power into the load from a source which has a resistance R_i, a transformer is interposed between the source and the load. Because of the resistances of the transformer windings and the losses in the magnetic core the transformer will consume a certain amount of energy itself. In addition to the energy lost in the transformer the magnetizing current causes a loss of power to the load at the low frequencies, and the leakage inductance causes a loss at the high frequencies. For a transformer of this type, intended to cover a range of frequencies, it is convenient to divide the theory and design into three phases, namely: low frequency, medium frequency, and high frequency. Figures 19*a*, *b*, and *c* represent the equivalent circuits that apply to each of these phases of discussion.

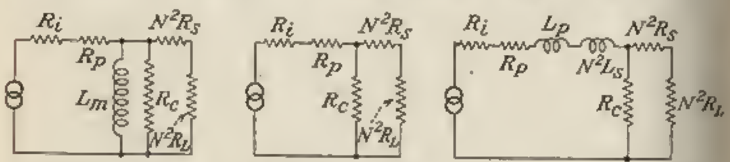


Fig. 19.—Equivalent circuits of an impedance-matching transformer.

In the figures R_i is the internal resistance of the source; R_p and R_s are the primary and secondary winding resistances; L_p and L_s are the leakage-flux inductances of the primaries and secondaries; L_m and R_c are the magnetizing inductance and core-loss resistance; and N is the ratio of primary turns to secondary turns.

The current in the transferred load resistance at the medium frequency is used as the reference level. Referring to Fig. 19*b* and letting $R_1 = R_i + R_p$,

$$R_2 = (R_s + R_L)N^2, \quad R_3 = \frac{R_2}{1 + \dfrac{R_2}{R_c}}, \text{ and } R_4 = \frac{R_3 R_1}{R_3 + R_1}$$

$$I_M = \frac{E}{R_2(R_3 + R_1)/R_3} \tag{22}$$

In many cases R_2/R_c is so small compared to 1 that $I_M = E/(R_2 + R_1)$.

For the low frequencies Fig. 19*a* applies, and the current I_L in terms of I_M is

$$I_L = \frac{I_M}{\sqrt{1 + \dfrac{R_4{}^2}{\omega^2 L_m{}^2}}} \tag{23}$$

Then $20 \log_{10} \sqrt{1 + \dfrac{R_4{}^2}{\omega^2 L_m{}^2}}$ is the loss due to L_m. Figure 20 shows the relation between L_m and R_4 for various losses at a frequency of 50 cps. For any other frequency multiply the ordinates by $50/f$, and locate L_m on the new scale. Also, because of the linear relation between L_m and R_4, both scales may be changed simultaneously by any factor x in order to provide a more convenient range for R_4. For most cases, since R_c is several times R_2, the quantity R_4 is equal to $R_1/(1 + R_1/R_2)$.

For the high frequencies Fig. 19*c* applies, and the current I_H in terms of I_M is

$$I_H = \frac{I_M}{\sqrt{1 + \dfrac{\omega^2 L_t{}^2}{(R_1 + R_3)^2}}} \tag{24}$$

Then $20 \log_{10} \sqrt{1 + \dfrac{\omega^2 L_t{}^2}{(R_1 + R_3)^2}}$ is the loss due to the leakage inductance.

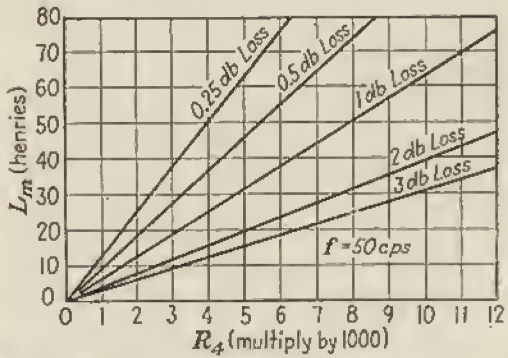


Fig. 20.—Loss at low frequency due to magnetizing inductance.

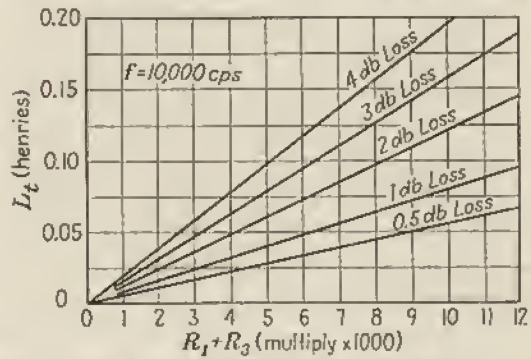


Fig. 21.—Loss at high frequency due to leakage inductance.

Figure 21 shows the relation between the total leakage inductance

$$L_t = L_p + N^2 L_s$$

and the resistance $R_1 + R_3$ for different decibel losses at 10,000 cps. For any other frequency f_x, multiply the ordinates by $10{,}000/f_x$, and read L_t on the new scale. Also both scales may be changed simultaneously by a factor x in order to provide a more convenient range for $R_1 + R_3$.

The procedure in designing a transformer of this kind is to first determine the size of core and number of primary turns in order to obtain a value of L_m which will limit the loss to a preassigned amount. In this procedure it is necessary to allow for the winding resistances R_p and R_s. The expression for L_m is given in Art. 13. The next step is to fix the ratio of turns and the number of secondary turns for the desired value of transferred load resistance. The final step is to determine the style of winding that will keep the leakage inductance within the limit which is allowed for a given loss at the highest frequency.

PUSH-PULL POWER AMPLIFIERS: CLASS A, CLASS AB, AND CLASS B

17. Graphical Analysis for Push-pull Power Amplifiers. The circuit diagram of the push-pull type of power amplifier is shown in Fig. 22. Whenever possible, the power stage of an amplifier should be operated in push-pull. There are several advantages of push-pull operation over a single-ended power stage. When a single tube is operated so that the efficiency of power conversion is reasonably high, the harmonics are also high. In push-pull operation the even harmonics cancel out in the final load resistor. Consequently, for a given percentage of distortion, the operating voltage can be adjusted so that each tube will deliver more power into a load resistor than a similar single tube will deliver into its optimum load in a single-ended power stage.

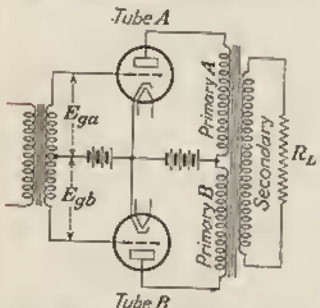


FIG. 22.—Schematic of a push-pull power amplifier.

The graphical analyses for all three classes of push-pull amplifiers are essentially the same. The magnetic field in the core of output transformer is a function of the algebraic sums of the a-c currents in the two sides of the primary windings. Hence the analysis is the same, except for d-c components, as though the tubes were replaced by a single class A tube which has I_p-E_p characteristics equivalent to the algebraic sum of the characteristics of the two push-pull tubes. These are called the *composite characteristics* of the push-pull unit and are illustrated in Fig. 23. Each composite curve represents the algebraic sum of the I_p-E_p curve of one tube for a grid potential of $E_c + \Delta E_g$ and the I_p-E_p curve of the other tube for a grid potential of $E_c - \Delta E_g$. E_c is the grid bias voltage. Then the load line, which is the load resistance measured across one primary winding of the transformer, is drawn across the composite I_p-E_p curves through the d-c operating points. Current values derived from the intersection of composite load line and the composite I_p-E_p lines are the algebraic sums of the a-c currents in the two primary windings. The effect is the same as though all the a-c current flowed through one primary winding and the source impedance were equal to the reciprocal of the slope of a composite I_p-E_p curve. Hence power output is given by the relation

$$P_0 = \frac{\mu^2 E_g^2 R_L N^2}{(R_L N^2 + r_p')^2} \tag{25}$$

where E_g = r-m-s a-c voltage from one grid to cathode
μ = amplification factor of either tube
r_p' = reciprocal of the slope of a composite I_p-E_p curve
R_L = load resistance
N = ratio of the turns of one primary winding of the output transformer to the turns of the secondary winding.

The power output can also be obtained from the peak values of a-c plate current and plate voltage which are labeled I_0 and E_0 in Fig. 23. It is $P_0 = (E_0 I_0)/2$. The distortion can be obtained by plotting the current in the load resistance and analyzing the curve according to the method given in a previous section of this handbook.

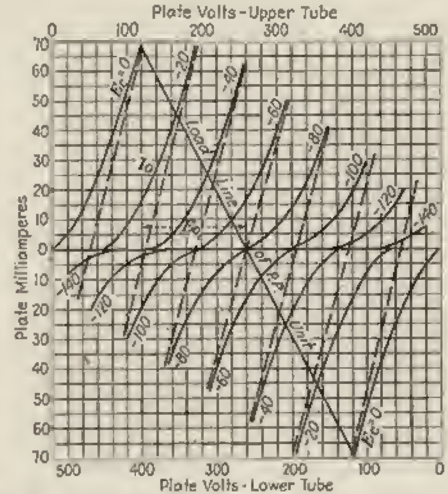


FIG. 23.—The composite characteristics of a push-pull amplifier.

18. Class A Push-pull Power Amplifier. In the class A push-pull amplifier a-c plate current flows for complete 360 deg. of the input cycle, and the characteristics of the tube are nearly straight lines over the complete range of the a-c plate and grid potentials. The composite I_p-E_p curves have approximately twice the slope of the separate I_p-E_p curves from which they are derived. Hence r_p' in Eq. (25) is approximately equal to $r_p/2$ or half the plate resistance of either tube. Then for class A push-pull

$$P_0 = \frac{\mu^2 E_g^2 R_L N^2}{\left(R_L N^2 + \frac{r_p}{2}\right)^2} \tag{26}$$

Class A operation gives the best wave form for the current in load resistor, but the efficiency is lower than that obtained by class AB or class B operation.

Equation (26) furnishes information for the design of the output transformer for the class A push-pull amplifier. Referring to the design relations and curves of Art. 16 the generator voltage is μE_g, and the resistance of the source, *viz.*, R_i, becomes $r_p/2$. The transferred load resistance is R_LN^2, where N is the ratio of turns for one primary winding to the total secondary turns. The value of R_LN^2 to be used in Eq. (26) is equal to one-fourth of the plate-to-plate load resistance which is usually specified in tube handbooks as the best value to use for a given tube. Hence the allowable values for the magnetizing and leakage inductances can be determined from Figs. 20 and 21.

The design of the input transformer becomes essentially the design of an interstage transformer which is given in Arts. 14 and 15. To preserve a balance in the magnitudes and phases of the two secondary voltages, the two secondary windings must be kept symmetrical with respect to leakage inductances to the primary, resistances, and distributed capacitances. If the two voltages applied to grids of the push-pull amplifier are less than 180 deg. apart, the net grid-to-grid voltage which will be effective in producing output power will be less than the algebraic sum of the two voltages.

19. Class AB Push-pull Power Amplifier. In the class AB push-pull power amplifier a-c plate current of each tube flows for less than 360 deg. but more than 180 deg of the input cycle. The grids may or may not be driven positive with respect to the cathode. For this type of amplifier the reciprocal of the slopes of the composite characteristics lies somewhere in between r_p and $r_p/2$, and the quantity r_p' in Eq. (25) must be determined for any particular set of operating conditions.

The design of the output transformer is carried out according to Art. 16, where R_i becomes r_p' and R_LN^2 is the load impedance which must be used in establishing L_m and L_l for each primary winding. Since the two primary windings are carrying unequal currents, care must be taken in the position of each primary winding with respect to each other and to the secondary winding. The effective leakage inductance will always be higher than it would be for the same transformer if both primary windings were carrying equal and opposite currents for all points of the input cycle.

Class AB push-pull amplifiers may be so driven that their grids go positive with respect to the cathode. Hence the input transformer design depends some on whether or not there will be grid current in the secondary during a part of each positive half cycle of the grid voltage. When there is no grid current, the design is the same as that given for the input transformer of class A push-pull amplifier. When there is grid current, the load on the driver tube varies over the cycle going from no load for a part of the cycle to a maximum load current which causes quite a drop in the grid voltage of the class AB tubes. Hence the input transformer must be so designed that magnetizing inductance will be high enough for no load conditions and have such a ratio of transformation that the output voltage of the driver tubes will not vary much over the cycle. The driver tubes should have as low plate resistance as possible. Low leakage inductance and winding resistances also help to reduce the flattening of the crest of the input voltage to class AB tubes.

20. Class B Power Amplifiers. For class B operation the d-c grid and plate potentials are adjusted so that plate current for each tube flows for only slightly more than 180 deg of the input cycle. In the graphical construction the I_p-E_p charts for the two tubes are adjusted, for a particular set of operating voltages, so that a large part of the I_p-E_p curves of the tubes coincide with the composite I_p-E_p curves. Only the low values of I_p of each tube will be different from their algebraic sum. Hence in this case $r_p' = r_p$ and the power output is

$$P_0 = \frac{\mu^2E_g^2R_LN^2}{(R_LN^2 + r_p)^2} \tag{27}$$

where r_p depends somewhat upon the amplitude of E_g and should be determined for a medium value of E_g.

For class B operation each primary of the output transformer carries current of the fundamental frequency for only alternate halves of the input cycle. Hence the effective leakage inductance of the transformer is materially higher than it would be if both windings always carried equal and opposite currents. The criterion on leakage inductance is the inductance measured across only one primary winding with the secondary winding shorted. This is the leakage inductance which enters into the characteristics of the transformer at the higher frequencies. In the design procedure given in Art. 16, r_p is the source resistance symbolized by R_i, and R_LN^2 is the load impedance transferred to one primary side. L_m and L_l are calculated or preassigned on the basis that only one primary winding is active at a time. The two primary windings should occupy similar positions with respect to the secondary and should be well interspaced with each other. The design of the input transformer is similar to that given for class AB operation. For zero-bias class B tubes the design of the input transformer is essentially the same as that of an impedance-matching transformer which is treated under Art. 16.

21. Pentode- and Beam-tube Power Amplifiers. The power sensitivity and the efficiency of power conversion for pentode and beam tubes in power amplifiers are usually much higher than for triode tubes. The expression for the power sensitivity of a pentode or a beam tube is the same as it is for a triode, as given in Art. 6. A method for determining the power output and the distortion from the I_p-E_p characteristics is given in Sec. 8, Art. 47.

The load resistor for pentode and beam tubes should be such that the instantaneous plate current does not fall below the knee of the I_p-E_p curve which is taken for the grid voltage reached on the peak value of the positive half of the input cycle. If the load resistor is higher than this value, there will be serious distortion of the output power. This limits the load resistor to a value considerably below the plate resistance of the tube. Also, if the load resistance is too low, the second harmonic distortion will be high. Consequently pentode and beam tubes should be used only when the load impedance remains fairly constant with frequency, or means should be taken to ensure that the load impedance remains within certain limits when the tube is driven for full power output. These tubes give much better results in push-pull arrangements than in single-ended circuits because of the even harmonic cancellation. Single-ended pentode and beam power amplifiers should incorporate degenerative feedback for the best results. Further points in favor of pentode and beam tubes are that they have lower input capacitances than similar triodes and require lower driving voltages.

22. Phase-inverter Amplifiers. When an amplifier requires a single-ended input, but it is desirable to have the output tubes in push-pull, it is necessary to derive voltages for the push-pull grids that are equal in magnitude and 180 deg. out of phase over the complete frequency range of the amplifier. This can be done by the use of a transformer with a single primary and center-tapped secondary. However, it is somewhat difficult to design such a transformer which will have secondary voltages from each end to the center tap that are equal in magnitude and 180 deg. out of phase over any considerable range of audio frequencies. Also it is often desirable to have resistance-capacitance coupling throughout an amplifier. This can be accomplished by the scheme of Fig. 24. Tube B is the phase-inverter tube. Its input voltage is derived from the output of tube A. The grid voltage for the phase inverter may also be derived from a portion of the grid resistor of the upper output tube or in the manner shown in Fig. 25. When all tubes are self-biased, the grid resistor and coupling condenser are not necessary when the grid voltage for the phase-inverter tube is derived from a portion of the grid resistor of the upper output tube.

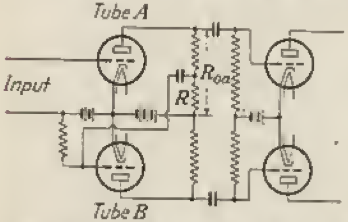


FIG. 24.—Phase-inverter type of push-pull amplifier.

It is preferable to derive the grid voltage for the phase-inverter tube from a point on the plate resistor, as shown, rather than from the grid resistor of the next tube because the magnitude and phase of this voltage will vary less with frequency. The proper grid voltage for the phase inverter tube is obtained by making $R = R_{oa}/A_b$, where R_{oa} is the plate resistor of the regular tube, A_b is the voltage amplification of the phase-inverter tube, and R is portion of R_{oa} between the point of pick-off for the phase-inverter tube and the d-c plate source. The grid voltages of the two output tubes will be unbalanced at the higher and lower frequencies because the voltage for the lower tube is influenced by two tube stages whereas the upper tube voltage is influenced by only one tube stage.

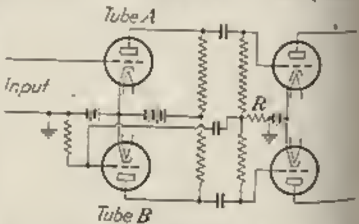


FIG. 25.—A self-balancing type of phase inverter.

Since R depends upon A_v, a correct balance will be had only when R is adjusted for a given tube. When the different tubes of the same type have large variations in constants, the self-balancing phase inverter of Fig. 25 is desirable. The value of R is not critical and may range from 0.1 to 0.5 of the grid resistors in the output stage.

23. Degenerative Feedback in Amplifiers. Controlled degenerative feedback is applied to a-f amplifiers for the purpose of improving their frequency characteristics, reducing wave-form distortion and phase shift, and increasing the stability. In the simplest case a voltage derived from the output of the amplifier is fed back so that it is effectively in series with the input or the grid circuit. Figure 26 illustrates degenerative feedback in its simplest form. For this simple circuit the general expression for the gain of the amplifier is

$$\text{Gain} = \frac{E_o}{E_i} = \frac{A_v}{1 - A_v\beta} \tag{28}$$

where A_v is the vector voltage amplification without feedback or is equal to E_o/E_g and has a negative real value, and $\beta = R/(R_f + R)$ when $1/\omega C << R_f$. $A_v\beta$ is called the *feedback factor*. The performance of the amplifier as to reduction of distortion, stability, etc., depends largely on the magnitude of $A_v\beta$.

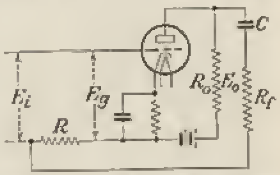


FIG. 26.—A simple amplifier with degenerative feedback.

That feedback improves stability is shown by the following example: In the amplifier circuit shown A_v has a negative numerical value. Hence the gain $= |A_v|/(1 + |A_v|\beta)$. Now assume $|A_v|\beta = 2$. The gain of the amplifier is equal to $|A_v|/3$. Suppose, owing to a change in d-c operating conditions or the substitution of another tube of the same type, $|A_v|$ is increased by 20 per cent. This will result in a 6.5 per cent increase in the gain of the amplifier. Greater values for $|A_v|\beta$ will produce less change in gain of the amplifier. When $A_v\beta$ becomes large compared to 1, the gain of the amplifier is equal to $-1/\beta$ and is entirely independent of the voltage gain of the tube.

Feedback also reduces wave-form distortion which is due to the nonlinear characteristics of the tube.

$$\text{Distortion output (with feedback)} = \frac{\text{distortion without feedback}}{1 - A_v\beta} \tag{29}$$

when the output voltage E_0 is kept the same with and without feedback.

Feedback is applied to different types of amplifiers and over one or more stages of an amplifier. In any case the equations are of the same forms as Eqs. (28) and (29), where A_v is the vector voltage amplification that the portion of the amplifier controlled by feedback would have without feedback and β is the vector ratio of the feedback voltage to the voltage which exists at the higher level point at which A_v is reckoned. Feedback must be so arranged that $A_v\beta$ has a negative real value over the useful frequency range of the amplifier and a value less than 1 for all other frequencies when it cannot be made negative. When the feedback becomes regenerative or the real value of $A_v\beta$ is positive although not sufficient to cause sustained oscillations, the gain *versus* frequency characteristics, the distortion output, and the stability of the amplifier will in general be worse than it is with no feedback.

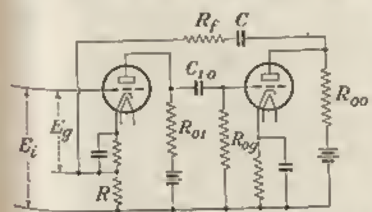


FIG. 27.—A two-stage amplifier with multiple degenerative feedback.

The many ways of applying simple and multiple feedback in amplifiers are too numerous to illustrate here. Figure 27 shows one of such circuits for a multistage amplifier. For more methods the reader is referred to the reference below.[1]

24. Power Supply to Tubes of an Amplifier. The design of the power supply is not included here. Only the things pertinent to the operation of the amplifier are given here.

Filament-power supply whether a.c. or d.c. should have good regulation. When using a.c., the leads should be low in resistance and twisted to avoid setting up disturbing magnetic fields.

For the B supply the importance of regulation depends upon the class of the amplifier, the class B type requiring the best regulation. It is important that the internal impedance of the supply, such as a rectifier, be small at the lowest a.f. as compared to the load impedance, particularly if the load impedance is somewhat inductive.

When using a common rectifier and also low capacity batteries for the B supply of a multistage amplifier, feedback will result unless means are taken to eliminate it. This feedback effect comes from a voltage set up largely by the plate current of the output tube flowing through the impedance of the B supply which is common to the plate circuits of the first stages of the amplifier. The feedback circuit is illustrated by Fig. 28. Z_c represents the impedance of the B supply which is common to all stages of the amplifier. A is the ratio of the plate voltage of the output tube to the plate voltage of the nth tube from the output end. Z_{on} is the impedance offered to the nth tube, and r_{pn} and μ_n are the constants of the nth tube. Then the over-all voltage gain of the amplifier is

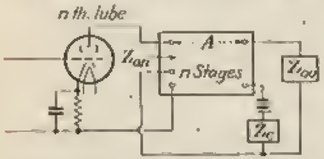


FIG. 28.—Diagram illustrating common impedance coupling between output stage and nth stage from output.

$$\text{Voltage gain of entire amplifier} = -\frac{\dfrac{\mu Z_{on}}{r_{pn} + Z_{on}} A}{1 - \dfrac{Z_c A r_{pn}}{Z_{oo}(r_{pn} + Z_{on})}} \tag{30}$$

The quantity $Z_{on}A/(r_{pn} + Z_{on})$ is the voltage gain when the common impedance coupling is zero. Hence feedback from common impedance coupling changes the gain and will cause sustained oscillations when the quantity

$$\frac{Z_c A r_{pn}}{Z_{oo}(r_{pn} + Z_{on})} = 1$$

The effect of common impedance coupling can be reduced and practically eliminated by the use of simple circuits of resistance, or inductive impedance, in series and capacitance in shunt with the plate supply to each tube, as shown in Fig. 29. These are called *decoupling circuits*, and the decoupling elements are C_{d1}, C_{d2}, R_{d1}, and R_{d2}. The reactances of the decoupling condensers should be small compared to the decoupling impedances at the lowest frequency for which the amplifier is designed.

[1] TERMAN, "Radio Engineering," 2d ed., Sec. 52, p. 248.

Then letting

$$D_1 = \frac{j\dfrac{1}{\omega C_{d1}}}{Z_{d1}} \quad \text{and} \quad D_2 = \frac{j\dfrac{1}{\omega C_{d2}}}{Z_{d2}}$$

the input tube of the amplifier of Fig. 29 will be decoupled by a factor $D = D_1D_2$, and the expression for the gain of the amplifier will be

$$\frac{E_{po}}{E_{g1}} = \frac{\text{gain with no common impedance coupling}}{1 - \dfrac{Z_c}{Z_{oo}} A \dfrac{Dr_{pn}}{(r_{pn} + Z_{on})}}$$

where A is the gain between the plate of input tube and plate of output tube. Hence, in order substantially to eliminate the trouble from common impedance coupling, it is necessary to make $\dfrac{Z_c}{Z_{oo}} AD \dfrac{r_{pn}}{(r_{pn} + Z_{on})}$ small compared to 1. This is usually accomplished quite well by making

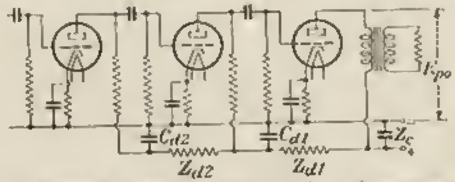


FIG. 29.—The use of decoupling circuits in a three-stage amplifier.

$D = 1/A$. In the circuit shown, the middle tube is decoupled from both the output and input tube. This may not be necessary, but a two-section decoupling circuit is much more effective than a single decoupling section having $C_d = C_{d2} + C_{d1}$ and $Z_d = Z_{d2} + Z_{d1}$. For a filtered rectifier plate supply the common impedance Z_c is the reactance of the output filter condenser. Some decoupling can be accomplished by connecting the individual stages of the amplifier across different points of the rectifier filter.

Self-bias resistors must be by-passed by condensers that have reactances (at the lowest frequency to be amplified) small compared to the resistors, or otherwise degeneration will result and the gain of the stage will be reduced at the lowest frequencies.

25. Direct-coupled Amplifiers. Under this classification are included all types of amplifiers in which the grid of one tube is connected to the plate of the preceding tube in such a manner that changes in d-c potential on the grid of the input tube will be amplified through the system. There are two important applications of such amplifiers. One application is an amplifying system for d-c purposes. The other application is an amplifier for a-c purposes when phase distortion at low frequencies is a consideration. It is difficult to obtain much amplification at low frequencies without phase distortion by the usual types of a-c amplifiers.

Direct-coupled amplifiers have high-frequency characteristics like a well-designed resistance-capacitance coupled amplifier. The tube capacitances shunt the coupling resistor and cause the amplification to decrease with increase in frequency above the frequency at which the effective shunt-capacitance susceptance is about three times the combined conductance of the coupling resistor and plate conductance.

The one common fault with many of the direct-coupled amplifiers when used for d-c work is instability. Small changes in the filament-, plate-, and grid-supply voltages cause false results in the output device. For amplifying low-frequency a.c. this particular characteristic is not so objectionable. Another common objection is the nature of the plate- and filament-supply voltages that are required.

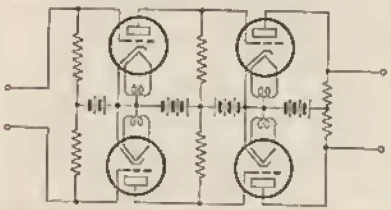


FIG. 30.—Direct-coupled push-pull amplifier.

The types of direct-coupled amplifiers that have been proposed are too numerous to discuss here. One type which seems to be free of some of the bad features enumerated above is a push-pull arrangement of tubes. This type possesses several advantages over ordinary single-tube-per-stage types. A two-stage push-pull type is shown in Fig. 30. For a balanced system, changes in plate current due to changes in the plate-supply voltage or to variation in cathode temperature are not amplified through the system. For balanced-output tubes there is no d-c component in the output device when no voltage is applied to the input. The output of the amplifier can be adapted to a high-impedance device such as the cathode-ray oscillograph or to a low-impedance device such as a milliammeter or the Duddell oscillograph. With the advent of twin tubes that have comparatively high transconductances, the push-pull arrangement becomes quite feasible. The main objection to push-pull input is that a device of high impedance must have balanced capacitances between its terminals and ground if the system is used at very high frequencies.

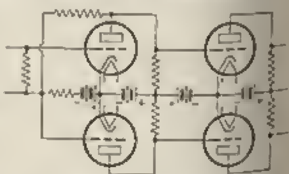


FIG. 31.—Direct-coupled amplifier with phase inverter for single-ended input.

The direct-coupled amplifier of Fig. 30 can be converted to a single-ended input by placing a resistor in series with the C battery of the input stage and connecting the plate of the upper tube through a resistor to the grid return end of the upper tube. This provides the voltage for the grid of the phase-inverter tube. The arrangement is shown in Fig. 31. This also introduces feedback into the upper tube and thereby increases its stability and reduces the distortion. Feedback can also be incorporated in the output tubes by placing resistors in the cathode circuits and adjusting the grid battery to give the proper operating grid potentials.

26. High-gain Amplifiers. Ingenious methods such as the one proposed by Schmitt[1] for obtaining practically the maximum possible voltage amplification from a high-mu pentode such as the 57 have

[1] SCHMITT, OTTO H. A., A Method of Realizing the Full Amplification Factor of High-mu Tubes, *Rev. Sci. Inst.*, December, 1933.

merit. The d-c plate potential is supplied to the pentode through a similar pentode which acts as a very high a-c impedance. The arrangement is shown in Fig. 32. The full gain of the tube is obtained only by a load of very high impedance.

27. Dynamic-coupled Amplifier.[1] A dynamic-coupled amplifier is one in which two tubes are coupled together as shown in Fig. 33. In this arrangement the input tube operates with negative grid bias, whereas the output tube operates with positive grid bias and therefore draws grid current. This grid current becomes the plate current of the input tube, or the input impedance of the output tube becomes the load impedance of the input tube. Moreover the load impedance of the input is all connected between

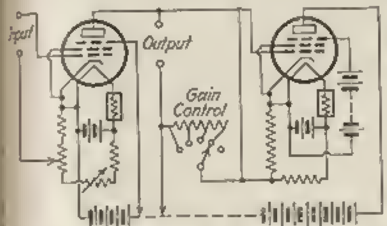


FIG. 32.—Direct-coupled high-gain amplifier of Schmitt.

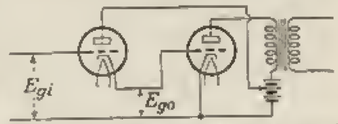


FIG. 33.—The dynamic-coupled amplifier.

the cathode and grid return. Hence the input tube has degenerative feedback, and its output voltage E_{go} (which is also the input voltage to the output tube) is

$$E_{go} = \frac{E_{gi}\mu_i \dfrac{R_{io}}{r_{pi}}}{1 + \dfrac{R_{io}}{r_{pi}}(1 + \mu_i)}$$

where μ_i and r_{pi} = amplification factor and plate resistance of the input tube

R_{io} = input resistance of the output tube.

This scheme of coupling gives satisfactory results only when r_{pi} and R_{io} depend upon E_{gi} and E_{go}, respectively, in such a way that R_{io}/r_{pi} is substantially constant over a complete cycle of E_{gi}. The output power of the system depends upon E_{go} and the output tube and load. Because the output tube is operating with the grid positive, the characteristics and load conditions for a triode are similar to those of a pentode instead of a triode with negative bias.

Usually the grid of the output should be connected through a resistor to the cathode in order to prevent high transient voltages during the warming-up period.

28. Frequency-response Control and Equalization in Amplifier Systems. By the use of certain expedients it is possible to design multistage amplifiers which will work with certain kinds of input and output devices and give over-all frequency-response characteristics of a desired type. Much can be done along this line when phase distortion is not a consideration. It may not always be desirable that the entire amplifier

[1] STROMEYER, C., General Theory and Application of Dynamic Coupling in Power Tube Design, *Proc. I.R.E.*, 1007, July, 1936.

or each stage thereof have a response which is constant over the entire frequency-band which is transmitted between the source and the load. One or more stages of transformer coupling of proper design can be used to accentuate the gain at the high frequencies. This can be done by making the Q_r factors of the transformers large and their resonant frequencies fall in the proper range. Other methods of accentuating the gain at the high frequencies can be accomplished by lowering the gain at the low frequencies. Shunting a portion of the plate resistor in a resistance-capacitance coupled amplifier with inductance will lower the gain at the low frequencies. The gain of an amplifier at the l-f end of the range can be accentuated by the use of one or more stages of the double-impedance coupling which is described in Art. 10. Condensers shunted across a portion the plate-coupling resistors in a resistance-capacitance coupled amplifier will result in higher gain at the high frequencies than at the low frequencies.

With most of the standard coupling methods, such as transformer, resistance-capacitance, and impedance-capacitance coupling, it is not so easy to control the gain at the medium frequencies without effect on the gain at the low or high frequencies. In other words the medium-frequency gain can be made greater or less than the gain at the low and high frequencies only by designing for lower or higher gain at low and high frequencies. A series circuit of resistance, inductance, and capacitance connected between the grid and cathode of one or more stages can be used to lower the gain over a small range of frequencies in the medium-frequency range. For such an arrangement the reduction in gain at the resonant frequency of the circuit depends upon the resistance of the circuit, and the band of frequencies over which the gain is reduced will depend largely upon the total effective resistance which includes the plate resistance of the tube immediately preceding the series circuit.

Variable gain control for the high frequencies, which is commonly known as *tone control*, is accomplished in its simplest manner by the use of a variable resistor and a fixed capacitance in series, both of which are placed in shunt with the coupling element of one stage of the amplifier. In a similar manner a variable resistor and a fixed inductance in series will serve as a gain control for the lower frequencies.

There are so many combinations of methods which may be employed to give frequency-response equalization in amplifier systems and to give any desired frequency response that it is impossible to cover all of them. Among these are the use of low-pass, high-pass, and band-pass filter circuits which are treated in another section of this handbook.

TESTING AND MEASUREMENTS

29. Frequency-response Measurements. A universal arrangement of equipment for making gain, or loss, measurements over a range of frequencies is shown in Fig. 34. The method is simply one of measuring the ratio of the output voltage to the input voltage. A calibrated potential divider or two calibrated resistors R_1 and R_2, so arranged that R_1 plus R_2 is constant, facilitates in making these measurements. For making gain measurements, S_1 is thrown in the position indicated by the full lines; for loss measurements in the dotted-line position. When the divider is so adjusted that the reading of the vacuum-tube voltmeter is the same for the two positions of S_2,

$$\text{Gain, or loss, in db} = 20 \log_{10} \frac{R_1 + R_2}{R_2}$$

For the full-line position of S_1 the resistance R_2 must always be small compared to the input impedance of the equipment under test plus R_s. To get the true gain, or loss, characteristic of a piece of equipment as it is actually used, it must be terminated as used and the termination included in the test. For example R_s and R_o represent the input and output resistance of the amplifier under test. These may also be any kind of impedances. When testing an input or interstage transformer it should be terminated in the tubes for which it is intended. Care must also be taken to limit the voltage applied to the equipment to the proper value.

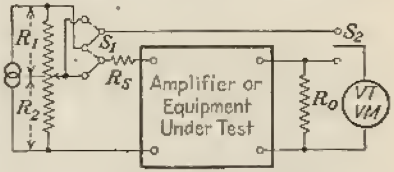


Fig. 34.—Method for making frequency-response measurements.

30. Measuring Distortion in Amplifiers. The simplest method for measuring the total harmonic distortion in the voltage across the output impedance in a power amplifier is shown in Fig. 35. For a given voltage impressed upon the grid of the amplifier, the vacuum-tube voltmeter is made to read a minimum by adjusting slide C of R_2 and the mutual inductance M. Then the reading of V.T.M. is a measure of the square root of sums of the squares of all the harmonic voltages across R_o. Mutual inductance M provides for a phase shift from 180 deg. through the amplifier tube. The vacuum-tube voltmeter must be as nearly an r-m-s meter as possible. The source should be reasonably free from harmonics. Switch S provides for measuring the total a-c voltage across R_o when V.T.M. has a multiplier to extend its range. A vacuum-tube voltmeter, using a type 56 or 76 tube and operated over a region in which the square root of the plate current plotted against grid voltage is nearly a straight line, makes an excellent meter for this purpose.

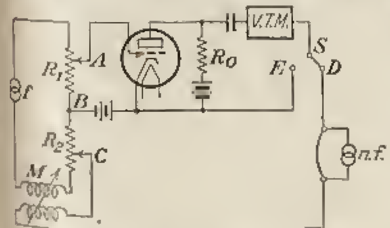


Fig. 35.—Circuit for measuring distortion.

When it is desired to know the separate harmonics in the output impedance, a voltage having a frequency nf almost equal to the harmonic sought may be introduced into the connection up to contact d, as illustrated. The voltage of nf will be equal to the particular harmonic voltage when the swing of the needle of V.T.M. is a maximum. The measurements may be carried out by means of a laboratory oscillator for nf and some filtering for the voltage obtained from the 60-cycle lighting circuit for f.

For the more refined measurements of distortion there are various types of wave analyzers on the market. These have a wider range of application than the simple method described above.

31. Measuring the Impedances of a Transformer and an Iron-core Reactance.[1] One of the simplest methods for measuring the impedance of an iron-core reactance at low frequencies and preferably the power frequency of 60 cycles is illustrated in Fig. 36. The circuit is arranged, when necessary, so that d.c. can be sent through the iron-core coil. When R_s is so adjusted that the reading of the vacuum-tube voltmeter is the same for both positions, a and b, of switch S, the absolute value of the impedance Z_x is equal to R_s, provided R_s is at least 20 times R. The error is less than 5 per cent. Oftentimes it is necessary to use an amplifier ahead of the vacuum-tube voltmeter. It is essential that the vacuum-tube voltmeter or amplifier be connected as shown, or false readings may result if the meter places too much stray shunt capacitance across Z_x.

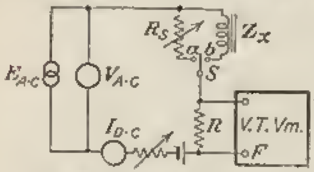


FIG. 36.—Circuit for measuring the impedance of iron-core coils.

The method of Fig. 36 may be used for measuring the impedances of the primary and secondary of a transformer. It is not possible, of course, to obtain the resistance and reactance separately by this method. Methods that place the standard resistance R_s in series with Z_x and require balancing the voltage drop across R_s against that across Z_x for the same current are objectionable except for quite low values of impedance. By such a method the d.c. through and a-c potential across Z_x are disturbed while adjusting R_s.

It is not generally safe to use the method described to measure the leakage inductance of a transformer. Leakage inductance is measured by shorting the secondary and measuring the impedance of the primary. Generally this measurement requires an inductance bridge because of the high value of R compared with X.

[1] See also Art. 29, Sec. 2.

SECTION 12

RADIO-FREQUENCY AMPLIFIERS

BY R. S. GLASCOW, M. S.[1]

1. Class A Amplifier. Amplifiers are divided into three general classes, A, B, and C, depending on the type of service in which they are to be used.

A class A amplifier is one which operates so that the plate output wave shapes of current are practically the same as those of the exciting grid voltage.

This is accomplished by operating the tube with sufficient negative grid bias so that some plate current flows at all times and by applying an alternating excitation voltage to the grid of such value that the dynamic operating characteristic is essentially linear. The grid must not go positive on excitation peaks, and the plate current must not fall low enough at its minimum to cause distortion due to curvature of the characteristic.

The characteristics of class A operation are freedom from distortion and relatively low power output. Practically all a-f amplifiers are operated in this manner. Radio-frequency amplifiers of the type used in receiving sets to amplify the signal voltage prior to detection are also of this class.

Class B and C amplifiers will be discussed under Power Amplifiers.

2. Radio-frequency amplifiers for receiving sets are usually classified as to the type of coupling employed between stages. This coupling means can be a resistance, an impedance, a transformer, or any combination of these elements. The circuit constants of the coupling means may be adjustable or fixed, giving rise to a further classification of a tuned or an untuned amplifier. In the latter the circuits are similar to those employed for a-f amplifiers. Special precautions must be taken in the circuit design if uniform amplification is to be obtained over an extended range of frequencies.

3. Resistance-coupled Amplifier. This type of amplifier is occasionally used where uniform amplification is desired over a moderate band in the lowest range of radio frequencies. In Fig. 1 the output voltage E_2 is given by

$$E_2 = \frac{\mu R_b}{r_p + R_b} E_1 \tag{1}$$

where μ and r_p are, respectively, the amplification factor and plate resistance of the tube used. Defining the voltage amplification per stage A as the ratio of the output voltage to the input voltage, we have

$$A = \frac{E_2}{E_1} = \frac{\mu R_b}{r_p + R_b} \tag{2}$$

[1] Professor of Electrical Engineering, Washington University, St. Louis.

As R_b is made very large compared to r_p the value A approaches μ as a limit, so that tubes having a large value of μ are necessary if reasonably high gain per stage is desired. Equation (2) presumes that the input impedance of the next stage which is shunted across R_b is enormously large, so that R_b is not appreciably reduced as a result of being shunted by this input impedance.

In a typical cascade amplifier as shown in Fig. 2, R_b is in effect shunted by the grid leak R_g in parallel with C_g, the input capacity of the tube. The reactance of the blocking condenser C_1 in series with them is negligibly small in comparison. For frequencies lower than 500 kc, with a

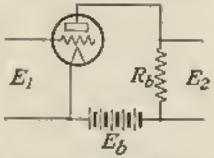


FIG. 1.—Resistance-coupled amplifier.

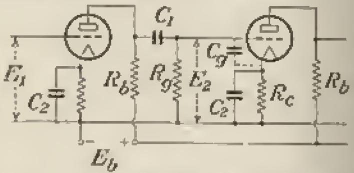


FIG. 2.—Resistance-capacity coupled cascade amplifier.

pure resistance in its plate circuit, C_g may be regarded as constant and independent of the frequency, and is given by

$$C_g = C_{gf} + C_{gp}\left(1 + \frac{\mu R_b}{r_p + R_b}\right) \tag{3}$$

where C_{gf} = capacity between grid and filament

C_{gp} = capacity between grid and plate.

These interelectrode capacities will be from 4 to 10 $\mu\mu$f depending on the type of tube and socket used; hence C_g may lie anywhere from 40 to 80 $\mu\mu$f. Thus at 1,000 cycles the input impedance of the tube alone will be about 3 megohms, while at 100 kc it has dropped to about 30,000 ohms. As a result the gain per stage diminishes as the frequency increases due to the reduction of the effective value of R_b by the short-circuiting effect of C_g.

The voltage amplification in Fig. 2 will be E_2/E_1, where

$$E_2 = \frac{E_1}{1 + r_p\frac{R_b + R_g}{R_bR_g} + j\omega C_g r_p} \tag{4}$$

This expression assumes that the reactance of the blocking condenser C_1 is negligible at these frequencies.

4. Resistance-coupled Amplifier Using Pentodes. The constants of available triodes render them very unsatisfactory in resistance-coupled r-f amplifiers, and suitable pentodes are accordingly used. These tubes have an input capacity which is substantially independent of the load impedance in the plate circuit and is composed of C_{gf} plus the capacity between the control grid and the screen grid. Accordingly a much smaller value of C_g can be obtained than with a triode. The value of r_p in r-f pentodes is usually in the vicinity of a megohm, which is far greater than the plate-load impedance which can be successfully used.

Consequently the alternating component of the plate current i_p is practically independent of the load in the plate circuit and is given by

$$i_p = g_m E_1 \tag{5}$$

where g_m = transconductance of the tube.

In Fig. 3, R is the equivalent resistance of the load in the plate circuit[1] and C is the total capacity shunted across the load, consisting of the input capacity C_g of the subsequent stage, plus the output capacity of the first tube, plus the stray capacity of leads, etc., to ground. The expression for the voltage amplification will be

$$A = \frac{g_m R}{\sqrt{\omega^2 C^2 R^2 + 1}} \tag{6}$$

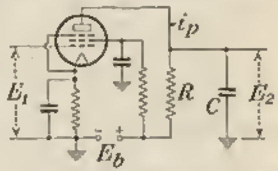


FIG. 3.—Resistance-coupled pentode amplifier.

When the frequency is such that

$$R = 1/\omega C$$

the denominator of Eq. (6) will be numerically equal to $\sqrt{2}$, and the amplification will have fallen to 70.7 per cent of its l-f value, or a reduction of 3 db. The frequency at this point is

$$f_0 = \frac{1}{2\pi RC} \tag{7}$$

and is a convenient relationship for design purposes, as it establishes the maximum frequency the amplifier will transmit without serious loss of gain. It corresponds to the cutoff frequency of a low-pass filter.

This upper limiting frequency may be extended by using a low value of resistance in the plate circuit so that the reactance of C is large in comparison to R. However, this will reduce the amplification over the entire frequency range. In practice, if f_0 is to be 1 or 2 Mc, R will perhaps be 2,000 ohms or less, as it is difficult to secure a value of C much less than about 25 $\mu\mu$f. Consequently a gain per stage of from 5 to 20 is about all that can be secured, using pentodes having a transconductance of 9,000 micromhos, such as types 1851 and 1852.

Combining Eqs. (6) and (7) the voltage amplification can also be expressed as

$$A = \frac{g_m R}{\sqrt{1 + \frac{f^2}{f_0^2}}} \tag{8}$$

5. Compensated Resistance-coupled Amplifiers. From Eq. (8) it is seen that the gain gradually falls off as f_0 is approached, which may be undesirable, particularly in video-frequency amplifiers for television. One method of diminishing the shunting effect of the load-circuit capacity C is to place an inductance L of proper size in series with the plate-circuit

[1] The value of this resistance is $R = \frac{R_bR_g}{R_b + R_g}$, but in practice R_g is usually very much larger than R_b, so that the value of plate-load resistance R_b may be substituted for R in most cases.

load resistance R, as shown in Fig. 4. At low frequencies the reactance of L is small, and the load impedance is essentially equal to R. But as the frequency increases, the impedance of the branch $R + j\omega L$ becomes progressively greater, thus tending to offset the shunting effect of C.

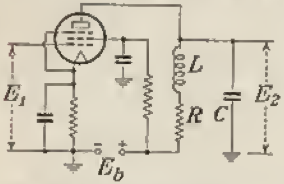


FIG. 4.—Compensated resistance-coupled amplifier.

The plate-load impedance is given by

$$Z_L = \frac{R + j\omega[L - C(R^2 + \omega^2 L^2)]}{\omega^2 C^2 R^2 + (\omega^2 LC - 1)^2} \tag{9}$$

and it may be shown that, by making the reactance of the coil L equal to half the load resistance R at the limiting frequency f_0, Z_L will substantially be constant in magnitude and equal to R. The gain will be approximately $g_m R$ up to a frequency of f_0. The required relations are

$$R = \frac{1}{2\pi f_0 C} = 4\pi f_0 L \tag{10}$$

Substituting Eq. (10) in Eq. (9), the load impedance becomes

$$Z_L = \frac{R\left[1 - j\left(\frac{f^3}{4f_0^3} + \frac{f}{2f_0}\right)\right]}{\left(\frac{f}{f_0}\right)^2 + \left(\frac{f^2}{2f_0^2} - 1\right)^2} \tag{11}$$

The accurate expression for the voltage amplification per stage is then

$$A = \frac{g_m R\sqrt{1 + \left(\frac{f^3}{4f_0^3} + \frac{f}{2f_0}\right)^2}}{\left(\frac{f}{f_0}\right)^2 + \left(\frac{f^2}{2f_0^2} - 1\right)^2} \tag{12}$$

Using the circuit constants mentioned at the end of Art. 4, L will be in the vicinity of 25 μh. Additional compensation methods are described in the reference below.[1]

6. Impedance-coupled Amplifier. The simplest amplifier of this type merely employs a choke coil in the plate circuit as shown in Fig. 5. The voltage amplification per stage in the case of a triode is given by

$$A = \frac{E_2}{E_1} = \frac{\mu\sqrt{R_b^2 + \omega^2 L_b^2}}{\sqrt{(r_p + R_b)^2 + \omega^2 L_b^2}} \tag{13}$$

where R_b and L_b are, respectively, the resistance and inductance (in henrys) of the choke coil. If the resistance of the coil is small compared to its reactance ωL_b and to the plate resistance r_p, Eq. (13) becomes

$$A = \frac{\mu\omega L_b}{\sqrt{r_p^2 + \omega^2 L_b^2}} \tag{14}$$

[1] SEELEY and KIMBALL, Analysis and Design of Video Amplifiers, *RCA Rev.*, pp. 290–308, January, 1939.

If ωL_b is very large compared to r_p, A approaches μ of the tube as a limiting value, as was the case with the resistance-coupled amplifier. By choosing L_b large enough so that the reactance of the coil is large compared to the plate resistance of the tube at the lowest frequency we are interested in, the gain will be constant for all higher values of frequency. Owing to distributed capacity effects and the shunting of the coil by the input capacity of the next tube, it is not possible to obtain uniform amplification as predicted above except at low frequencies. For high frequencies such as the present broadcast band the effect of this capacity is to produce a parallel resonant circuit whose impedance is high at the resonant frequency but which drops off rapidly for frequencies higher than resonance. This results in a reduction of the gain for frequencies above resonance. To avoid this, it becomes necessary to use a value of choke-coil inductance such that resonance occurs somewhat below the highest frequency to be amplified. This value of inductance is governed chiefly by the input capacity of the next tube which may be of the order of 10 to 20 $\mu\mu f$, depending on the type of tube used and the nature of the load in its plate circuit. For this reason there is little to be gained by reducing the distributed capacity of the coil if it is already small compared to the tube input capacity. At broadcast frequencies the value of inductance thus obtained results in too low a reactance to give good amplification for frequencies much below resonance.

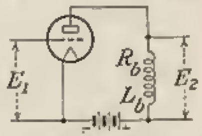


FIG. 5.—Impedance-coupled amplifier.

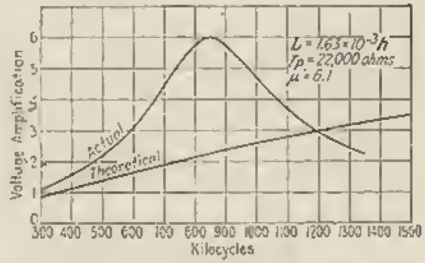


FIG. 6.—Amplification of a choke-coupled amplifier tube.

This is illustrated in Fig. 6. The coil used was a single layer solenoid closely wound with 173 turns of No. 28 wire having an inductance of 1.63×10^{-3} henry and about 10 ohms d-c resistance. The distributed capacity of the coil was 3.5 $\mu\mu f$. The curve shows the measured amplification[1] using a Western Electric 215-A "peanut" tube which had an amplification factor of 6.1 and a plate resistance of 22,000 ohms. The input capacity of the vacuum-tube voltmeter which used a tube of the same type was 18 $\mu\mu f$, including leads, which lowered the natural period of the choke coil to 850 kc. The lower curve shows the theoretical amplification that would be obtained if these shunting capacities were absent.

[1] FRIIS and JENSEN, *Bell System Tech. Jour.*, **3**, 187, April, 1924.

7. Tuned Impedance-coupled Amplifier. Since the coil in the plate circuit will be shunted by some capacity which will cause the combination to have a resonant frequency at which the amplification will be a maximum, the circuit of Fig. 7 is sometimes used in receiving circuits. This form of circuit, with various modifications in the details, is commonly used as the coupling means between amplifier stages in radio transmitting circuits. The condenser C is not usually variable, except in low-power transmitters, and tuning is accomplished by varying the position of taps on the coil L. A vernier adjustment of inductance is provided in the form of a heavy copper or aluminum disk which acts as a short-circuited secondary of a single turn and which can be rotated within the coil. By rotating the plane of this disk, a fine adjustment of inductance is obtained for tuning purposes. The position of the taps should be chosen so that the plane of the disk is displaced nearly 90 deg., at resonance, from that of the coil, so as to minimize the I^2R loss in the disk.

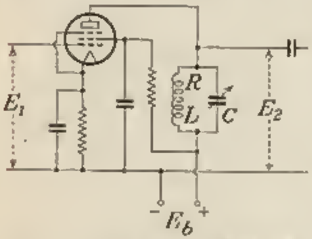


FIG. 7.—Tuned impedance-coupled amplifier.

In receiving circuits, particularly where space may be at a premium, a fixed condenser may be used for C, tuning being accomplished by adjusting the position of a suitable cylindrical core of molded iron dust which is arranged so that it can be moved in and out of the coil.[1] Improvements in the quality of these iron-dust cores in recent years enables them to be used to advantage at much higher frequencies. Values of Q $(= \omega L/R)$ in the vicinity of 100 can be obtained in the broadcast range of frequencies, using a coil diameter of about ⅜ in. and a length of about 1 in. This method, known as *permeability tuning*, has been used in connection with push-button tuning, particularly in automotive receiving sets.

The voltage amplification of the circuit in Fig. 7 at any frequency is given by

$$A = \frac{E_2}{E_1} = g_m Z_L \tag{15}$$

where Z_L is given in vector form by Eq. (9). Transforming Eq. (9) into its scalar magnitude, Eq. (15) becomes

$$A = g_m \sqrt{\frac{R^2 + \omega^2 L^2}{\omega^2 C^2 R^2 + (\omega^2 LC - 1)^2}} = g_m \sqrt{\frac{\omega^2 L^2 (1 + Q^2)}{\omega^4 L^2 C^2 + Q^2(\omega^2 LC - 1)^2}} \tag{16}$$

At resonance the voltage amplification will be

$$A = g_m \frac{R^2 + \omega^2 L^2}{R} = g_m \omega L \frac{1 + Q^2}{Q} \tag{17}$$

8. Tuned Transformer-coupled Amplifiers. A typical tuned transformer-coupled amplifier is shown in Fig. 8. Receiving sets of the

[1] POLYDOROFF, W. J., Ferro-inductors and Permeability Tuning, *Proc. I.R.E.*, p. 690, May, 1933.

t-r-f type are seldom used today as the great majority now use the superheterodyne circuit. However, most of the better receivers of the latter type employ a stage of t.r.f. ahead of the first detector. In its original form, which is still used to some extent in some of the short-wave bands of an all-wave receiver, the primary coil L_p in Fig. 8 was of lower inductance than the secondary. The secondary inductance for a frequency range of 1,600 to 550 kc. is ordinarily from 200 to 250 μh.

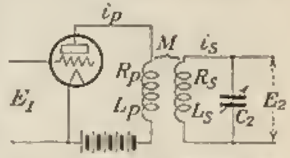


FIG. 8.—Tuned transformer-coupled amplifier.

Since the resistance R_p and the reactance ωL_p of the primary can be neglected in comparison to the plate resistance r_p of the tube, the voltage E_2 across the secondary at any frequency is

$$E_2 = \frac{-jE_1 \mu M}{C_2\left[r_p R_s + \omega^2 M^2 + jr_p\left(\omega L_s - \frac{1}{\omega C_2}\right)\right]} \tag{18}$$

At resonance, where $\omega L_s = 1/\omega C_2$, the voltage amplification becomes

$$A = \frac{E_2}{E_1} = \frac{\mu \omega^2 M L_s}{r_p R_s + \omega^2 M^2} \tag{19}$$

If the mutual inductance M in Eq. (19) is adjusted to satisfy the condition

$$\omega M = \sqrt{r_p R_s} \tag{20}$$

the optimum value of voltage amplification will be obtained and Eq. (9) reduces to

$$A_{\text{opt.}} = \frac{\mu \omega L_s}{2\sqrt{r_p R_s}} \tag{21}$$

which is the maximum amplification it is possible to obtain with a given tube and coil.

When M is adjusted to its optimum value, it will be noted that the figure of merit of the tube is $\mu/\sqrt{r_p}$. Therefore if two tubes have equal values of transconductance, the one having the higher amplification factor will give the greater gain. Tetrodes and pentodes will accordingly produce a greater gain than a triode. With M less than optimum the gain becomes more nearly proportional to the transconductance of the tube. When optimum coupling is employed, the amplification is directly proportional to the ratio of the coil reactance to the square root of its resistance, instead of Q, of the coil. With values of M considerably less than optimum, as when pentodes are used, the gain becomes more nearly proportional to the figure of merit Q, of the coil.

The impedance looking into the primary coil in Fig. 7 is

$$Z_p' = R_p + j\omega L_p + \frac{\omega^2 M^2}{R_s + j\left(\omega L_s - \frac{1}{\omega C_2}\right)} \tag{22}$$

At resonance, with optimum coupling, $Z_p' = r_p$ of the tube. This condition differs from the resistance- and impedance-coupled amplifiers in that, in the latter two, optimum amplification is approached by making the impedance of the load very large compared to r_p.

If a pentode is used in the circuit of Fig. 8, the above equations are still applicable. Since these tubes have plate resistances r_p approaching a megohm in value and amplification factors varying from several hundred to several thousand, the coupling that can be used between primary and secondary without causing instability in the form of oscillations is far below the optimum value. The preceding equations can therefore be simplified. Since $r_p >> \omega M$ in the case of a pentode, the expression for the secondary voltage in Eq. (18) becomes

$$E_2 = \frac{-jE_1 g_m M}{C_2\left[R_s + j\left(\omega L_s - \frac{1}{\omega C_2}\right)\right]} \tag{23}$$

and the voltage amplification at resonance is

$$A = g_m Q_s \omega M \tag{24}$$

These tubes enable values of amplification per stage to be obtained which are much larger than can be obtained with triodes. With a given

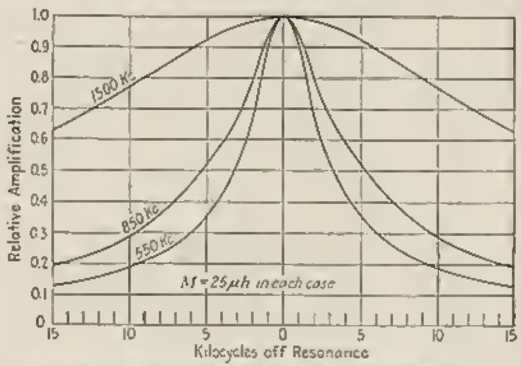


FIG. 9.—Selectivity as a function of frequency.

secondary coil the selectivity in the case of a pentode is better than with a triode; or for equal selectivities the pentode circuit can use a smaller and less expensive coil. These advantages, together with their freedom from oscillation without the use of neutralizing circuits, have caused triodes to be virtually abandoned in the field of r-f amplifiers for receiving circuits. Triodes are still used in the higher power amplifier stages of radio transmitters.

9. Variations in Selectivity. The type of t-r-f transformer considered in Art. 8, which had a primary inductance smaller than that of the secondary, causes an appreciable variation in selectivity over the tuning range, as shown in Fig. 9. The amount of gain also tends to fall at the l-f end of the tuning range. Both of these difficulties can be reduced by the use of a primary coil of large inductance—about 3 or 4 mh. This value of L_{p}, in conjunction with its distributed capacity, combined with the output capacity of the tube, resonates the primary circuit to a frequency somewhat below the l-f tuning limit of the secondary. The electrical circuit is the same as Fig. 13 with the primary permanently

tuned to a fixed low frequency. The variation in selectivity of a tuned transformer having a large primary is shown in Fig. 10. The variation in gain throughout the tuning is also much less than with the transformer of Fig. 9.

The expression for the voltage amplification is

$$A = \frac{\mu M}{C_2\sqrt{a^2 + b^2}} \tag{25}$$

$$\text{where } a = R_2[R_1 + r_p(1 - \omega^2 L_1 C_1)] - \omega(L_1 + r_p R_1 C_1)\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \omega^2 M^2$$

$$b = \omega R_2(L_1 + r_p R_1 C_1) + [R_1 + r_p(1 - \omega^2 L_1 C_1)]\left(\omega L_2 - \frac{1}{\omega C_2}\right) + \omega^3 M^2 C_1 r_p$$

From a practical point of view it is evident from the complexities of this expression that it would be much easier to determine the character-

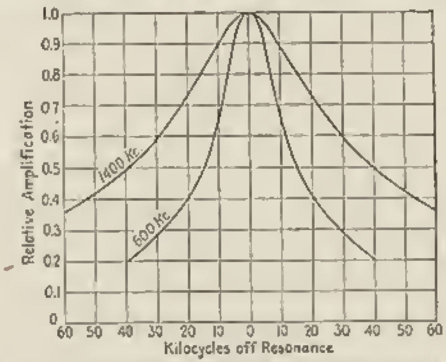


FIG. 10.—Variation in selectivity in t-r-f amplifier having high-inductance primary.

istics of these transformers experimentally by laboratory measurements. The variation of the resistances with frequency will have to be measured in any event, so one might just as well determine the over-all performance by measurement. In this way the effects of regeneration, stray couplings, the proximity of shielding, etc., may be included. The presence of small amounts of feedback, if not sufficient to produce oscillations, will often cause the actual gain of an amplifier to depart from its computed value by a considerable amount.

The use of a primary operated above its resonant frequency results in a plate-load impedance which has capacitive reactance. A load of this nature results in negative feedback in the case of triodes, so that neutralizing circuits have to be employed to prevent the gain from being reduced to a fraction of its theoretical value. Ordinarily, these circuits are used to balance out the effects of positive feedback and prevent oscillation. Pentodes are free from these troubles.

10. Combinations of Inductive and Capacitive Coupling. To secure better performance in tuned amplifiers without resorting to moving parts other than the tuning condensers, combinations of inductive and capacitive coupling between stages have been used.[1] By a proper choice of circuit elements it is possible to make the effective coupling vary with the frequency in a predetermined manner. In this way the variation of gain with frequency can be given almost any desired characteristic.

Two examples of such circuits are shown in Fig. 11. In Fig. 11*a* the coil L_b has a large value of inductance; hence its distributed capacitance C_1, augmented by C_{pf} of the tube, resonates it to a frequency somewhat below the tuning range of the set. The output current of the tube divides between L_b and the path through the coupling condenser C_m. At low frequencies a larger portion of the output current flows through this second path because of the high impedance offered by L_b as parallel

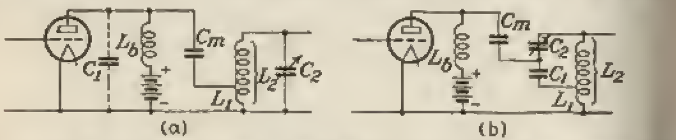


FIG. 11.—Tuned amplifiers using combinations of inductive and capacitive coupling.

resonance is approached in the latter. This causes the voltage induced in L_2 to remain more nearly constant over the tuning range.

The circuit of Fig. 11*b* accomplishes the same results in a somewhat different manner. The coil L_b and condenser C_m merely serve as choke coil and blocking condenser of an amplifier using parallel feed. The amplified output current divides between C_1 and C_2 and then recombines to flow through the primary L_1 of the autotransformer. The capacity of the tuning condenser C_2 is increased as the signal frequency is lowered, which causes a progressive increase in the effective coupling. C_1 is about twenty times larger than the maximum value of C_2, while L_1 includes about a turn or two of the coil L_2.

11. Cascade Amplifiers. If two or more identical stages of amplification are connected in cascade, the over-all voltage amplification is given by

$$A = A^n \tag{26}$$

where n = number of stages
A = amplification per stage.

This expression presumes that the various stages do not react on each other, which is not always the case in practice owing to small unavoidable couplings between input and output circuits. If the various stages are not all identical, the over-all amplification will be the product of the individual values of A per stage. The response curve of a multistage amplifier composed of identical stages is readily obtained from the curve of an individual stage by raising its ordinates to the nth power, where n is the number of stages.

[1] WHEELER, H. A., and W. A. MCDONALD, Theory and Operation of Tuned Radio-frequency Coupling Systems, *Proc. I.R.E.*, **19**, 738, May, 1931.

The use of several stages of cascade tuned r-f amplification enables both the selectivity and fidelity of the amplifier to be increased, provided the tuning of each stage is made broader as the number of stages is increased. This is illustrated in Fig. 12, both amplifier circuits being alike except for the values of mutual inductance between the primary and secondary of the t-r-f transformers. The necessity for broader tuning per stage in multistage amplifiers in order to avoid too great a sacrifice in fidelity permits the use of coils of rather compact dimensions wound with relatively small wire. The increased coil resistance thus produced will reduce the gain per stage, but this can be offset if necessary by increasing the mutual inductance to more nearly the optimum value. At frequencies sufficiently remote from resonance so that the gain per stage becomes less than unity, a cascade amplifier acts as an *attenuator* of the signal. An increase in the number of stages will therefore actually

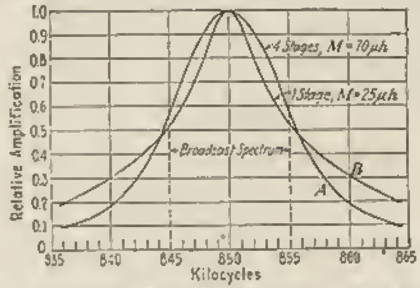


FIG. 12.—Increase in selectivity with cascading.

decrease the strength of interfering signals whose frequencies are above or below the band where the gain per stage is equal to or greater than one. All signals whose frequencies lie within this band will be strengthened by an increase in the number of stages. For this reason two types of selectivity may be recognized: the *adjacent-channel selectivity*, and the *distant-channel selectivity*. It is therefore possible in a comparative test of two amplifiers of equal sensitivity to find that the first will produce less interference from interfering signal of, say, 30 kc away from resonance than the second; while for a signal of, say, 60 kc away there may be more interference present than in the second amplifier.

The attenuation of signals remote from the resonant frequency requires that the amplifier be well shielded in order to prevent short portions of the lead wires and circuits of the output stage from acting as antennas and picking up energy. Thus a few inches of exposed wire running to the grid of the detector tube might have a voltage induced in it from an interfering powerful local station which is much greater in magnitude than these same signals after passing through the amplifier.

12. Band-pass Filters. A rectangular response curve would be ideal for the r-f amplifier of a receiving set designed for entertainment purposes. The use of a pair of tuned circuits as a coupling means between stages results in a flatter response curve with steeper sides than can be obtained with a single tuned circuit. Such an arrangement is shown in Fig. 13, and the general appearance of the resultant response curves is given in Fig. 14. Owing to the more uniform amplification obtained over a wider band of frequencies, these circuits are often referred to as band-pass filters. This form of circuit is commonly used in the i-f amplifier of superheterodynes.

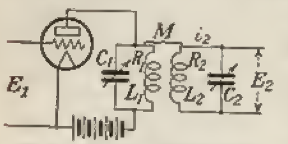


FIG. 13.—Transformer-coupled amplifier with primary and secondary tuned.

When the primary and secondary are both tuned to the same frequency, the width of the transmitted band depends upon the magnitude of the coupling between them. A double-humped response curve results if M is greater than the critical value, and, as M is increased, the two peaks move farther apart and the hollow between them becomes deeper, particularly if the resistance of the two coils is low.

In practice, both primary and secondary are tuned to the same frequency; consequently $\omega L_1 = 1/\omega C_1$ and $\omega L_2 = 1/\omega C_2$. If the resultant common resonant frequency is called f_0, the selectivity characteristic can be determined, assuming a pentode to be used, from

$$E_2 = \frac{g_m E_1 M}{\omega C_1 C_2} \times \frac{1}{R_1R_2\left[1 - 4Q_1Q_2\left(\frac{f-f_0}{f_0}\right)^2 + j(Q_1+Q_2)\frac{f-f_0}{f_0}\right] + \omega^2M^2} \tag{27}$$

where f = frequency in question
Q_1 and Q_2 = values of $\omega L/R$ of the primary and secondary circuits.

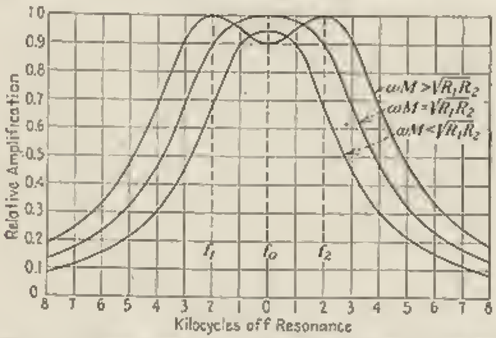


FIG. 14.—Response curves of doubly-tuned r-f stage.

At resonance when $f = f_0$, the voltage amplification will be

$$A = \frac{g_m M}{\omega C_1 C_2 (R_1R_2 + \omega^2 M^2)} \tag{28}$$

A single-humped curve results when $\omega M = \sqrt{R_1R_2}$, as shown in Fig. 14. In the case of this figure the value of Q for the two circuits is somewhat higher than would be employed in the i-f amplifier of a receiving set designed for entertainment purposes. While the selectivity would be excellent, the resulting attenuation of side bands in a two-stage amplifier using three such transformers would greatly impair the fidelity of reception. It is sometimes the practice to use transformers of slightly different characteristics in the several stages. Thus one or two of the transformers may have a more or less pronounced hollow at f_0, while the other may have a single hump. In this way the amplification may be made fairly uniform throughout the band between f_1 and f_2 and then fall off sharply on either side. In television receivers where the i-f amplifier is called upon to transmit uniformly a band of frequencies about 4.5 Mc in width, a transformer consisting of three tuned coupled circuits may sometimes be required to secure the desired uniformity over the transmitted band. The design problems involved in these applications are discussed in the references below.[1]

With the trend toward higher fidelity in the better grade of broadcast receivers, it is highly desirable to have some adjustable control over the shape of the i-f response curve so as to be able to increase the fidelity on local reception when high selectivity, to prevent interference and noise, is not required. But the broad response curve required would be unsatisfactory in many cases of distant reception where high selectivity might be needed to avoid adjacent channel interference. One scheme of securing adjustable selectivity is to vary the coupling between the primary and secondary coils by mechanical means. Another method, illustrated in Fig. 15, is to have a small coil L_3 tightly coupled to L_1. By rotating the switch arm to points 2 and 3, a progressive increase in M between L_1 and L_2 is secured. This will result in a slight detuning of the secondary circuit as the response curve is widened, but this is of no serious consequence. Individual iron-dust cores in L_1 and L_2 enable close coupling to be had between L_1 and L_3 with a comparatively small number of turns in the latter.

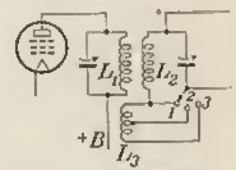


FIG. 15.—Method for varying band width in i-f transformer.

13. Regeneration in Amplifiers. The three-electrode vacuum tube is not a perfect unilateral device but permits the amplified output energy to react upon the input circuit. The grid-to-plate capacity of the tube serves to couple electrostatically the input and output circuits as shown in Fig. 16. If some of the output voltage is fed back into the input circuit so as to be in phase with e_g, the total, or regenerative amplification, may be expressed by

$$A_r = A\frac{S}{1 - AS} \tag{29}$$

where S is the fraction of the output which is fed back into the input circuit and A is the gain of the amplifier if feedback were absent. If the quantity AS is unity, the total amplification becomes infinite, and a continuous oscillation will result. In addition to feedback due to C_{gp} which almost always has to be balanced out to secure stability, feedback due to coupling resulting from the use of a common B or C battery may

[1] MOUNTJOY, G., Television Signal-frequency Circuit Considerations, *RCA Rev.*, p. 204, October, 1939; and Simplified Television I-f Systems, *ibid.*, p. 299, January, 1940.

be sufficient to cause instability. Small electrostatic or electromagnetic couplings between the input and output circuits of the amplifier can also give rise to oscillation even if each stage has been perfectly neutralized. For example, a four-stage amplifier having a gain of 10 per stage will oscillate if as much as 0.01 per cent of the output voltage succeeds in getting into the input circuit in the proper phase. Consequently multistage amplifiers of high over-all gain must be carefully shielded to avoid instability, particularly at the higher frequencies.

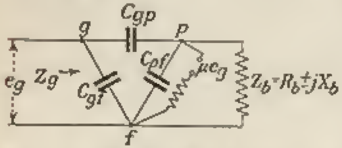


FIG. 16.—Equivalent circuit of a triode.

The oscillation of a single-stage amplifier can occur only if the plate circuit is sufficiently inductive. If the impedance in the plate circuit is pure resistance or a condensive reactance, no oscillations can take place, although in the latter case antiregenerative feedback may occur of sufficient magnitude greatly to reduce the resultant gain. The effect of feedback may be looked upon as being due to the input impedance Z_g of the grid-filament terminals of the tube. This impedance is of the form

$$Z_g = \pm r_g - j\frac{1}{\omega C_g} \tag{30}$$

When the plate circuit is inductive, the sign of r_g is negative, so that the tube is then capable of annulling part or all of the positive resistance of the associated input circuit. In the latter event, oscillations occur. The effect of the various circuit elements of Fig. 16 on Z_g is given by

$$Z_g = \frac{C_{gp} + C_{pf} - j\frac{1}{\omega}\left(\frac{1}{R_b \pm jX_b} + \frac{1}{r_p}\right)}{\frac{\mu C_{gp}}{r_p} + (C_{gf} + C_{gp})\left(\frac{1}{R_b \pm jX_b} + \frac{1}{r_p}\right) + j\omega(C_{gf}C_{gp} + C_{gp}C_{pf} + C_{pf}C_{gf})} \tag{31}$$

When Z_b is capacitive and has sufficient resistance associated with it, r_g is positive, and the tube may introduce rather large losses into the input circuit, even though the grid is biased sufficiently negative so that no conductive grid current flows.

14. Methods of Avoiding Oscillation. Circuits designed to combat the effects of regeneration are of two general types. Either sufficient resistance is introduced into the input circuit to offset the negative resistance introduced by the tube or else a suitable network of circuit elements is employed so as to isolate electrically the input and output circuits by making them two pairs of opposite points of an a-c bridge. The most common method of the first-mentioned group is to insert a resistance of several hundred ohms in series with the grid of the tube. In a tuned amplifier designed to cover a range of frequencies, this resistance must be sufficiently large to secure stability at the highest frequency, which means that it is much larger than necessary at the lower frequencies. This results in loss of amplification at these frequencies.

15. Neutralizing Circuits. One form of bridge circuit due to C. W. Rice is shown in Fig. 17 where are given the actual circuit and the electrical equivalent with the tube electrodes omitted. The filament terminal of the tube, instead of being connected to the lower end of the input circuit, is connected to an intermediate point which divides the inductance into two parts, L_a and L_b. The lower terminal n of the input circuit is connected to the plate through a small balancing condenser C_n. The terminals g and n of the input circuit and f and p of the output circuit constitute two pairs of opposite points of a bridge. An inspection of the latter figure indicates that no voltage can exist across the input terminals gn due to a voltage between fp if the arms are balanced.

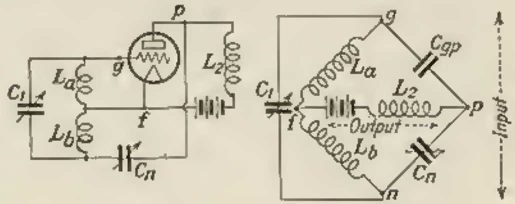


FIG. 17.—Rice neutralized amplifier.

Hence the energy which is fed back through C_{gp} is opposed in phase by that which flows through C_n. The conditions for a balance are

$$\frac{L_a}{L_b} = \frac{C_n}{C_{gp}} \tag{32}$$

This balance is not entirely independent of frequency as Eq. (32) would indicate unless the coupling between L_a and L_b is substantially unity. This is because L_a is shunted by the input capacity of the tube. With certain arrangements a h-f parasitic oscillation may take place which will impair the performance of the amplifier at the frequencies for which it was designed. A small capacity of about the size of C_n shunted across L_2 will often prevent such parasites in receiving circuits. The Rice circuit is commonly used in neutralizing r-f power amplifier circuits in transmitter sets.

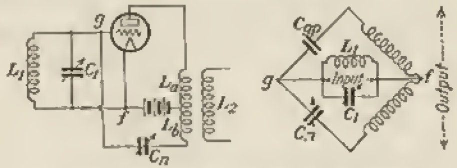


FIG. 18.—Hazeltine neutralized amplifier.

Another form of balancing circuit due to L. A. Hazeltine known as the *Neutrodyne* is shown in Fig. 18. This type of circuit applies the same principle to the output circuit as the previous method did to the input. The conditions for balance are the same as Eq. (32). The coupling between L_a and L_b should again be approximately unity if the circuit is to remain balanced for a wide range of frequencies with a fixed adjustment of C_n, as L_a is shunted by the output impedance of the tube. This circuit has the advantage over the Rice circuit for receiving sets in that one set of plates of the tuning condenser is at filament or ground potential. This enables the rotors of the condensers to be mounted directly on a common shaft without requiring insulating bushings or couplings. A modification of this circuit has the neutralizing condenser C_n connected to a tap at some intermediate point in L_2 thus dispensing with the coil L_b. Lack of tight coupling between L_n and L_2 with this arrangement makes it more difficult to secure complete neutralization for a wide range of frequencies.

A circuit wherein all four of the bridge arms are condensers is shown in Fig. 19. The grid-plate capacity as well as the grid-filament capacity of the tube is involved, these two capacities serving as a pair of ratio arms. The conditions for a balance are

$$\frac{C_n}{C_a} = \frac{C_{gp}}{C_{gf}} \tag{33}$$

The value of C_a is usually about 100 $\mu\mu$f, which requires a value of C_n somewhat larger in size than the neutralizing condensers of the preceding circuits. In order to avoid the accumulation of a charge on

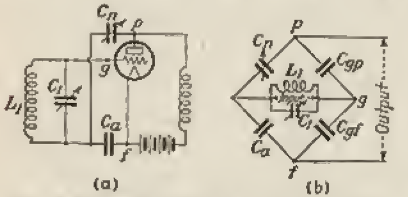


FIG. 19.—Capacity bridge neutralization of grid-plate capacity.

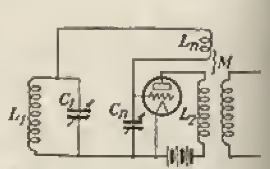


FIG. 20.—Mutual inductance bridge circuit.

the grid which may cause the tube to "block," C_n is usually shunted by a 250,000-ohm grid leak. The distributed capacity of a suitable choke coil whose natural frequency is below the frequency to be amplified can also be substituted for the condenser C_a.

Another form of circuit involving the principle of a mutual inductance bridge is illustrated in Fig. 20. The conditions for a balance are

$$\frac{M}{L_2} = \frac{C_{gp}}{C_{gp} + C_n} \tag{34}$$

Since C_n is in parallel with the grid-filament capacity of the tube, it is possible to utilize C_{gf} in place of an actual neutralizing condenser C_n and balance by proper adjustment of the mutual inductance between L_n and L_2.

16. Neutralizing Adjustments. The most convenient method of neutralizing the above circuits is to tune the amplifier to a signal in the h-f range of the receiving set. The tube filament of the stage to be neutralized is then opened, usually by slipping a piece of paper between the filament pin and the filament terminal in the tube socket. This destroys the repeater action of the tube and converts that portion of the circuit into its equivalent electrical network. The neutralizing condenser is then adjusted until the signal disappears. The filament is then lighted, and the procedure is repeated with the next stage.

When stray couplings are present, the value of balancing capacity required may vary with the frequency; hence, when exact neutralization is obtained at one frequency, the stage may be sufficiently unbalanced at some other frequency so that oscillations occur. In this case a compromise adjustment of C_n must be found which will hold the stage out of oscillation for the entire tuning range. This may not be possible if considerable stray coupling is present together with high gain per stage.

17. Neutralizing Power Amplifiers. Radio-frequency power amplifiers, such as are used in transmitting sets where sufficient power is available, can be neutralized by means of a suitable r-f ammeter in the output tank circuit. In these circuits provision is usually made to remove the plate voltage from the tube to be neutralized rather than to switch off the filament.

Figure 21 shows the last two stages of power amplification of a typical 1-kw broadcast transmitter. The first stage consists of two 75-watt screen-grid tubes in parallel which require no neutralization. The second stage is

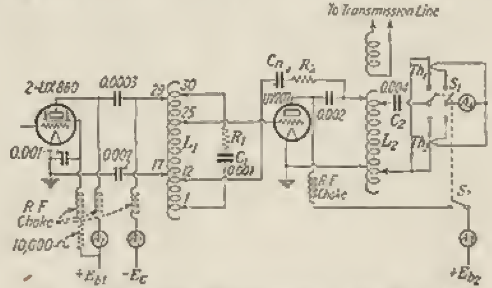


FIG. 21.—Broadcast transmitter power amplifier.

neutralized by means of the condenser C_n, which connects to the input tank circuit L_1C_1 at the point shown. The principle is the same as that of Fig. 17. The turns to which the various taps on L_1 are connected are indicated by the numbers. A 30-ohm resistance R_2 is connected in series with C_n to secure a more exact phase balance, since C_{gp} of the tube will have some losses associated with it and will therefore have a phase angle of less than 90 deg.

The neutralizing adjustment is made as follows: The switch S_1 is thrown to the top position inserting a low-range thermocouple Th_1 in the output tank circuit L_2C_2. At the same time the galvanometer A_1 is connected to the thermocouple, and the plate circuit is opened by S_2 which is mechanically connected with S_1. With excitation applied to the grid, the balancing condenser C_n is then adjusted until A_1 reads zero. The switch S_1 is then thrown to the lower position, closing the plate circuit and inserting a high range thermocouple Th_2 in the tank circuit, and at the same time transferring A_1.

18. Pentodes as R-f Amplifiers. The triode was superseded by the screen-grid tetrode, owing to the higher gains per stage obtainable without the need of neutralizing circuits. Still higher gains on the part of the pentode have enabled it to replace the tetrode in this field. The freedom from oscillation in these tubes is due to the reduction in the capacity between plate and control grid. This capacity is broken up in effect into two series condensers with the mid-point grounded to the filament, so far as r-f potentials are concerned, as will be seen from Fig. 22.

In r-f pentodes the suppressor grid is of further assistance in reducing this capacity, and values of C_{gp} of 0.01 $\mu\mu$f, or less, are obtained. Feedback of amplified output energy through the tube is thereby reduced to the point where stable operation with fair gain can be obtained at wave lengths of a few meters. These tubes may oscillate if too high a value of gain per stage is attempted. Capacitive coupling between grid and plate leads external to the tube must be carefully avoided by the use of adequate shielding.

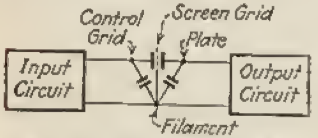


Fig. 22.—Elimination of coupling between input and output circuits by means of screen-grid tube.

The majority of these tubes for receiving purposes are of the remote cutoff or variable-mu type.[1] This feature enables a variable negative bias to be impressed on the control grid as a means of volume control without producing cross modulation and distortion when strong local signals are being received. With the conventional type of tube on strong signals the bias would have to be adjusted almost to cutoff in order to reduce the transconductance sufficiently to avoid overloading the last stage. Serious distortion of the modulated envelope would result if the tube were operated in this region of high curvature.

19. Radio-frequency Power Amplifiers. The low output and plate efficiency of class A amplifiers preclude their use in transmitters, and class B or class C operation is employed.

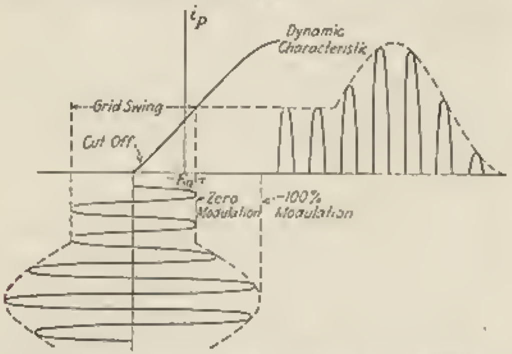


Fig. 23.—Characteristics of class B amplification.

Class B amplifiers are operated with a negative bias approximately equal to cutoff so that the plate current is almost zero when the alternating grid excitation is removed. With a sinusoidal voltage applied to the grid, the plate current consists of a series of half-sine waves,

[1] Ballantine and Snow, Reduction of Distortion and Cross-talk in Radio Receivers by Means of Variable-mu Tetrodes, *Proc. I.R.E.*, **18**, 2102, December, 1930.

similar to the output of a half-wave rectifier. The load impedance is adjusted so as to obtain an approximately linear dynamic characteristic, as shown in Fig. 23. The grid swings positive on excitation peaks, causing grid current to flow. Class B amplifiers are used in radiotelephone transmitters following the modulated stage. The power output obtainable from a given tube is much greater than with class A operation and the plate efficiency is much higher, having a theoretical maximum value of 78.54 per cent. As with a-f power amplifiers, tubes operating as class B r-f amplifiers may also be operated in push-pull.

A class C amplifier is one in which high output and plate efficiency are the primary considerations. The grid is negatively biased to a point considerably beyond cutoff, as shown in Fig. 24, so that the plate current is zero with no grid excitation. The latter is quite large and is often sufficient to cause the plate current to reach saturation on positive swings. Plate efficiencies in the vicinity of 90 per cent may be obtained with the larger tubes. These high efficiencies are made possible by

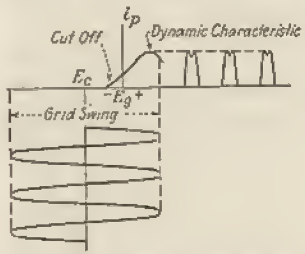


Fig. 24.—Class C operation.

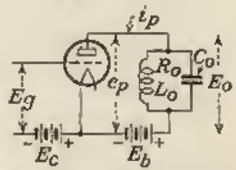


Fig. 25.—Schematic circuit of r-f power amplifier.

allowing the plate current to flow during less than 180 deg. of the cycle and only at a time when the plate potential is comparatively low. In radio-telegraph transmitters all stages are operated class C, while with radio telephony only the modulated amplifier and the stages preceding it are so operated.

The plate-current wave shapes in both cases are badly distorted, particularly with class C operation, and the output contains both odd and even harmonics. However, the tank circuit L_oC_o in Fig. 25 is resonant to the fundamental to which it offers a high impedance of the nature of a pure resistance. The impedance offered to the plate-current harmonics diminishes rapidly with the order of the latter; hence the voltage drop E_o across the tank circuit is very nearly sinusoidal in shape. The instantaneous plate voltage e_p will be the algebraic difference between the plate-supply voltage E_b and the drop E_o across the load.

Either triodes or screen-grid tetrodes may be used as power amplifiers. The latter have the advantage of not requiring neutralization. The screen-grid voltage in transmitting tubes is usually about 15 per cent of the plate-supply voltage, which is proportionally much lower than in receiving tubes. These tubes are difficult to construct for power outputs much greater than 500 watts, and, where larger outputs are required, triodes must be used.

20. Current and Voltage Relations. The instantaneous current and voltage relations for a class C amplifier are shown in Fig. 26. The potential e_p of the plate with respect to the filament is at a minimum during the time plate current is actually flowing. The power loss within the tube will be equal to the product of e_p and i_p averaged over a complete cycle. It is evident from Fig. 26 that this loss can be kept small by limiting the angle $2\theta_1$ during which time plate current actually flows. This will vary from 180 deg. in the case of a class B amplifier to perhaps as low as 60 deg for class C operation. It will also be noted that the grid-excitation voltage E_g is at its positive maximum when the plate voltage is a minimum. The minimum plate voltage should not

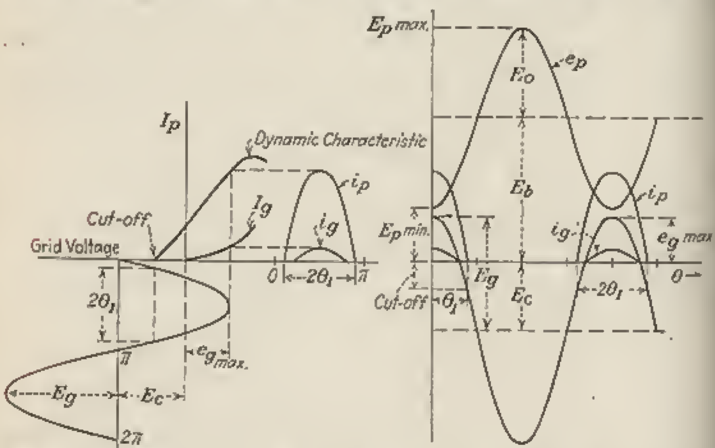


FIG. 26.—Instantaneous values of current and voltage in class C amplifier.

be allowed to fall below the value of $e_{g\,\max.}$ if excessive grid current is to be avoided. Ordinarily $e_{g\,\max.}$ is limited to about 80 per cent $E_{p\,\min.}$.

21. Circuit Calculations. In the design of a power amplifier the given data will include the frequency, the type of tube to be used, and the plate-supply voltage. The minimum plate voltage and the maximum positive value of the grid voltage are then selected, also the angle s_1. The required grid-excitation voltage will be

$$E_g = \frac{E_b}{\mu} + \frac{1}{1 - \cos\theta_1}\left(\frac{E_{p\,\min.}\cos\theta_1}{\mu} + e_{g\,\max.}\right) \tag{35}$$

The required C bias will be

$$E_c = E_g - e_{g\,\max.} \tag{36}$$

and the voltage across the tank circuit is given by

$$E_0 = E_b - E_{p\,\min.} \tag{37}$$

Corresponding pairs of plate and grid voltages can then be computed for increments of 5 or 10 deg. over the time interval $2\theta_1$ during which plate current flows. Since the various current and voltage waves are symmetrical

on either side of the vertical axis, it is only necessary to do this from zero to θ_1. A suitable table for this purpose is given below:

TABLE I

Given data:	Assumed values:	Computed values:
Tube..................	$E_{p\,\min}$..................	E_g.............. Eq. (35)
μ..................	$e_{g\,\max}$..................	E_c.............. Eq. (36)
E_b..................	θ_1..................	E_0.............. Eq. (37)

1	θ	0°	10°	20°	30°	40°	θ_1
2	$\cos\theta$	1	0.9848	0.9397	0.8660	0.7660	
3	$E_0 \cos\theta$						
4	$e_p = E_b - E_0 \cos\theta$	.					
5	$E_g \cos\theta$						
6	$e_g = E_g \cos\theta - E_c$						
7	i_p	y_0	y_1	y_2	y_3	y_4	0
8	i_g	..	......	......	......	......	0
9	$i_p \cos\theta$	y_0'	y_1'	y_2'	y_3'	y_4'	0
10	$i_g \cos\theta$	..	......	......	......	......	0

The values of plate and grid currents in lines 7 and 8 are obtained from the static characteristics of the tube for the computed pairs of instantaneous values of e_p and e_g in lines 4 and 6. The grid-current characteristic will also be necessary if the power required for grid excitation is to be determined.

The d-c component of plate current I_b will be the average value of i_p over a complete cycle and is given by

$$I_b = \frac{1}{18}\left(\frac{y_0}{2} + y_1 + y_2 + \cdots + y_{n-1}\right) \tag{38}$$

using the trapezoidal rule to determine the area under the curve for i_p. If 5-deg. intervals are used in Table I, the coefficient of Eq. (38) would be $\frac{1}{36}$.

The d-c component of grid current I_c can be found in a similar manner by substituting as ordinates the items of line 8 in Eq. (38).

The maximum amplitude of the fundamental component of the plate current is given by

$$\begin{aligned} I_{p1} &= \frac{2}{\pi}\int_0^{\pi} i_p \cos\theta\, d\theta \\ &= \frac{1}{9}\left(\frac{y_0'}{2} + y_1' + y_2' + \cdots + y_{n-1}'\right) \end{aligned} \tag{39}$$

using the trapezoidal rule to evaluate the definite integral. If 5-deg. intervals are used in Table I, the coefficient of Eq. (39) becomes $\frac{1}{18}$.

The maximum amplitude of the fundamental component I_{g1} of the grid current can be obtained in the same way by substituting the items of line 10 in Eq. (39).

22. Power Relations. The d-c power supplied to the circuit from the source of E_b is

$$P_{\text{input}} = E_b I_b \tag{40}$$

The power output to the tank circuit at the fundamental frequency is

$$P_{\text{tank}} = \frac{E_0 I_{p1}}{2} \tag{41}$$

since the tank impedance is of the nature of a pure resistance R_b at resonance. The required value of R_b is evidently

$$R_b = \frac{E_0}{I_{p1}} \tag{42}$$

and is related to the constants of the tank circuit by

$$R_b = \frac{L_0}{C_0R_0} \tag{43}$$

where R_0 is the apparent resistance of the tank coil and includes coupled resistance introduced by the useful load which is either inductively or conductively coupled to the tank coil. In the circuit of Fig. 25 the value of coupled resistance reflected into the tank coil would be the power absorbed from the tank divided by the square of the oscillatory tank current.

The resistance of the load required to fulfill the assumed operating conditions, as given by Eq. (42), will bear no simple relation to the plate resistance r_p of the tube as used in computations relating to class A power amplifiers, since r_p is infinite during the greater portion of the cycle under class C operation. From Eq. (43) it is seen that load impedance of the tank circuit may be varied by varying the ratio of L_0 to C_0. As the latter item is often a mica condenser of fixed capacity, a variation may be made in the value of R_b by using the tank inductance as an autotransformer, as illustrated in Fig. 27. The ratio of transformation will be approximately the turns ratio P/S, and, by moving the plate tap so as to alter the number of turns included in P, it is possible to change the load impedance as viewed from the tube by the square of the transformation ratio.

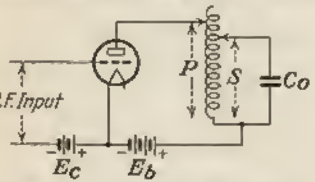


FIG. 27.—Tank-circuit inductance used as autotransformer to vary load impedance.

The power input to the grid is

$$P_{\text{grid input}} = \frac{E_gI_{g1}}{2} \tag{44}$$

The power amplification will be Eq. (41) divided by Eq. (44) and is

$$A_p = \frac{E_0I_{p1}}{E_gI_{g1}} \tag{45}$$

Power amplifiers are practically always operated with a fixed-bias voltage E_c instead of being self-biased by means of a grid leak and condenser, as with oscillators. This is because in the event of failure of the excitation voltage the self-bias would no longer function and the tube would be injured. A portion of the power input to the grid would be consumed across E_c and would charge the bias battery, if one were used. This power lost across the bias is E_cI_c, and the power consumed within the tube due to the flow of grid current is

$$P_g = \frac{E_gI_{g1}}{2} - E_cI_c \tag{46}$$

Since the grid is enclosed by the plate, the heating of the grid by P_g must be radiated by the plate in addition to its own losses.

The power loss within the tube which is to be dissipated at the plate in the form of heat, exclusive of the power loss in the filament, is

$$\text{Tube loss} = E_bI_b - \frac{E_0I_{p1}}{2} + \frac{E_gI_{g1}}{2} - E_cI_c \tag{47}$$

This expression may be used to check the assumed operating conditions from the standpoint of allowable plate dissipation.

The plate efficiency is defined as the ratio of the output to the tank circuit to the power supplied to the plate and is given by

$$\text{Plate efficiency} = \frac{E_0I_{p1}}{2E_bI_b} \tag{48}$$

With the allowable plate dissipation fixed, a moderate improvement in the plate efficiency will materially increase the useful output, and the maximum output will be obtained when the plate efficiency is made a maximum.

The effective value of the oscillatory current in the tank will be

$$I_L = \frac{E_0}{\sqrt{2(R_0^2 + \omega^2L_0^2)}} \tag{49}$$

Where the effective value of Q for the coil is high, the currents in the coil and condenser are approximately the same and will be given with sufficient accuracy for most purposes by

$$I_L = I_c = E_0\omega C_0 = \frac{E_0}{\omega L_0} \tag{50}$$

The preceding discussion has been based upon the series-fed circuit of Fig. 25, but the same equations and method of analysis will likewise apply to the case of parallel feed in Fig. 21. This latter arrangement is the one usually employed.

23. Class B Amplifiers. In order not to distort the envelope of the applied modulated wave in Fig. 23, the dynamic characteristic must be essentially linear, and the operating conditions are chosen so as to bring this about. When this is the case, the maximum amplitude of the fundamental component of the plate current is given by

$$I_{p1} = \frac{\mu E_g}{2r_p + R_b} \tag{51}$$

to a fair degree of approximation. The d-c component of plate current will then be

$$I_b = \frac{2}{\pi}I_{p1} = 0.637I_{p1} \tag{52}$$

The plate efficiency, from Eq. (48), becomes

$$\text{Plate efficiency} = \frac{E_0I_{p1}}{2E_bI_b} = \frac{\pi}{4}\frac{E_0}{E_b} \tag{53}$$

Since E_0 approaches E_b as a limiting value, it follows that the plate efficiency of a class B amplifier approaches 78.54 per cent as a limiting value. In actual practice it is usually about 65 per cent on excitation peaks at 100 per cent modulation and falls to about 33 per cent when the applied excitation voltage is unmodulated.

24. Tuning Adjustments. The tank circuit should always be adjusted to unity power factor so that minimum plate voltage may coincide with maximum plate current. A departure from this relation will lower the plate efficiency. This adjustment is usually made by tuning the tank circuit for minimum d-c plate current. Strictly speaking, minimum I_b may be used as an accurate measure of unity power factor only when C_0 of the tank is the element varied. The usual tuning adjustment is L_0, which is varied by means of a copper or aluminum disk rotated within the tank coil and acts as a single short-circuited turn. In this case maximum impedance will not occur at unity power factor, and L_0 should be adjusted to a value slightly lower than that which produces minimum I_b. If the effective value of Q for the tank is fairly high, the adjustments for maximum impedance and unity power factor practically coincide, in which case the current may be adjusted for minimum plate current with either tuning element the variable.

25. Modulated Amplifiers. If an a-f voltage is superimposed upon the d-c plate-supply voltage E_b of a class C amplifier having constant

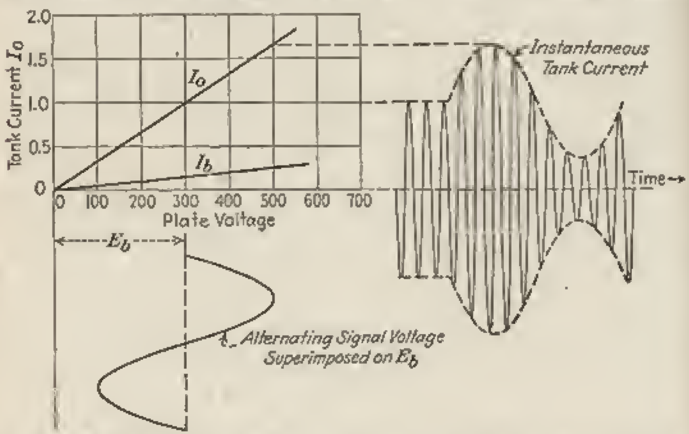


FIG. 28.—Modulation of class C amplifier by superimposing a-f signal voltage on plate-supply voltage

r-f excitation applied to its grid, the tank current I_0 may be made to rise and fall in amplitude as illustrated in Fig. 28. The schematic diagram of the circuit is shown in Fig. 29. A linear relation must exist between tank current and plate voltage if distortion is to be avoided. The relation between the plate voltage and I_b should also be fairly linear so that the modulator tube supplying the a-f power shall work into a constant load resistance, which will be equal to E_b/I_b, or, in general, $\Delta E_b/\Delta I_b$.

The grid excitation, grid bias, and tank-circuit impedance are adjusted so as to obtain the desired linear relations. The adjustments may be checked by varying E_b from zero to twice normal value and plotting I_0 and I_b against E_b as in Fig. 28. The value of plate-supply voltage impressed upon the modulated amplifier is somewhat lower than the normal value used for unmodulated operation in order to avoid excessive plate heating on modulation peaks. The grid bias E_c required is approximately twice the value of cutoff for the tube, and the tank impedance is usually higher than with unmodulated operation. The plate efficiency is lower than with unmodulated amplifiers and is usually in the neighborhood of 60 per cent, depending upon the size of the tube used. Either triodes or tetrodes may be used.

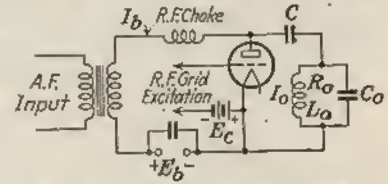


FIG. 29.—Plate-modulated class C amplifier.

The continuous power output with 100 per cent modulation is 1.5 times the power at zero modulation. The output on modulation peaks will be four times the unmodulated carrier output. This increase in the power output when modulated must be furnished by the a-f input

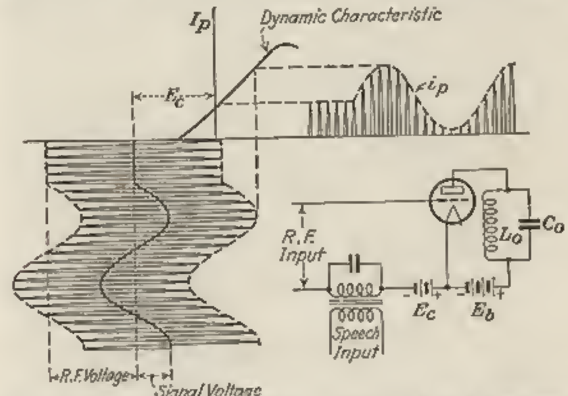


FIG. 30.—Schematic circuit and operation details of grid-modulated amplifier.

from the modulator tubes. The amount of a-f power required varies with the square of the modulating factor, so that the modulator tubes must be capable of furnishing a sizable amount of audio power if 100 per cent modulation is to be attained.

While the plate-modulated amplifier has been the most widely used, other methods requiring very much less audio power can be employed. Instead of varying the voltage applied to the plate of the modulated amplifier, it is possible to secure similar results by varying the magnitude of the C bias at an a-f rate. The schematic circuit is shown in Fig. 30, together with the details of operation. The signal voltage cyclically adds to and subtracts from the fixed biasing voltage E_c, causing the amplitude of the plate-current impulses to rise and fall. The plate-current wave shapes will be similar to those of the class B amplifier of Fig. 23, except that the angle $2\theta_1$ during which plate current flows will vary with the modulation. The mode of operation changes from an underexcited class C amplifier when unmodulated to a class B amplifier on modulation peaks, assuming complete modulation. The advantage of this method over plate modulation is that very little a-f energy is required for complete modulation. The modulating source is only required to furnish a portion of the grid-excitation losses of the amplifier in this case. The plate efficiency is somewhat lower, and freedom from distortion is more difficult to secure.

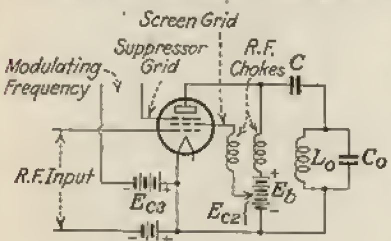


FIG. 31.—Screen-grid pentode used as modulated class C amplifier.

Another method is to insert the modulating voltage in the suppressor-grid circuit of a screen-grid type of power pentode operating as a class C amplifier, as shown in Fig. 31. The suppressor grid is biased negatively by a moderate amount and swings positive on modulation peaks, during which time suppressor-grid current flows. The power represented by this current has to be furnished by the modulating source, but it is negligible in comparison to the demands of a plate-modulated amplifier. The distortion is low with moderately high percentages of modulation but becomes appreciable at 100 per cent.

26. Doherty High-efficiency Amplifier.[1] The plate efficiency of a class B amplifier varies between about 33 and 65 per cent from 0 to 100 per cent modulation, resulting in a rather low all-day efficiency in view of the average per cent modulation of a broadcast station. Consequently an appreciable reduction could be effected in the energy requirements of a transmitter if this efficiency could be raised and kept constant. The Doherty amplifier accomplishes this desirable result in the following manner:

Two tubes, effectively in parallel, supply power to a common tank circuit, as shown schematically in Fig. 32. Tube T_1 is operated so that its output voltage E_1 is at its maximum permissible value when the unmodulated carrier voltage E_c is applied to the grid. The grid bias on T_2 is made sufficiently negative so that the output current I_2 is about zero at this value E_c of the carrier voltage. This high value of E_1 with an excitation voltage of E_c impressed is brought about by having the tube work into a load impedance of $2R$, or twice the value of tank impedance that would be ordinarily used. This value of $2R$ is brought about by the properties of the impedance-inverting network in the plate circuit, which is the equivalent of a quarter-wave line. These lines have a sending-end impedance Z_s which is given by the relation

$$Z_s = \frac{Z_L^2}{Z_r} \tag{54}$$

[1] DOHERTY, W. H., A New High Efficiency Power Amplifier for Modulated Waves, *Proc. I.R.E.*, p. 1183, September, 1936.

where Z_r = terminating impedance ($R/2$)
Z_L = characteristic impedance of the line, which is equal in this case to R when the reactive series and shunt arms of the simulating network have the values given in the figure.

When the carrier voltage increases to a value greater than E_c (reaching $2E_c$ at 100 per cent modulation), T_2 begins to furnish power to the tank circuit. However, this causes the impedance of the tank, as viewed from the end of the network, to *rise*. But this apparent rise in Z_r, from Eq. (54),

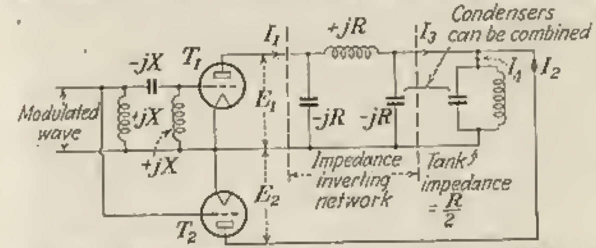
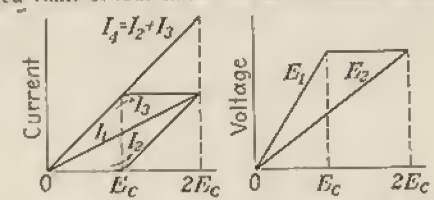


FIG. 32.—Schematic diagram of a Doherty amplifier.

causes a *reduction* in Z_s. Consequently, the output current I_1 of tube T_1 rises, even though E_1 remains constant. The increasing grid excitation maintains E_1 as the plate load impedance Z_s falls. As the excitation increases beyond the unmodulated amplitude E_c, T_2 contributes more and more power to the tank and thereby permits T_1 also to supply more power. When the excitation reaches a value of $2E_c$, corresponding to the instantaneous peak of a completely modulated wave, half of the power in this tank is being contributed by T_2. The network is at that instant effectively terminated in R ohms instead of the original value of $R/2$, permitting T_1 to deliver twice its initial power output. The total power delivered to the tank circuit is then the required value of four times the unmodulated value.



(a)-Grid excitation (b)-Grid excitation

FIG. 33.—Current and voltage relations in a Doherty amplifier.

The variations in the various currents in Fig. 32 are shown in Fig. 33*a*, and the voltages E_1 and E_2 vary as shown in Fig. 33*b*. One of the characteristics of the impedance-inverting network shown in Fig. 32 is that the current I_2 will lag 90 deg. behind E_1. Consequently a network producing a similar phase shift, but in the opposite direction, is inserted in the grid circuit of T_1, so that the currents I_2 and I_3 will be in phase with each other. Reversing the signs of the reactances in the series and shunt arms of the grid network, as shown, will produce the desired leading phase shift of 90 deg.

The Doherty method of operation enables a plate efficiency of slightly more than 60 per cent to be secured when the carrier is unmodulated. The distortion is somewhat greater than with class B operation, Fig. 33.

depicting the ideal characteristics, but by using reversed feedback it is possible to meet all the requirements of high-fidelity broadcasting.

27. Frequency Multipliers. The plate current of a class C amplifier is badly distorted and contains a large percentage of harmonics. It is possible to resonate the tank circuit to one of these harmonics and cause it to absorb power at the harmonic frequency. The impedance offered to the fundamental and the balance of the harmonics will be small; hence little power will be absorbed at these frequencies.

Frequency multipliers are used to obtain higher frequencies than can be readily produced by crystal-controlled oscillators. Quartz crystals for high frequencies become rather fragile and are apt to crack in service. To secure crystal control of the frequency in the case of short-wave transmitters, the crystal is ground to oscillate at some l-f multiple of the transmitted frequency. The output of the crystal-controlled oscillator is then impressed on one or more amplifiers connected in cascade and adjusted to multiply the frequency. The usual practice is to double the frequency with each stage, and, while greater multiplications than this can be obtained, the output falls off rapidly as higher multiplications per stage are attempted. If a push-pull circuit is being used as the frequency-multiplying stage, the output tank circuit will have to be tuned to the third harmonic of the input voltage, since even harmonics will cancel in the output circuit. A class C amplifier having a plate efficiency of 80 per cent would show an efficiency of about 70 per cent when used as a frequency doubler. The instantaneous current and voltage relations in a frequency doubler will be similar to Fig. 26 except that the frequency of e_p will be twice as great and will therefore be low in value for a shorter time interval. This requires a smaller value of θ_1 in order to keep the losses within the tube small. These losses are proportional to the product of the instantaneous values of e_p and i_p and can be minimized by restricting the flow of plate current to a smaller interval of time. This calls for values of E_g and E_c somewhat higher than with the conventional type of class C amplifier. Either triodes or tetrodes can be used. The former will not need to be neutralized, as the input and output circuits are tuned to different frequencies and hence will not oscillate.

SECTION 13

RECEIVING SYSTEMS

By G. L. Beers, B.S.[1]

1. Classification. The following is a classification of radio receivers according to their operating principle.

1. Tuned radio frequency.
2. Superheterodyne.
3. Regenerative.
4. Superregenerative.

2. Tuned-radio-frequency Receivers. Tuned-radio-frequency (t-r-f) receivers are those which obtain their selectivity and r-f amplification through the use of circuits which function at the frequency of the incoming signal.

Tuned r-f receivers use from two to six circuits which are tuned simultaneously by means of a single tuning control. A gang condenser, which consists of several variable condensers assembled in a single unit, is used to vary the frequency of the tuned circuits. The series resistance of a conventional tuned circuit, whose frequency is varied by means of a variable condenser, increases with frequency. The selectivity of t-r-f broadcast receivers varies in a ratio of about 3:1 from one end of the broadcast range to the other. One or two of the tuned circuits in a t-r-f receiver are generally used in the antenna-input system and the remainder are used to provide the coupling between the stages of the r-f amplifier. One or two stages of a-f amplification are used in the audio portion of the receiver. Tuned r-f receivers are best suited for use where the selectivity requirements are not extreme.

3. Superheterodyne Receivers. In the superheterodyne receiver the received voltage is combined with a voltage from a local oscillator and converted into a voltage of a lower or intermediate frequency which is then amplified and detected to reproduce the original signal wave.

The superheterodyne receiver utilizes the essential components of a t-r-f receiver and, in addition, a frequency converter and i-f amplifier. The frequency converter consists of a variable-frequency oscillator and a detector. The function of the frequency converter is to change the frequency of the received signal to the i.f. The oscillator and t-r-f circuits in superheterodyne receivers are usually tuned simultaneously by means of a gang condenser. A constant frequency difference is maintained between the oscillator and r-f circuits either through the use of a combination of fixed shunt and series condensers in the oscillator circuit in conjunction with a gang condenser in which all of the variable condensers are identical in capacity or through the use of a gang condenser

[1] Engineering Department, RCA Manufacturing Company, RCA Victor Division, Camden, N. J.

in which specially shaped plates are used in the oscillator variable condenser. The i-f amplifier uses two or three transformers, which usually contain two coupled circuits with the coupling adjusted to provide the so-called *band-pass filter* characteristics. The i-f amplifier provides the major portion of the amplification and selectivity. Since the characteristics of this amplifier are independent of the frequency to which the receiver is tuned, the sensitivity and selectivity of a superheterodyne receiver are usually very uniform throughout its tuning range. The r-f circuits are used primarily for eliminating certain types of interference which are common to this type of receiver. The performance of the superheterodyne receiver is in general superior to that of any other type of receiver in use today.

4. Regenerative Receivers. In a regenerative receiver the following action takes place: The received voltage is impressed on the grid of a vacuum tube. A portion of the resultant voltage which appears in the plate circuit of the tube is fed back to the grid circuit in the proper phase relation to increase the applied grid voltage. The effect of this action is to reduce the effective resistance of the resonant circuit to which the signal is applied and, thereby, provide considerable amplification of the received signal.

Regenerative receivers are usually provided with two controls, one for tuning the receiver and the other for controlling the amount of feedback energy. If the feedback is increased beyond a certain value, sustained oscillations are produced. It is common practice to tune regenerative receivers while sustained oscillations are being produced, as the beat frequency produced between the carrier wave of the transmitting station and the locally produced oscillations indicates when the receiver is properly tuned. This method of tuning is called the "zero-beat" method as the tuning of the receiver is adjusted so that the beat note decreases in frequency till it is no longer audible. When a conventional regenerative receiver is tuned in this way, interference is produced in near-by receivers which are tuned to the same station. A stage of tuned r-f amplification is sometimes used between the antenna and the regenerative circuit to reduce the possibility of producing this type of interference. The regenerative receiver is quite sensitive considering the number of tubes which are used. It is not very selective since only a single tuned circuit is generally used. They are now practically obsolete as broadcast receivers, although they are still used to a limited extent in marine receivers and in short-wave work.

5. Superregenerative Receivers. A superregenerative receiver is a regenerative receiver in which sustained oscillations are prevented by the periodic variation of the effective resistance of the resonant circuit to which the received signal is applied.

In the superregenerative receiver oscillations are permitted to build up at a periodic rate in a resonant circuit tuned to the frequency of the received signal wave. Sustained oscillations in this circuit are prevented by the application of a quenching frequency potential to the grid of the superregenerative tube which periodically affects the tube characteristics in such a way as to stop the oscillations. The quenching frequency may be supplied either by a separate oscillator or by the superregenerative tube itself. The audio system of this type of receiver is usually provided with an a-f filter to remove the quenching frequency from the audio output. An r-f stage is frequently used ahead of the detector to prevent

energy being transferred from the superregenerative circuit to the antenna. A signal input of 50 to 100 μv will give an intelligible signal, although an input of 500 to 1000 μv is generally necessary to reduce the noise to a satisfactory value. Harmonics of the quench frequency beating with the received signal make a source of interference if the ratio between signal and quench frequencies is not 100:1 or more. The superregenerator is still used in some police automobile installations but is being replaced by the superheterodyne because of the better signal-to-noise ratio and selectivity which this receiver provides.

6. Method of Rating. Receiving sets are generally rated on the basis of the following characteristics: (1) sensitivity; (2) selectivity; (3) fidelity; (4) overload level; (5) power consumed.

1. The *sensitivity* is that characteristic which determines to how weak a signal it is capable of responding. It is measured quantitatively in terms of the input voltage required to give a standard output.
2. The *selectivity* is the degree to which the receiver is capable of differentiating between the desired signal and signals of other carrier frequencies. This characteristic is not expressible by a single numerical value but requires one or more graphs for its expression.
3. The *fidelity* of a radio receiver is the degree to which it accurately reproduces at its output terminals the signal which is impressed upon it. As applied to a radio receiver, fidelity is measured by the accuracy of reproduction at the output terminals of the modulation of the received wave.
4. The *overload level* of a receiver is the maximum power output which can be obtained from it when the output voltage does not contain more than 10 per cent of total harmonics.

7. Method of Testing. A standardized method of testing radio receivers has been established by the Institute of Radio Engineers and is described in detail in the Year Book of the Institute. The following is a brief summary of the procedure:

1. *Definition of Terms.*

a. Sensitivity, selectivity, fidelity, and maximum undistorted output (see Method of Rating).

b. Normal test output: An a-f power output of 0.5 watt in a standard dummy load connected across the output terminals of the receiver is the normal test output of a broadcast radio receiver except when the maximum power output is less than 1 watt and more than 0.1 watt, in which case the normal test output is 0.05 watt.

c. Normal radio-input voltage: This term represents the r-m-s r-f voltage modulated 30 per cent at 400 cycles which results in normal test output at resonance.

d. Standard test frequencies: In the testing of a broadcast radio receiver, the seven standard carrier frequencies are 540, 600, 800, 1,000, 1,200, 1,400, and 1,600 kc. When tests at only three carrier frequencies are required, the carrier frequencies of 600, 1,000, and 1,400 kc are used.

2. *Equipment Required.*

a. A signal generator: This consists of a shielded vacuum-tube oscillator whose frequency can be varied from 500 to 1,600 kc. An a-f oscillator is provided to modulate the r-f oscillator by a known amount at any frequency from 40 to 10,000 cycles. A calibrated attenuator is used to impress a known potential on the standard antenna connected to the receiver. The attenuator system should be such as to allow a range of voltage impressed on the standard antenna unit from 1 to 200,000 μv.

b. Standard antenna: The standard antenna for a broadcast radio receiver not having a self-contained antenna is an antenna having substantially the same impedance as a series circuit containing a capacity of 200 $\mu\mu$f, a self-inductance of 20 μh, and a resistance of 25 ohms.

c. Standard dummy load: This is a pure resistance whose value is equal to the 400-cycle impedance of the loud-speaker which is supplied with the radio receiver. The load resistor should be capable of dissipating the maximum power output of the receiver without an appreciable change in resistance. An output filter is provided for preventing the flow of d.c. through the load resistor when testing sets which normally have d.c. in their output circuit. A vacuum-tube voltmeter or equivalent device is used for determining accurately the r-m-s voltage across the load resistor.

d. Harmonic-measuring circuit: For this purpose a harmonic analyzer capable of measuring frequencies up to 15,000 cycles is recommended. The instrument should have sufficient frequency discrimination to measure harmonics which are 0.5 per cent or less of the fundamental.

3. *Tests.*

a. Sensitivity: The sensitivity is determined by impressing an r-f voltage, with 400 cycles, 30 per cent modulation, in series with a standard antenna and adjusting the intensity of the input voltage until normal test output is obtained for carrier frequencies between 550 and 1,500 kc.

b. Selectivity: The selectivity of a receiver is determined by tuning it to each test frequency in succession, with the receiver in the same condition as in the sensitivity test, and measuring the r.f. necessary to give normal test output at steps not greater than 10 kc at least up to 100 kc on either side of resonance or until the radio-input voltage has increased to 10,000 times or more if the measuring equipment permits.

c. Electric fidelity: This is determined by tuning the radio receiver to each standard test frequency in succession with the receiver in the same condition as in the sensitivity and selectivity tests, adjusting the impressed voltage to the normal radio-input voltage and then varying the modulation frequency from 40 to 10,000 cycles at 30 per cent modulation and constant r-f input voltage throughout, taking readings of relative output voltage at convenient modulation frequencies.

4. *Additional Tests.*

a. Determination of the overload level: This is determined by increasing in successive steps the r-f input to the receiver (with modulation adjusted to 30 per cent at 400 cycles) and measuring both the power output and the percentage harmonics. The overload level of the receiver is the least power output which contains a total harmonic distortion of 10 per cent (r-m-s voltage).

b. Volume-control tests: This test is a determination of the effect of the volume control on the sensitivity, selectivity, and fidelity.

c. Test for hum: For determining the hum voltage, a filter is connected between the output of the receiver and the voltmeter. This filter has a characteristic which evaluates the various hum components according to their quantitative effect on the human ear.

8. Design of Receiving Systems. The majority of receiving sets in use today are broadcast receivers designed to cover the frequency range of from 550 to 1,500 kc. The essential electrical elements of a modern broadcast receiver may be classified as follows:

1. Radio-frequency system.
2. Audio-frequency system.
3. Volume-control system.
4. Power-supply system.
5. Loud-speaker.

9. Radio-frequency System. *Antenna-input Systems.* The antenna-input system transfers the signal wave intercepted by the antenna to

the grid of the first tube in the receiver. The antenna-input system also contributes to the over-all performance as follows:

1. One or more t-r-f circuits in the antenna-input system provide selectivity for the separation of stations as well as the prevention of cross modulation.
2. A reduction in tube noise for a given sensitivity is obtained through the step-up in voltage provided by the use of tuned circuits in antenna-input systems.

A typical antenna-input system is illustrated in Fig. 1. Since there is considerable variation in the characteristics of receiving antennas used, the value of the antenna-coupling inductance is chosen so that the antenna system is always tuned to a frequency below the tuning range of the receiver. If the antenna circuit becomes resonant in the tuning range of the receiver, the first tuned circuit in an unicontrolled receiver will be thrown out of alignment with the remainder of the receiver and the over-all performance will be seriously affected. Figure 2 shows the voltage step-up between the antenna and the grid of the first tube which is obtained from such an arrangement. Two coupled tuned circuits are sometimes used between the antenna and the grid of the first tube. This reduces the voltage gain to approximately half that obtained with the single tuned circuit but

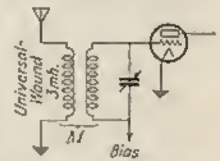


FIG. 1.—Antenna-input system.

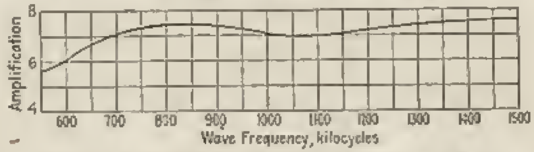


FIG. 2.—Amplification of input system of Fig. 1.

increases the selectivity and therefore reduces the possibility of cross modulation in the first tube of the receiver. An antenna-input system is shown in Fig. 3, which provides considerably greater coupling between the antenna and the first tuned circuit. This system is employed in automobile receivers where the signal intercepted by the antenna is usually quite small. By connecting a small inductance in series with the antenna so that a series-tuned circuit is formed which is resonant at approximately 2,000 kc, this system will provide a voltage gain which varies from 10 at 600 kc to 20 at 1,400 kc. Another antenna input system which is used extensively in automobile receivers, particularly those which are designed for a specific car and antenna, is to connect the antenna to a tap, approximately 30 to 50 per cent on the coil in the first tuned circuit.

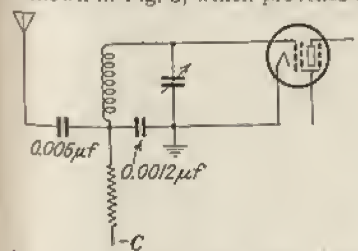


FIG. 3.—Closely coupled antenna-input system.

A number of broadcast radio receivers employ a loop antenna. In this case the loop comprises the inductance in the first tuned circuit. In some console receivers the loop is rotatable by means of a control on the front panel of the receiver. In midget-type receivers the position of the loop is fixed, and it is sometimes necessary to orient the complete receiver to obtain the maximum signal voltage from a desired station. The directional properties of the loop antenna can frequently be used to minimize interference, provided the source of the interference, the station transmitting the desired signal, and the receiver are not in the same plane.

10. Radio-frequency Amplifiers. The types of r-f amplifiers in use in broadcast receivers may be classified as tuned, fixed-tuned, and untuned.

Tuned r-f amplifiers are those which amplify a narrow band of frequencies and are provided with a control by which the position of this band of frequencies may be moved over a wide frequency range.

Untuned r-f amplifiers are not provided with a tuning control and are designed to amplify a wide band of frequencies.

Fixed-tuned r-f amplifiers are those which pass a narrow band of frequencies and whose resonant frequency is not varied with the tuning of the receiver. The i-f amplifier of a superheterodyne receiver is an amplifier of this type.

11. Single-tuned Circuit T-r-f Amplifiers. The selectivity and amplification which can be obtained from a conventional t-r-f amplifier stage are a function of the effective resistance of the tuned circuit used in the interstage transformer. Since the selectivity provided by a t-r-f amplifier cannot be increased beyond a certain limit without serious attenuation of the high modulation frequencies, the useful amplification which can be obtained from an amplifier stage is therefore limited. The selectivity and amplification which a t-r-f amplifier will provide can be calculated. From a practical standpoint of receiver design, however, it usually requires less time and is more accurate to determine the characteristics of a particular transformer experimentally by laboratory measurements since a determination of the effective resistance of the tuned circuit is necessary even if the characteristics of the transformer are to be calculated. It is likewise difficult to take into consideration the effects of regeneration and the proximity of shielding, etc., in a mathematical consideration of r-f transformer characteristics. The ratio of reactance to effective resistance or $\omega L/R$ of the tuned circuits used in r-f transformers for broadcast receivers is usually between 75 and 150 throughout the broadcast frequency range. The diameter of the coils used in the t-r-f circuits of broadcast receivers varies from ½ to 1 in. and the size of the copper wire used for winding the coils is usually between Nos. 20 and 35 B. & S., the larger wire being used in the short-wave coils of "all-wave" receivers. Litz wire and cores molded of finely divided iron particles are frequently used to improve the Q of t-r-f transformer coils. The position of the core within the coil is usually variable so that the inductance of the coil can be adjusted to a desired value.

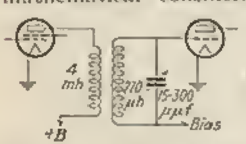


Fig. 4.—T-r-f interstage transformer.

Considerable shielding is required in screen-grid r-f amplifiers to prevent coupling between circuit elements and wiring which may likewise cause oscillations. It is common practice to locate the grid circuits and plate circuits associated with each tube in separate metal compartments to prevent coupling between them.

Figure 4 illustrates the type of t-r-f transformer which is used in the majority of broadcast receivers. The primary of the transformer is a small "universal-wound" coil which is either wound on a form of small diameter so that it can be mounted inside the secondary or is wound directly on the end of the same form as the secondary. The secondary is wound on a piece of tubing made of bakelite or some similar material. The primary is coupled electromagnetically to the secondary. The amplification and selectivity characteristics obtained with this transformer when used with an r-f pentode, having a transconductance of 1,000 micromhos, are shown in Fig. 5.

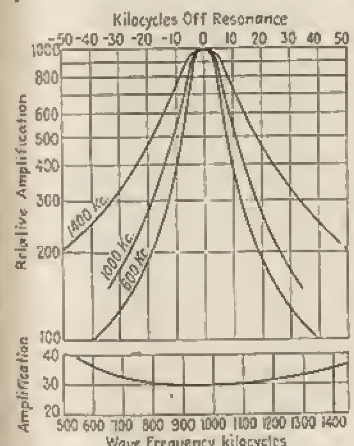


Fig. 5.—Characteristics of transformer in Fig. 4.

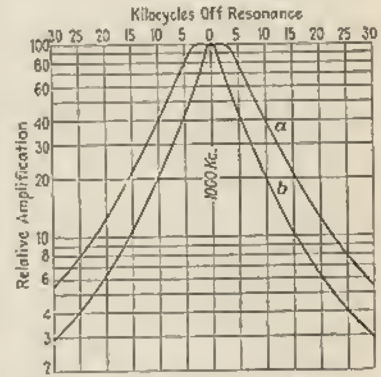


Fig. 6.—Selectivity comparison of single and coupled tuned circuit.

12. Coupled Tuned-circuit T-r-f Amplifiers. A number of broadcast receivers use one or more transformers in which two tuned circuits are used. The two circuits are coupled near the point of critical coupling. The advantage obtained through the use of this type of transformer is that a considerable improvement is obtained in the shape of the selectivity characteristic. Figure 6 illustrates this improvement. Curve *a* shows the characteristic obtained with two coupled tuned circuits, and curve *b* shows the characteristic obtained with two similar tuned circuits in cascade. The width of the top of the resonance curve of a coupled tuned-circuit transformer depends on the coupling between the two circuits. The flatness of the top of the curve depends on the effective resistance of the tuned circuits. By using slightly greater than critical coupling at the l-f end of the broadcast range and less at the h-f end of the range, the selectivity of this type of transformer can be made more uniform over the broadcast range than one using a single tuned circuit. Figure 7 shows the selectivity characteristic obtained from a transformer of this type. The voltage gain provided by a coupled tuned-circuit t-r-f transformer is approximately one-half that which can be obtained from a transformer using a single tuned circuit.

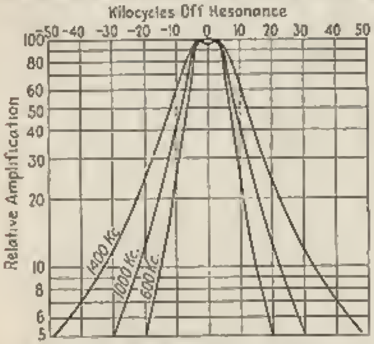


Fig. 7.—Selectivity characteristics of coupled tuned-circuit t-r-f transformer.

13. Untuned R-f Amplifiers. A stage of untuned r-f amplification is sometimes used in receivers where additional gain is desired without the need for the additional selectivity which would be provided by a stage of t-r-f amplification. Figure 8 shows an untuned amplifier stage which has been used in broadcast receivers. A tube in this amplifier stage will provide a gain of approximately 6 throughout the broadcast frequency band; 4.5 at 5 Mc and 1.5 at 15 Mc.

14. The i-f amplifier in a superheterodyne is the major factor in determining the receiver sensitivity and selectivity.

Modern superheterodyne receivers use an i.f. at or near either 175 or 455 kc. One hundred seventy-five kilocycles is used to a limited extent in receivers which are designed to cover only the tuning range from

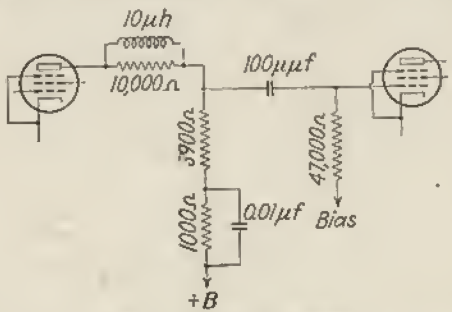


Fig. 8.—Untuned r-f stage.

550 to 1,500 kc, while 455 kc is used in receivers whose tuning range includes the international short-wave bands. Nearly all i-f amplifiers make use of transformers employing two coupled tuned circuits. The selectivity characteristic provided by a transformer of this type may be made substantially flat-topped if the coupling between the two tuned circuits is adjusted to near the critical value.

The two characteristics which are given the most consideration in the design of an i-f amplifier are gain and selectivity. These characteristics may either be calculated or determined experimentally. The gain in a coupled

tuned-circuit i-f stage with both circuits tuned to resonance is equal to

$$\left|\frac{E_1}{E_2}\right| = S_m \times \frac{\omega M}{r_1 r_2 + \omega^2 M^2} \times \frac{1}{\omega^2 C_1 C_2}$$

The selectivity characteristic may be determined by

$$\left|\frac{E_1}{E_2}\right| = S_m \times \frac{\omega M}{r_1 r_2[1 - 4Q_1 Q_2 B^2 + j2(Q_1 + Q_2)B] + \omega^2 M^2} \times \frac{1}{\omega^2 C_1 C_2}$$

where E_1 is the voltage developed across the secondary of the transformer; E_2 is the voltage applied to the grid of the amplifier tube; S_m is the transconductance of the amplifier tube; M is the mutual inductance between primary and secondary; r_1 and r_2 are the effective series resistances of the primary and secondary; Q_1 and Q_2 are the $\omega L/r$ of the primary and secondary, respectively; B is $(f - f_0)/f_0$, where f_0 is the common resonant frequency and f is any other frequency; and C_1 and C_2 are the primary and secondary capacities.

To obtain maximum gain in an i-f amplifier stage, the L/C ratio should be the maximum which will give the desired frequency stability. If the L/C ratio of the tuned circuits is made too high, the variations in the inter-electrode capacity of the tubes may cause a serious misalignment of the tuned circuits. The capacity used to tune the intermediate frequency circuits is therefore seldom less than 30 or 40 μμf.

The width of the frequency band which a coupled tuned-circuit transformer will pass is controlled by the coupling between the two tuned circuits and the effective resistance of the circuits. If increasing the coupling between the circuits until the transformer passes the desired frequency band causes the top of the selectivity characteristic to become double-peaked; it can be made flat by increasing the effective resistance of one or both of the tuned circuits. To obtain the same selectivity characteristic in kilocycles at 455 kc as at 175 kc, the Q of the tuned circuits must be approximately 2.5 times as great. To secure compact tuned circuits having the Q required (80 to 100) to give satisfactory selectivity at 455 kc, the coils are frequently wound in sections using Litz wire. A two-to-one improvement in the Q of coils suitable for a 455-kc i-f transformer can generally be obtained through the use of cores molded of finely divided iron particles and an insulating binder.

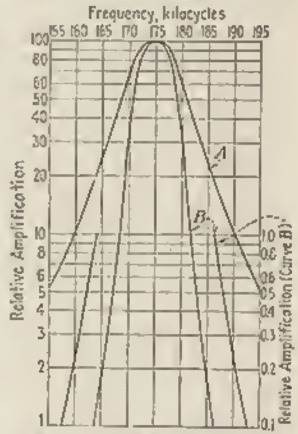


Fig. 9.—Intermediate-frequency selectivity characteristics: Curve A, one stage; curve B, three stages.

A typical i-f transformer consists of two universal-wound coils assembled on an insulating support such as a wooden rod or piece of bakelite tubing. These two coils constitute the inductive elements of two tuned coupled circuits. One of the tuned circuits is connected in the plate circuit of the amplifier tube and the other in the grid circuit of the succeeding tube. The electromagnetic coupling between these circuits is determined by the spacing between the coils. The tubing or rod on which the coils are wound is mounted on a plate of insulating material such as porcelain or isolantite. On this plate are frequently mounted two small adjustable condensers that are used to tune the two coupled tuned circuits. Care must be exercised in the design of these condensers to ensure that the capacity of the condensers remains constant after adjustment. In another design fixed condensers are used, and each circuit is tuned by adjusting the position of a molded iron core associated with each coil. The entire transformer assembly is enclosed in a metal container which serves both to protect the unit and shield it electrically.

The selectivity characteristic provided by a typical 175-kc i-f transformer is shown in curve A (Fig. 9). These characteristics are also representative of the best 455-kc transformers. A voltage amplification of several hundred can readily be obtained with a single i-f transformer and a modern r-f pentode having a transconductance in excess of 1,500. The voltage gain for the usual i-f amplifier, consisting of three transformers and two amplifier tubes, when measured from the grid of the first detector to the grid of the second detector is usually from 15,000 to 30,000. The voltage gain in the amplifiers using two transformers and one amplifier tube is 10,000 or less. The amplification in the three-transformer amplifiers is usually held considerably below the optimum value to prevent instability.

15. Frequency Converters. In a superheterodyne receiver the received signal wave is changed to a signal wave of an i.f. This change is accomplished through the medium of a frequency converter, which consists of a detector and variable-frequency oscillator. The detector is frequently called the *first detector* owing to its position in the circuit.

In some receivers the first detector is a negatively biased r-f pentode and operates due to the curvature of the E_g-I_p characteristic. The received signal voltage and a voltage from the local oscillator are both impressed on the grid of this detector. The beat-frequency potential produced by the rectification of these two currents is impressed on a tuned circuit connected in the plate circuit of the detector. The majority of receivers, however, employ a pentagrid converter as the combined oscillator and first detector. The coupling between the oscillator and first detector, when this tube is used, is obtained through the electron stream in the tube. The reaction frequently encountered with two-tube frequency converters that employ electromagnetic or electrostatic coupling between the oscillator and first detector circuits is thus avoided. This freedom from direct coupling between the oscillator and first detector resulting from the use of a pentagrid converter makes it possible to prevent the radiation of the oscillator energy by the antenna system without employing an r-f amplifier stage ahead of the first detector. The efficiency of a frequency converter is a function of the conversion transconductance of the tube employed as the first detector. Conversion transconductance is defined as the ratio of the i-f current through the i-f transformer primary in the plate circuit of the first detector to the r-f signal applied to its grid. The conversion transconductance of a typical pentagrid converter is generally somewhat higher than that obtained from an r-f pentode used as a first detector and may vary from 300 to 500 micromhos, depending on the potentials applied to the several electrodes.

In several receivers a separate oscillator tube is used in conjunction with a pentagrid converter. Greater flexibility in the design of the oscillator circuits is thus permitted since the separate tube has a considerably higher transconductance than the triode portion of the pentagrid converter. This advantage is particularly important in receivers designed to cover frequency ranges up to 30 Mc, owing to the difficulty of obtaining a stable oscillator with the desired output and frequency stability at such frequencies.

The major problems in the design of the frequency converter for a unicontrolled superheterodyne receiver are:

1. To maintain a constant-frequency difference between the oscillator and r-f circuits.
2. To minimize variations in the oscillator frequency with variations in the supply voltage and variations in tubes, etc.
3. To maintain a constant oscillator voltage on the detector grid throughout the tuning range of the receiver.
4. To minimize radiation from the oscillator in order to prevent interference in near-by receivers.

16. Methods of Maintaining Constant-frequency Difference. Three methods have been used to maintain a constant-frequency difference between the oscillator and first detector in unicontrolled superheterodyne receivers.

The *first method* makes use of straight-line-frequency condensers and requires that the oscillator rotor be displaced with respect to the r-f circuit rotors by an amount sufficient to give the proper frequency difference. This arrangement has the disadvantage that the useful tuning range of the condensers is reduced by the amount that the rotors are displaced. For this reason this method cannot be used where the i.f. is high.

The *second method* uses a gang condenser in which the oscillator condenser plates have a special shape. This method of oscillator and t-r-f circuit alignment is not suitable for use in all-wave receivers which cover several wave bands but is frequently used in receivers which cover only the normal broadcast frequency range.

The *third method* makes use of condensers of equal capacity for both the t-r-f and oscillator circuits. This method is used exclusively in all-wave receivers and in many receivers which cover only the normal broadcast frequency range. It is suitable for an i.f. of either 175 kc or 455 kc. The constant-frequency difference between the t-r-f and oscillator circuits is obtained through the use of a combination of shunt and series condensers in the oscillator circuit. The oscillator in superheterodyne receivers is generally tuned to a higher frequency than the t-r-f circuits, since a smaller percentage change in frequency is required and a smaller change in capacity is therefore necessary to produce the desired variation in the oscillator frequency. The oscillator tuning inductance is therefore smaller than that of the r-f circuits, and its value is such that the correct frequency difference between the oscillator and t.r.f is obtained at the middle of the tuning range with equal capacity in each circuit. The combination of shunt and series condensers used in the tuned oscillator circuit maintains the frequency difference constant throughout the tuning range of the receiver. A different combination of shunt and series condensers is used with each tuning range in all-wave receivers.

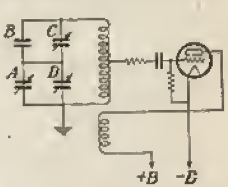


FIG. 10.—Typical superheterodyne oscillator circuit.

These condensers are shown in Fig. 10. Condenser A is the main tuning condenser. Condenser B is the fixed-series capacity. Condenser C is a small adjustable condenser for accurately adjusting the total series capacity. Condenser D is the small adjustable shunt condenser.

Typical values to maintain a frequency difference of 175 kc are as follows:

Main tuning capacity A	15–350 $\mu\mu$f
T-r-f tuning inductance	270 μh
Oscillator tuning inductance	215 μh
Fixed-series capacity B	750 $\mu\mu$f
Adjustable-series capacity C	15–70 $\mu\mu$f
Adjustable-shunt capacity D	5–40 $\mu\mu$f

The equations for calculating the circuit constants in a system of this type are given in Sec. 6.

Figure 10 shows a typical oscillator circuit used in superheterodyne receivers. It will be noted that the tube is connected across only a portion of the tuned circuit so as to minimize the effect of tube variations on the oscillator frequency.

17. Tone Controls. A considerable number of broadcast receivers are equipped with a h-f tone control, which is a device that enables the user of a receiver to vary the over-all fidelity characteristic of the receiver. The usual tone control operates on some portion of the a-f system in such a manner as to vary the h-f response. Figure 11 shows the most general method of accomplishing this result. In receivers employing resistance-coupled amplifier stages, the variable RC combination is shunted across the plate load resistor.

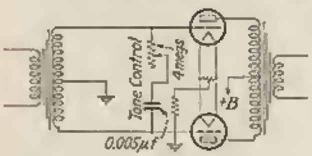


Fig. 11.—Tone-control circuit.

The advantages of a h-f tone control are as follows:

1. Noise encountered when receiving distant stations can be reduced considerably by decreasing the h-f response of a receiver through the use of a tone control.
2. All broadcast transmitters do not have the same fidelity characteristics and a tone control permits the user to compensate for some of these variations.
3. The frequency-response characteristic of the ear varies with the intensity of the sound. A tone control compensates for this characteristic.

A l-f tone control is used in some receivers so that l-f interference can be minimized. Such interference can be caused by a l-f hum on the

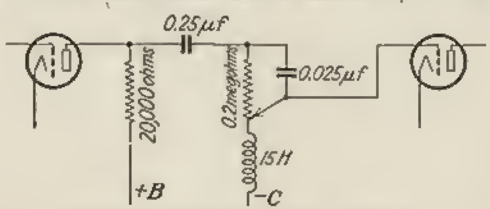


Fig. 12.—Low-frequency tone control.

carrier wave of a transmitter or by the beat note between two transmitters operating on the same channel. The intelligibility of the speech

reproduced by a broadcast receiver is frequently improved by decreasing the receiver's l-f response. Figure 12 shows a l-f tone control which has been used in broadcast receivers. A switch having two or more positions is sometimes used instead of the potentiometer.

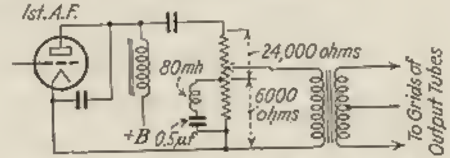


Fig. 13.—Tone-compensated volume control.

Acoustically Compensated Volume Control. A volume-control arrangement has been used in a number of broadcast receivers in which the over-all frequency-response characteristic of the receiver varies with the audio output level. This type of volume control has been called an *acoustically compensated volume control* and is intended to compensate for the variation in the frequency-response characteristic of the ear with amplitude.

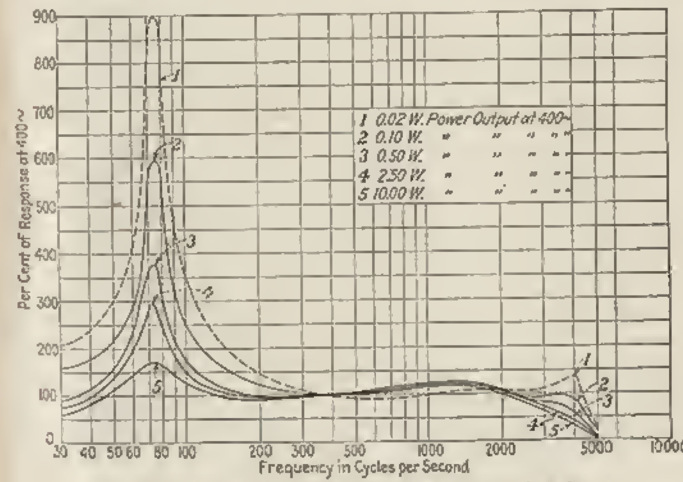


Fig. 14.—Variation of l-f and h-f response with volume.

Reducing the audio output of a receiver to a low value with a typical volume-control system gives the listener the impression that the very low and high frequencies have been attenuated and the middle frequency range has been correspondingly accentuated. The acoustically compensated volume control was devised to correct this effect. Figure 13 shows one of the arrangements which has been used to accomplish this result. This volume-control system makes use of a resonant circuit which attenuates the middle frequency range more than the high and low frequencies when the audio output is reduced. The effect of this type of control is illustrated by the curves in Fig. 14, which show the relation between the audio output and frequency-response characteristic of the receiver. The l-f compensation shown by these curves was used, not only to compensate for the variation in the frequency-response characteristic of the ear with amplitude, but also to correct for the acoustic deficiencies of the cabinet in which the receiver was installed. Since a definite relation should exist between the audio output level and the frequency-response characteristic of a receiver equipped with an acoustically compensated volume control, it is necessary that the audio output for a given setting of the volume control be substantially independent of the strength of the received signal. Some form of a.v.c. is necessary to meet this requirement.

18. Volume-control System. The two types of volume control which are used in broadcast receivers are manual and automatic.

The control of volume in both types is generally accomplished by varying the transconductance of the amplifier tubes through a change in the potential applied to the control grids. This method makes it possible to apply volume control to a number of tubes simultaneously using a single potentiometer or variable resistor. The source of the variable control grid potential does not need to supply power which is a prerequisite of any simple a-v-c system.

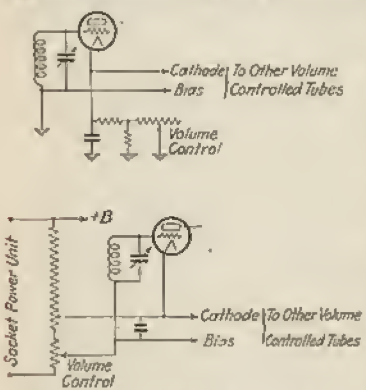


Fig. 15.—Volume-control circuits.

Serious distortion and cross modulation may be introduced through the use of this type of volume control if an amplifier tube is biased near the cutoff point and the applied signal potential is large. This distortion and cross modulation are functions of the third and higher derivatives of the E_g-I_p characteristic of the tube. To minimize this distortion, it is advisable to proportion the volume-control potential applied to the grid of the individual tubes inversely with the signal voltage on each tube. The use of remote cutoff amplifier tubes is desirable in a control-grid-bias volume-control system which must take care of a wide variation in the strength of received signals. Two arrangements which are frequently used to obtain manual volume control are illustrated by Fig. 15.

19. Automatic Volume Control. Automatic volume control is used almost universally in broadcast receivers. It has the advantage that practically the same audio output is obtained from the receiver irrespective of the input. This is an advantage in tuning from one station to another where a considerable difference exists in the relative field strength of the stations. It also has the advantage of compensating for some of the more serious effects of fading. Automatic volume control also makes the manual adjustment of volume less critical since the entire

range of the manual control is used only to vary the actual audio output. With the manual type of volume control, only a small fraction of the total variation of the control may be required to change the sound output from minimum to maximum. The manual type of control is therefore likely to be very critical to adjust.

Figure 16 shows a typical a-v-c- arrangement. In this system the d-c component of the rectified output of a detector is used as additional control grid bias for the r-f and i-f amplifier tubes. A single tube performs the dual function of providing the control grid bias and demodulat-

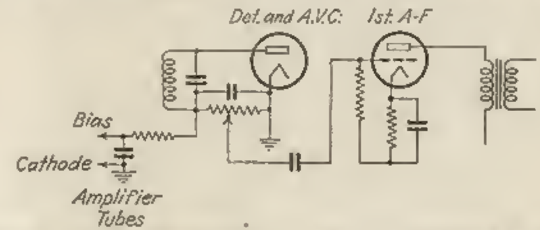


Fig. 16.—Combination detector—volume-control tube circuit.

ing the received signal. The output level is controlled by varying the audio amplification. For the receiver to reproduce faithfully the dynamic range of a received program, the rectifier from which the a-v-c control potential is derived must have a substantially linear input-output characteristic. A diode rectifier, with a load resistance of several hundred thousand ohms, provides a rectifier that is sufficiently linear.

Response and Recovery Characteristic. A resistance-capacity filter is usually used in the output circuit of an a-v-c rectifier. This filter prevents the a-f components in the output circuit of the rectifier from being applied to the amplifier grids. The time constant of the a-v-c rectifier output circuit should be such that the lowest modulation frequencies will not cause variations in the amplifier grid bias. It should not be so slow, however, as to give a noticeable delay when the system recovers from a crash of static. A time constant between $\frac{1}{10}$ and $\frac{1}{5}$ sec. is usually considered satisfactory.

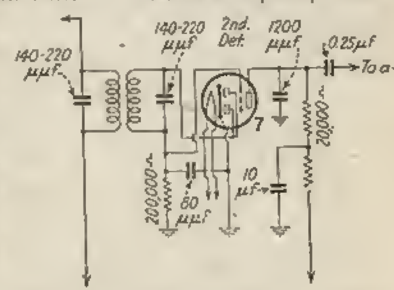


Fig. 17.—Avoiding detector distortion.[1]

20. Delayed Automatic Volume Control. The system illustrated by Fig. 18 is an example of delayed a.v.c. in which no control potential is derived until the signal level at the a-v-c rectifier has reached a predetermined value. The control grid of the double-diode triode is directly connected to the diode output resistor so that its bias becomes more negative with an increase in the amplitude of the signal applied to the diode.

[1] For description of this circuit see Art. 48, p. 455.

When no signal is applied to the diode, the control grid is at cathode potential and a d-c drop of between 50 and 100 volts occurs across the cathode resistor. The diode anode A is connected through a suitable resistor to the plate-supply system at a point sufficiently negative with respect to the cathode to give the desired delay. When a signal is applied to the signal diode, the control grid becomes negative and the drop across the cathode resistor decreases. When the amplitude of the received signal exceeds the predetermined level, the cathode of the tube becomes negative with respect to the anode A, and current flows through resistor R causing an increase in the negative bias on the amplifier grids.

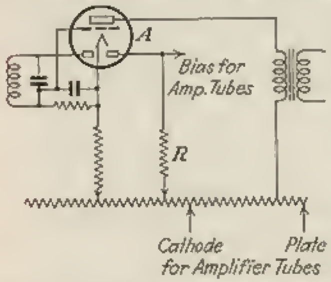


FIG. 18.—Double-diode triode as a-v-c. tube.

21. Selectivity Ahead of A-v-c System. In some receivers employing a separate a-v-c rectifier, this rectifier is connected to a point in the receiver which is preceded by less selectivity than is used ahead of the audio detector. The advantage of this system is that, when the receiver is tuned off resonance with a desired signal, the noise which is normally encountered is reduced. Under this condition the a-v-c potential is proportionately greater than the signal potential at the audio detector, and the receiver sensitivity and audio output are less than would have been obtained if the same selectivity was used ahead of the a-v-c rectifier and audio detector. This difference in selectivity should not exceed 10:1, otherwise the reduction in sensitivity, when tuned off resonance from a strong signal, will be so great as to prevent the reception of a weak signal on the adjacent channel.

22. Biasing the Amplifier Tubes at Different Rates. To minimize the type of distortion frequently encountered in volume-control systems due to the curvature of the E_g-I_p characteristic, it is desirable to proportion the volume-control grid bias for each amplifier tube inversely as the signal potential applied to the tube. The method generally used for approximating this relation is to provide one or more taps on the a-v-c bias resistor. The r-f amplifier tube is connected to the resistor; hence the entire potential drop is applied to its grid. The i-f amplifier tubes are connected to the tap or taps on the resistor so that they receive one-half or less of the total a-v-c voltage.

23. Separate Channel or Parallel A-v-c Systems. In some receivers a separate i-f amplifier stage is used to feed the a-v-c diode. The use of the separate channel, which is usually designed to have higher gain than the normal signal channel, makes it possible to provide a delayed a.v.c. having a very flat characteristic. The use of the separate channel also makes it easy to provide less selectivity in the a-v-c channel than in the signal channel and still provide a high signal voltage at the a-v-c rectifier. Another expedient which can be used with the separate channel a-v-c system to give a very flat a-v-c characteristic is to apply a part of the a-v-c potential to the amplifier tube in the signal channel following the point at which the additional a-v-c amplifier tube is connected. Care must be exercised in determining the control potential to be applied to an

amplifier stage following the point in the normal signal channel from which the control potential is derived. If the control potential applied to such a stage is too great, the a-v-c system may be overcompensated and the receiver output may actually decrease as the strength of a received signal increases. Figure 19 illustrates an a-v-c system employing a separate amplifier stage. In this arrangement a portion of the

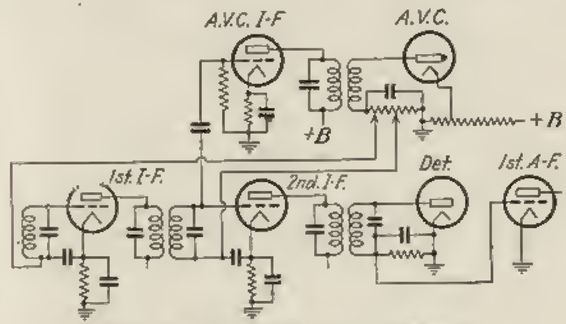


FIG. 19.—Amplified a-v-c arrangement.

control potential is applied to the signal amplifier tube subsequent to the point to which the separate a-v-c amplifier tube is connected.

24. Tuning Indicators. The majority of console radio receivers are provided with a tuning indicator which enables the user of the receiver to tune it accurately by eye to a desired station. The visual indication of resonance is usually obtained through the use of a 6U5 or similar electronic device in which the shape of the pattern on a fluorescent screen is controlled by the potential applied to one of the device's electrodes. The control potential in the majority of receivers is obtained from the a-v-c system. Receivers designed to pass a wide frequency band, such as high fidelity receivers, are usually provided with a special control circuit which is much more selective than the normal signal channel. This selective control circuit is tuned to the center of the i-f pass band, and the deflection of the pattern on the fluorescent screen of the 6U5 thus accurately indicates when the receiver is in resonance with a desired signal.

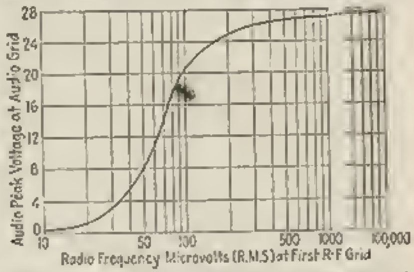


FIG. 20.—Automatic volume-control characteristic.

25. Noise in Receiving Systems. The source of the noise which is frequently obtained in the output circuit of a receiving system may either be located external to or within the receiving system. The two general sources of noise which are external to the receiver are as follows:

1. Atmospheric static.
2. Man-made static.

The expedients which are employed in receiving systems to minimize noise due to these types of interference without sacrificing the fidelity of the system are to employ an antenna system which will provide as favorable a signal-to-noise ratio as possible and to use sufficient shielding on the receiver chassis to prevent the noise being picked up by the receiver circuits.

The two chief sources of noise which are located within a receiving system are *thermal agitation* and *shot* effect.

Thermal-agitation noise is due to the random motion of the electrons within a conductor. The noise voltage introduced into a circuit by this cause may be calculated from the equation,

$$\bar{e}^2 = 5.49 \times 10^{-23} TZ\, df$$

where $\bar{e}^2$ = mean square thermal-agitation voltage
T = absolute temperature of the conductor (273 + °C.)
Z = resistance of the conductor or the resonant impedance of a tuned circuit
df = frequency band width factor.

The number of electrons emitted by the cathode of a thermionic tube varies from instant to instant, and this variation in emission introduces a voltage in the circuit through which these electrons pass. This variation in electron emission has been called *shot effect*.

The following equation gives the voltage introduced in a circuit by this cause:

$$\bar{E}^2 = 3.18 \times 10^{-19} IZ^2\, df$$

where $\bar{E}^2$ = mean-square shot-effect voltage (without charge)
I = electron current
Z = resonant impedance of the tuned circuit
df = frequency band width factor.

The space charge obtained in a vacuum tube under normal operating conditions reduces the shot-effect voltage to about one-half the above value.

The thermal-agitation and shot-effect noise found in the output circuit of a receiver usually originates in the grid and plate circuits, respectively, of the first tube. Where the gain in this tube is very low, the second tube may also contribute to the noise.

Since both types of noise are introduced as a series of pulses, the circuits in which the noise is introduced are excited at the frequency to which they are tuned.

The shot-effect voltage developed in the plate circuit of a tube varies in proportion to the square root of the plate current. Changing the plate load impedance has no direct effect on the signal-to-noise ratio since both factors are changed in the same ratio. High gain in the first tube with low plate current is therefore desirable to minimize shot-effect noise.

Thermal-agitation noise varies as the square root of the impedance across which the noise is developed. The merit of an antenna-input system from the thermal-agitation noise standpoint may be expressed as the ratio of $g/\sqrt{Z}$, where g is the voltage gain between the antenna

and the grid of the first tube and Z is the effective impedance in the grid circuit of this tube.

26. Complete Receiving System. The usual broadcast receiver consists of the following elements:

1. The receiver chassis.
2. The loud-speaker.
3. The cabinet.

In the majority of receivers the r-f, i-f, a-f, and power supply circuits are assembled as a single unit. In a few receivers the power supply rectifier and filter system and the power output tubes are mounted on a separate base.

The tuning condenser in a large number of broadcast receivers is flexibly mounted, with respect to the chassis, by means of soft rubber washers. The complete chassis in many receivers is also flexibly mounted in the cabinet. These precautions are used to prevent acoustic feedback in receivers which are capable of producing a high power output. Acoustic feedbacks are caused by the loud-speaker vibrations being transmitted through the cabinet to the receiver chassis and thence to tuning condenser or some other circuit element which is caused to vibrate sufficiently to intermittently detune the receiver at an a-f rate. If the proper phase relations exist between the loud-speaker vibrations and the variations in signal intensity which result from the vibration of the condenser plates, sustained oscillations may be produced.

27. Shielding and Filtering. It is common practice to confine the r-f and i-f circuits in metal containers which provide both electromagnetic and electrostatic shielding. Tube shields are used with "glass" type tubes to prevent coupling between tubes and between the grid and plate portions of individual tubes. When metal tubes are used, these shields are not required. In some instances shielded leads are used to provide the connections to the grids or plates of amplifier tubes, but in general the necessity for such shielding is avoided by so locating these leads that they are electrically isolated by the tube shields and the metal containers for the r-f and i-f circuits.

Care must be exercised in locating the power transformer and filter reactor on the receiver chassis, otherwise the electromagnetic field produced by these units may induce an appreciable hum voltage in the a-f circuits. It is desirable to keep these units separated from the a-f circuits as much as possible, and it is frequently necessary to determine experimentally the best location for these components by connecting them into the circuit with flexible leads and orienting them until a position is established which reduces the hum to the desired minimum.

Resistance-capacity filters are frequently used in the voltage supply leads for the tube electrodes. These filters are employed to prevent coupling between points in the system which differ in signal potential and to provide additional filtering for the voltage fluctuations which may exist at the output of the B supply filter. The d-c drop which can be tolerated in a given circuit is frequently a limiting factor in the use of such filters. When r-c filters are used in circuits in which the average current varies during the operation of the receiver, it is essential that the recovery characteristic of the filter be such that the voltage on the electrode can return to its normal value in approximately ⅙ sec., otherwise noticeable interruptions in the received program will be obtained when sudden changes in the average current occur. This problem is most frequently encountered when r-c filters are used in the plate or screen circuits of tubes which are controlled by the a-v-c system.

28. Loud-speaker. The electrodynamic loud-speaker is used in substantially all the broadcast receivers which are produced today.

29. Cabinet. The cabinet for a broadcast radio receiver must fulfill three requirements:

1. It must house and protect the receiver chassis and loud-speaker mechanism.
2. It must provide sufficient baffle area for the loud-speaker to give the desired l-f response.
3. It must serve as a piece of furniture which will harmonize with the furnishings in the room in which it is to be placed.

30. Single-dial Tuning Problem. One of the major problems in the design of a unicontrolled broadcast receiver is the maintenance of the proper alignment of the tuned circuits throughout the broadcast frequency range. To maintain such alignment normally requires that the inductances and variable condensers be made very uniform. It is common practice to sort the coils in groups so that the variation in inductance between them is less than 0.5 per cent. Coils are also employed which are wound in two sections, such as *a* and *b* in Fig. 21. One or more of the turns in section *a* can be moved with respect to section *b* so as to increase or decrease the spacing between the two coil sections. The total inductance of the coil can thereby be adjusted to any desired value. Adjustable iron cores are used in some receivers to give the coils the desired inductance values.

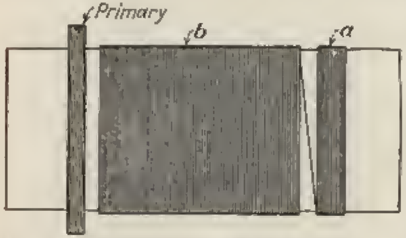


Fig. 21.—Method of varying inductance slightly for tracking purposes.

31. Push-button Tuning Controls. The majority of automobile radio receivers and a large number of home receivers are equipped with push-button tuning controls. Three general types of push-button tuning arrangements have been used.

The first type makes use of separate tuned circuits for each push button. These circuits are tuned to the desired station by variable condensers or variable iron-core inductances which are provided with screw-driver adjustments. The push buttons operate switches which select the groups of tuned circuits that have been pretuned to the desired station. The number of stations which can be tuned in on such a receiver is limited to the number of push buttons and groups of tuned circuits.

The second type of push-button tuning control makes use of a gang condenser so that the receiver can be tuned continuously through the entire frequency range covered by the receiver. The push-button tuning is accomplished by cams which rotate the gang condenser to the proper position to tune in the desired station. The cams are generally semicircular or U-shaped. When a push button is depressed, the top edge of one leg of the U is brought in contact with a rocking plate which is geared to the rotor of the condenser. As the push button is further depressed, the cam causes the condenser rotor and the plate to rotate until the other leg of the U also comes in contact with the plate. The U-shaped cams are held in a clamping arrangement which permits easy adjustment to any position which will cause a desired station to be tuned in. In some receivers a solenoid is used to supply the force which causes the cam to rotate the gang condenser to the desired position. In this case the only function of the push button is to select the cam and close the electrical circuit through the solenoid. This arrangement is likewise suitable for ganged iron-core tuning.

In a third type of push-button tuning control a gang condenser is also used, but the condenser is rotated by an electric motor. The push button makes contact with a slip ring attached to the condenser rotor shaft and closes the circuit through the motor. When a push button is depressed, the condenser rotates until an insulated segment on the slip ring opens the circuit. The angular displacement between the insulated segment on slip ring and the condenser rotor can be adjusted so that any desired station can be tuned in. This type of push-button tuning control is frequently used to provide remote tuning control. The remote-control unit with the required number of push buttons is connected to the receiver through a multiconductor cable.

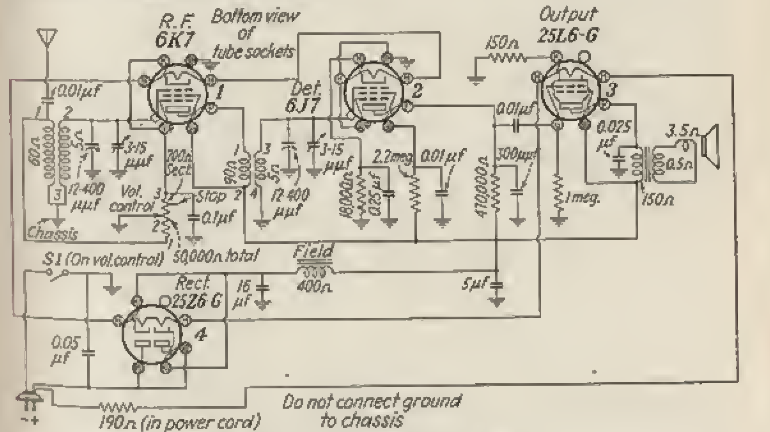


Fig. 22.—Typical t-r-f receiver.

32. Tuned-radio-frequency Receivers. Tuned r-f receivers are no longer produced except as the least expensive receivers of the midget type. Figure 22 shows the schematic diagram of a typical receiver.

33. Superheterodyne Receivers. The ease of obtaining high amplification and a high degree of selectivity with a minimum of shielding allows considerable flexibility in the design of a superheterodyne receiver. Sufficient amplification can be obtained in the r-f and i-f circuits so that a detector and single stage of a-f amplification are sufficient to provide the desired sensitivity. The general tendency in the design of superheterodyne receivers has been to take advantage of the high degree of selectivity which this type of receiver can provide at a corresponding sacrifice in fidelity. The superheterodyne receiver, however, lends itself just as well to the design of a high-fidelity receiver since the advantages of coupled tuned circuits can readily be realized in this type of receiver.

34. Superheterodyne Characteristics. The adjacent-channel selectivity and fidelity of a superheterodyne receiver can be determined readily from the characteristics of the individual components of the receiver.

Figure 23 shows the gain from the antenna to the grid of each tube. Figure 24 shows similar curves giving the total selectivity contributed by the tuned circuits between the antenna and the grid of each tube. To obtain the curves in Fig. 24, the selectivity curves of the individual circuits are plotted to the same scale on logarithmic coordinates. The over-all selectivity characteristic curves are then obtained by laying off for each frequency a distance which is equal to the sum of the distances which represent the ordinates of the individual selectivity characteristics for the same frequency. From these two sets of curves (Figs. 23 and 24) it is possible to determine the voltage on the grid of each tube from a local station when the receiver is tuned to a distant station on an adjacent channel. Such a determination is frequently desirable in this type of receiver where the selectivity contributed by the circuits between each tube is not uniform. This relation between gain and selectivity between each tube must be properly proportioned; otherwise, the signal from a local station may be sufficient to draw grid current on one of the tubes even if the over-all selectivity of the receiver is sufficient to separate the signals from the local and distant stations before they reach the second detector.

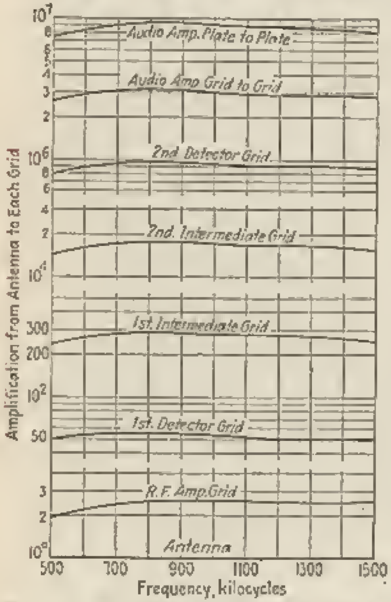


FIG. 23.—Voltage gain in superheterodyne receiver.

35. Superheterodyne Interference Problems. The selectivity of a superheterodyne receiver as determined in Fig. 24 is not a true indication of the actual selectivity of the receiver under all conditions, as this type of receiver is susceptible to certain types of interference which are not encountered with a t-r-f receiver. The susceptibility of these interferences is a result of converting the received signal to an i.f. The following classification gives the more important possible sources of interference common to a superheterodyne receiver in which the i.f. is lower than any frequency in the tuning range of the receiver.

1. *Image-frequency interference:* If f is the oscillator frequency in a superheterodyne and IF the i.f., signals impressed on the first detector, having frequencies of either $f + IF$ or $f - IF$, will be heterodyned to the i.f. and pass through the receiver. It is therefore necessary to prevent one of these signals from reaching the first detector; otherwise, *image-frequency* interference will result. Radio-frequency circuits, tuned to the signal which it is desired to receive, are the usual arrangement for preventing image-frequency interference. Since the oscillator in superheterodyne receivers is usually tuned to a higher frequency than the r-f circuits, a signal which can produce image-frequency interference must have a frequency of $f_1 + 2IF$, where f_1 is the frequency of the desired station.

When a received signal is successively heterodyned to two intermediate frequencies, as is the case in some superheterodyne receivers used in communication work, there is more than one signal that can cause image-frequency interference with any desired signal. For example, if f_1 is the frequency of the desired signal and IF_1 and IF_2 the two intermediate frequencies, then interference can be caused by signals whose frequencies are $f_1 + 2IF_1$ and $f_1 - 2IF_2$. It is assumed that both oscillators are tuned to a higher frequency than the signal frequency. The circuits ahead of the second heterodyne oscillator and associated detector must provide the selectivity necessary to avoid interference by the $f_1 - 2IF_2$ signals.

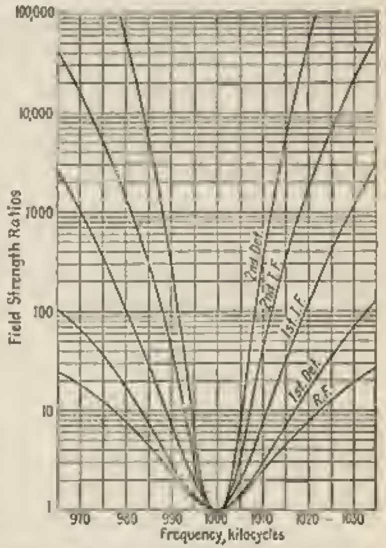


FIG. 24.—Superheterodyne selectivity characteristics.

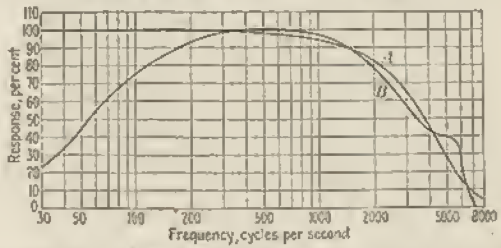


FIG. 25.—*A*, side-band attenuation due to r-f circuits of superheterodyne; *B*, over-all fidelity characteristic.

2. *Interference due to harmonics of the oscillator heterodyning undesired stations:* If a signal having a frequency of $2f \pm IF$ is impressed on the first detector, it will cause interference with the signal being heterodyned by the fundamental oscillator frequency f. Tuned r-f circuits ahead of the first detector reduce the possibility of this type of interference.

3. *Interference due to stations which are separated by the i.f.:* Combinations of signals are sometimes encountered which are separated by the intermediate frequency, and, if such signals are permitted to reach the first detector, interference will result. Tuned r-f circuits ahead of the first detector are also used to prevent this type of interference.

4. *Interference due to harmonics of the i.f. produced by the second detector:* When the i.f. is lower than any frequency in the tuning range of the receiver, certain harmonics of the i.f. fall in the broadcast frequency band. If these harmonics, which are produced by the second detector, are of sufficient amplitude and are fed back to the input system of the receiver, they will cause interference when a station is received whose frequency is equal to a particular harmonic of the i.f. With an i.f. of 175 kc this type of interference is likely to be encountered at 700, 875, 1,050, 1,225, and 1,400 kc. This type of interference is eliminated by careful shielding of the second-detector circuits.

5. *Responses when the difference frequency is less than the i.f.:* When the frequency difference between the oscillator and the signal impressed on the first detector is one-half or one-third the i.f., a second or third harmonic of the beat frequency may be produced in the first detector which will be amplified by the i-f amplifier. Interference with a desired signal may be produced in this way. If sufficient selectivity is used ahead of the first detector to prevent image-frequency interference, interference of this type will also be avoided.

36. Sources of Interference When the I.F. Is Higher Than the Signal Frequency. In some all-wave receivers the i.f. is higher than the signal frequency throughout one tuning range. When this condition exists, the potential sources of interference differ from those enumerated above. Interference may result from the following causes:

1. *Interference due to harmonics of the received signal:* If the tuning range includes a signal frequency equal to one-half or one-third the r.f., such a signal may produce harmonics in the first detector which will be amplified by the i-f amplifier. Intermediate-frequency signals are thus produced without the use of the heterodyne oscillator. The frequency of the signals produced in this way does not vary as the receiver is tuned. The local oscillator also heterodynes the signal to the i.f., but the i.f. thus produced varies as the receiver is tuned. When the receiver is tuned through such a signal, a beat note is produced by the two i-f signals. Selectivity ahead of the first detector will restrict the tuning range over which this interference is encountered but cannot eliminate it when the desired signal is the signal causing the interference.

2. *Interference due to two signals whose sum frequency equals the i.f.:* When two signals are impressed on the first detector and produce a sum frequency equal to the i.f., a beat note is produced as the receiver is tuned through a desired signal. Under this condition two i-f signals are produced, one of which remains fixed in frequency while the other varies as the receiver tuning is changed. Since the signals which can produce this interference may be on adjacent channels, the selectivity which must be used ahead of the first detector to avoid entirely this interference is equivalent to that normally used in the complete receiver.

37. Choice of the I.F. The choice of the intermediate frequency for a superheterodyne receiver is a compromise between the following factors:

1. With a given t-r-f system ahead of the first detector the possibility of encountering image-frequency interference is reduced as the i.f. is increased.

2. Under the above conditions the possibility of interference due to two stations separated by the i.f. is also reduced as the i.f. is raised.

3. The possibility of interference due to harmonics of the i.f. being fed back from the second detector to the input of the receiver increases as the i.f. is raised, since lower harmonics appear in the broadcast band and the amplitude of the harmonics which can cause interference is therefore increased.

4. The difficulty of obtaining a high degree of selectivity and amplification in an i-f amplifier is increased as the i.f. is raised.

The majority of broadcast receivers employ intermediate frequencies at or near either 175 or 455 kc. The higher i.f. is used in all-wave receivers to minimize image-frequency interference and reduce reaction between the oscillator and first detector circuits when the receiver is tuned to high signal frequencies. With an i.f. of 175 kc the fourth harmonic is the first to appear in the broadcast range from 550 to 1,600 kc. The second and third harmonics of a 455-kc i.f. appear in this tuning range.

38. Tuned-radio-frequency Circuits. The t-r-f circuits ahead of the first detector in a superheterodyne receiver are used primarily for

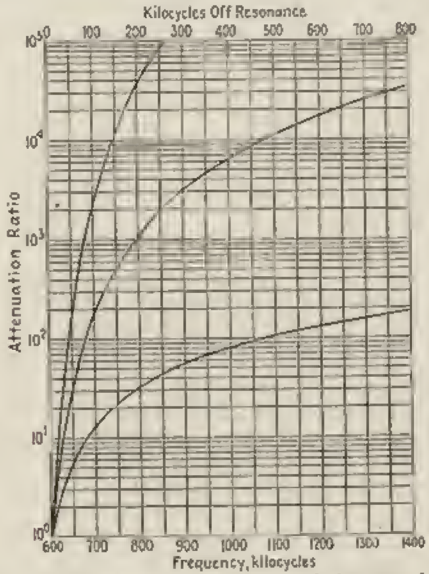


FIG. 26.—Attenuation of one, two, and three t-r-f circuits.

eliminating certain types of interference common to the superheterodyne type of receiver. Figure 26 shows the attenuation of one, two, and three t-r-f circuits for frequencies up to 800 kc off resonance when tuned to 600 kc. From curves of this type it is possible to obtain the image-frequency ratio for any given r-f system which may be used ahead of the first detector. *Image-frequency ratio* has been termed the ratio between the field strength necessary to produce standard output from a superheterodyne at the image frequency and that necessary to produce standard output at the frequency to which the receiver is tuned. The image-frequency ratio provided by modern broadcast receivers is usually about 20,000:1 in the tuning range from 540 to 1,600 kc. With an image frequency of

460 kc this ratio can be obtained with two tuned r-f circuits. This combination provides an image-frequency ratio of between 100:1 and 200:1 in the tuning range from 10 to 20 Mc. Care must be exercised in the design of a superheterodyne receiver to use sufficient shielding so that the actual selectivity of the t-r-f circuits is realized. If a reasonable amount of shielding is not used, signals which will cause image-frequency interference may be picked up directly on the first detector circuits, and the benefit of the t-r-f circuits between the antenna and this detector will be lost.

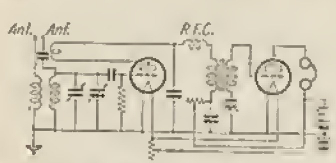


FIG. 27.—Regenerative circuit with resistance control.

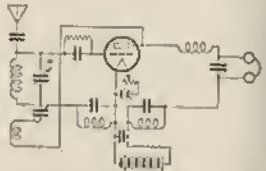


FIG. 28.—Single-tube superregenerator.

39. Regenerative Receivers. A typical regenerative-receiver circuit is shown in Fig. 27. In this arrangement a variable resistance is used to vary the plate potential on the tube and thereby control the regeneration. The coupling between the tickler coil and the inductance of the tuned circuit is fixed. This arrangement is generally used in receivers which make use of plug-in coils to cover a wide frequency range since the tickler coil can then be wound on the same form as the tuned circuit inductance.

40. Superregenerative Receivers. Figure 28 shows the circuit diagram of a single-tube superregenerative receiver in which the quenching

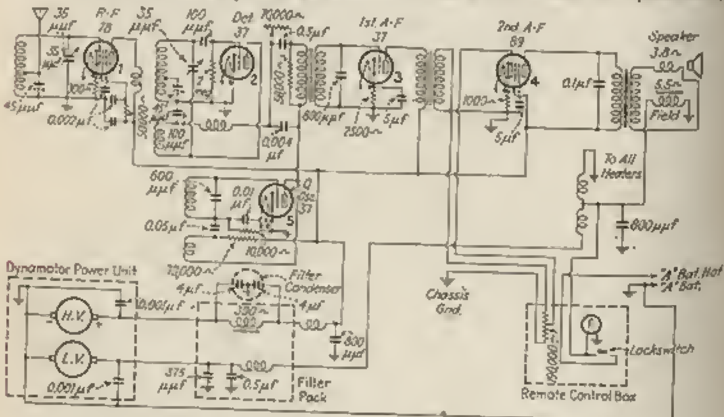


FIG. 29.—Superregenerator receiver for police cars.

frequency is produced by the same tube which provides the superregeneration. The quenching frequency is usually between 5,000 and 20,000

cycles. A filter is generally used in the output circuit of the superregenerative tube to eliminate the quenching frequency so that it does not appear in the receiver output.

In Fig. 29 is shown the complete circuit diagram of a superregenerative receiver used in police cars for the reception of signals on frequencies between 30 and 40 Mc. This receiver employs a tuned r-f stage ahead of the superregenerative detector to prevent radiation. The 20-kc quench frequency is provided by a separate oscillator. Two stages of a-f amplification are employed. Amateur practice on 56 Mc is to use a single tube in which the periodic blocking of the tube is produced by the proper choice of grid leak and condenser.

41. All-wave Receivers. A large number of the broadcast receivers being produced at the present time cover one or more short-wave ranges in addition to the normal broadcast frequency band (540 to 1,600 kc). These short-wave ranges include frequencies up to 48,000 kc.

In the majority of all-wave receivers separate coils are employed in the r-f system for each tuning range. A few receivers use a tapped coil for each tuned circuit. When such coils are utilized, the unused portion of the coil is always short-circuited. When separate coils are employed, the coils for two or more of the frequency bands are frequently wound on a single form. The coil windings differ considerably with the frequency range which the coils are designed to cover. Wire as small as No. 35 Brown and Sharpe is used in the inductances for the tuning range from 540 to 1,600 kc, while wire as large as No. 22 Brown and Sharpe is used in some of the short-wave coils. The turns on the short-wave coils are usually spaced to minimize the coil losses.

All-wave receivers are provided with a gang switch for simultaneously connecting the coils used for each tuning range to the associated tuning condensers and tubes. Such a switching arrangement is illustrated by Fig. 30, which shows the complete circuit diagram of a typical all-wave receiver.

Receivers of this type are usually equipped with tuning mechanisms which permit the user to change the drive ratio between the tuning knob and the variable condenser from 10:1 to 50:1. The 50:1 ratio is necessary to tune the receiver accurately to a short-wave station since the frequency band covered in a single h-f tuning range may be over ten times that covered in the range from 540 to 1,600 kc.

Special tuning dials are necessary on all-wave receivers since a separate scale is required for each tuning range. In some receivers all the scales are visible to the user regardless of the tuning range which is being used, and an indicator which is actuated by the range switch knob is used to designate the correct scale. In the dials used on other receivers of this type only the scale corresponding to the tuning range being used is visible. With this arrangement the dial scales are movable with respect to the dial opening, and the range switch is mechanically connected with the dial scales so that, as the tuning range is switched from one frequency band to another, the proper scale is moved into place.

42. Automobile Radio Receivers. Compactness and ruggedness are two of the essential requirements of an automobile radio receiver. Compactness is required because of the small space which is usually available in which to mount the receiver, and ruggedness is necessary because of the vibration and road shocks to which the receiver is subjected.

Since the strength of the signals intercepted by an automobile antenna varies greatly with the location of the car, it is essential that an automobile radio receiver be equipped with an effective a-v-c system.

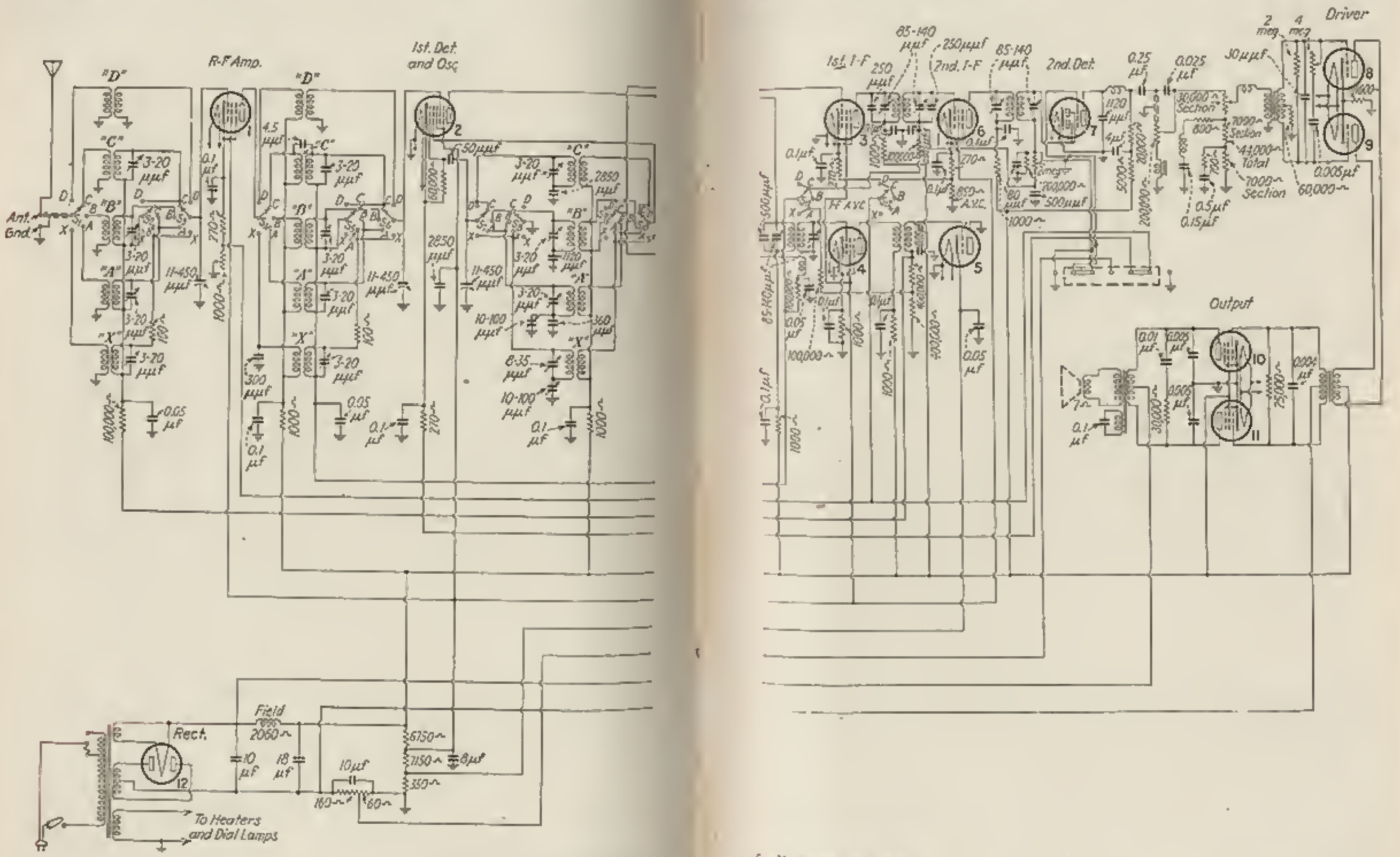


Fig. 30.—Circuit diagram of all-wave receiver.

The plate and bias potentials for the tubes in automobile receivers are generally obtained from vibrator B-supply systems. Two arrange-

ments which are frequently used are illustrated by Fig. 31. In the system shown by diagram (a), a vibrator and transformer are used to derive a high-voltage alternating potential from the 6-volt storage battery. This voltage is rectified by a vacuum-tube rectifier and supplied to the receiver circuits through a conventional filter. In diagram (b) the vibrator is not only used to provide the high voltage but also performs the function of rectification.

The chassis of an automobile radio receiver is generally completely shielded to prevent the pickup of ignition interference on the receiver circuits. Two methods have been employed for preventing the ignition systems of automobiles from causing excessive interference in automobile radio receivers. In the first method the interference radiated by the ignition system is mimimized through the use of suppressor resistors in the spark plug and distributor leads. An r-f filter is used in the leads connecting the receiver to the storage battery. All portions of the automobile electrical system which may radiate the interference such

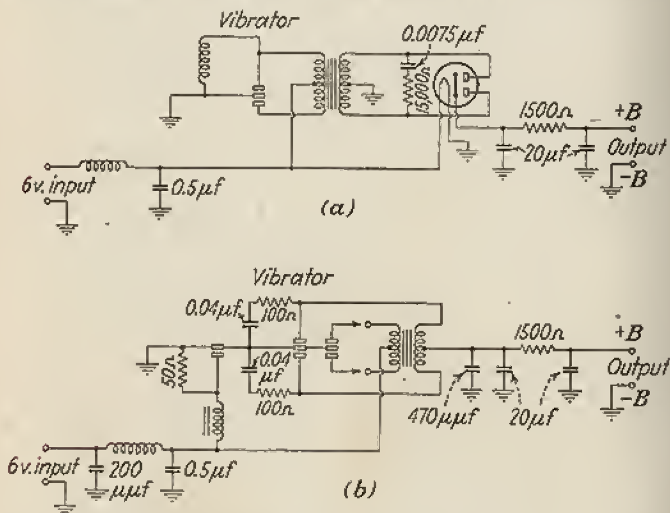


FIG. 31.—B-supply systems for automobile receivers.

as leads to the dome light, etc., are by-passed with a suitable by-pass condenser. The objection to this method is that the resistors which are used to suppress the h-f oscillations may decrease the effectiveness of the ignition system to the point where a loss in engine efficiency occurs. In the second method which has been used to minimize this type of interference, a special antenna filter system is employed which discriminates between h-f ignition interference and the desired signal.

In a large number of automobile radio receivers all the receiver elements are assembled in a single unit. This unit is generally designed for a specific line of automobiles and is arranged to mount behind an opening in the dash.

43. Radio Phonograph Combinations. Many radio receivers of the console type and a limited number of the table type are provided with a turntable and phonograph pickup. The a-f voltage developed by the pickup is usually applied to the grid of the first a-f amplifier tube in the

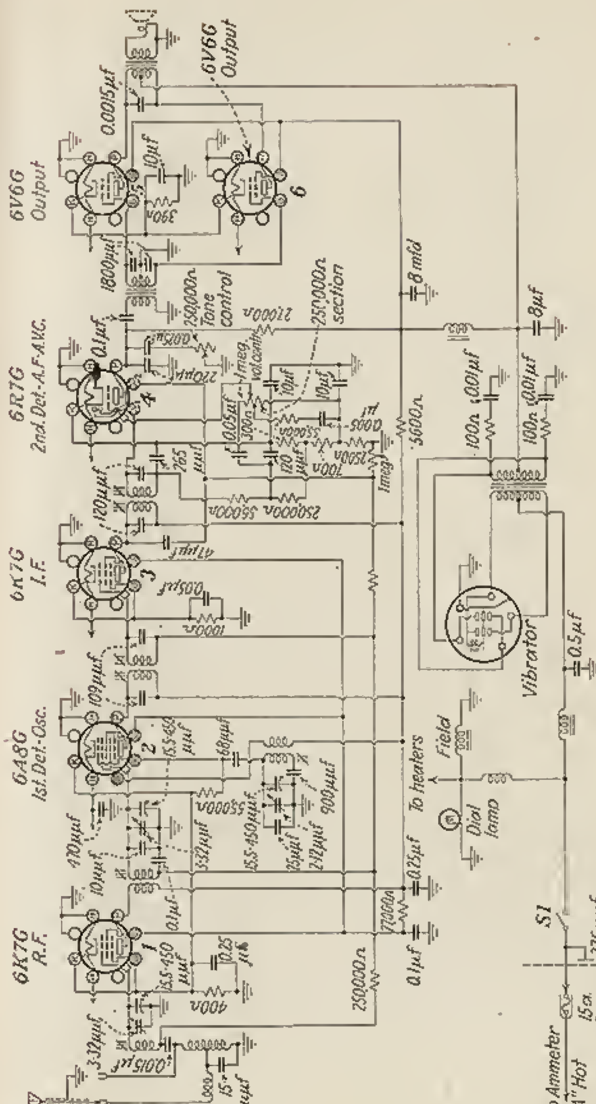


FIG. 32.—Typical automobile radio receiver.

receiver. Many standard radio receivers are provided with terminals so that connections can be made to a phonograph pickup.

44. Universal or A-c-d-c Receivers. Figure 33 shows the circuit diagram of a five-tube superheterodyne receiver which may be operated from either a.c. or d.c. The heaters of the tubes used in receivers of this type are connected in series. Since the voltage required for the heaters is considerably less than the line voltage, the heaters are connected to the supply line through a series resistor which provides the desired voltage drop. This resistor must usually dissipate considerable heat. In a number of the smaller universal receivers the series heater resistor is included as a third conductor in the power cord thus facilitating the

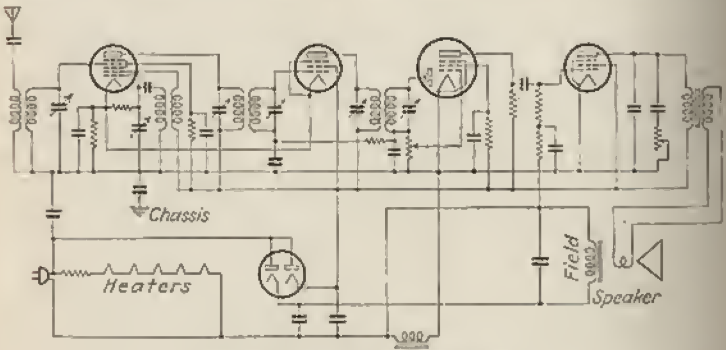


FIG. 33.—A-c-d-c receiver circuit.

dissipation of heat. The rectifier tube prevents the electrolytic condensers from being damaged in case the power plug is not inserted correctly in a d-c outlet.

HIGH-FIDELITY RECEIVERS

45. Audio-frequency Response Range. The term "high fidelity" has been associated with broadcast receivers which provide reasonably uniform reproduction of frequencies from 50 to 8,000 cycles. It is impractical to exceed the 8,000-cycle h-f limit as long as the 10-kc spacing between broadcasting stations is maintained. When the range of high frequencies reproduced by a broadcast receiver is extended, a corresponding increase in l-f response range must be made to maintain a proper acoustic balance.

46. Variable Selectivity. To obtain reasonable freedom from cross talk and h-f interference when receiving weak signals, it is necessary that a high-fidelity receiver be provided with some means whereby its selectivity can be increased over that required for the reception of high-fidelity programs from local transmitters. This change in selectivity is generally secured by altering the effective coupling between the primary and secondary of one or more of the i-f transformers. This is accomplished by moving one i-f coil with respect to the other, by the use of a transformer employing a third winding which is shunted by a variable resistor, or by varying the coupling between two windings of a variometer which

contains a portion of the primary and secondary inductance of the i-f transformer.

47. "Monkey Chatter" Interference. Receivers reproducing high audio frequencies are susceptible to "monkey chatter" interference. This interference is produced either by the side bands of an undesired signal beating with the desired carrier wave or by the side bands of the desired signal beating with an undesired carrier wave. A large percentage of the energy in this interference is found in the frequency band from 8,000 to 10,000 cycles. A low-pass filter giving an attenuation of at least 40 db for frequencies above 8,000 cycles is generally employed in high-fidelity receivers to minimize this interference. Such a filter is also effective in reducing other h-f interference such as man-made static, tube hiss, and beat notes caused by carrier waves on adjacent channels.

48. Minimizing Distortion in High-fidelity Receivers. As the range of frequencies reproduced by a broadcast receiver is increased, the distortion which can be tolerated is reduced. The most important source of distortion in such receivers is the detector and a-f amplifier. The diode is the most satisfactory of the detectors used in broadcast receivers from the standpoint of distortion. To minimize the distortion when receiving signals having a high percentage modulation with a diode detector, it is necessary that the a-f impedance of the diode output circuit be the same as its d-c resistance. This condition may be obtained through the circuit shown in Fig. 17, in which the grid of the a-f amplifier tube is directly connected to the diode output resistor thus avoiding the shunt a-c path formed by the conventional coupling condenser and leak. The distortion introduced by this detector and a-f amplifier stage can be kept well under 5 per cent when the percentage modulation of the received signal is approximately 100 per cent. A push-pull class A output stage is generally used where minimum distortion is desired.

The precautions for minimizing distortion as outlined under Automatic Volume Control (Art. 19) must also be observed.

49. High-fidelity Loud-speakers. Special loud-speakers are employed in high-fidelity receivers to reproduce the wide range of frequencies. In some receivers a special cone speaker is used employing a voice coil wound with aluminum wire. In other receivers two loud-speakers are utilized to reproduce the desired frequency range. One of these speakers is designed to reproduce frequencies from 50 to 2,500 cycles and the other, frequencies from 2,500 to 8,000 cycles. The h-f energy radiated by the usual cone loud-speaker is concentrated in a relatively narrow beam. Sound diffusers are used in a number of high-fidelity receivers to disperse this beam and produce a more uniform distribution of h-f energy. These devices consist of a number of vertical or horizontal slats placed in front of the loud-speaker at various angles with respect to its axis.

50. Frequency Modulation Receivers. A frequency modulation receiver utilizes the essential elements of an amplitude modulation receiver and in addition requires some means for converting the variation in frequency into a change in amplitude. The balanced detector shown in Fig. 34 is one arrangement used for this purpose. This detector system minimizes the effect of amplitude variations in the received signal. A voltage limiter is frequently used ahead of the balanced detector to further minimize amplitude variations.

51. Commercial Receivers. The principles underlying the design of commercial receivers are the same as those employed in the design of broadcast receivers.

Ruggedness and reliability are among the chief considerations in the design of commercial receivers, since such receivers must usually remain in continuous operation for long periods of time. Simplicity of tuning is not so important in this type of receiver as in broadcast radio receivers, since commercial receivers are generally used by skilled operators. Commercial radio receivers are generally designed to use battery-operated tubes. The plate potential for such receivers is supplied by either batteries or a motor generator. In some transoceanic receiving systems, three complete receiver and antenna combinations are used to overcome the effects of fading. In an installation of this type the antennas are separated by several wave lengths. An automatic-volume-control arrangement is provided so that only the output of the receiver which is receiving the strongest signal is used.

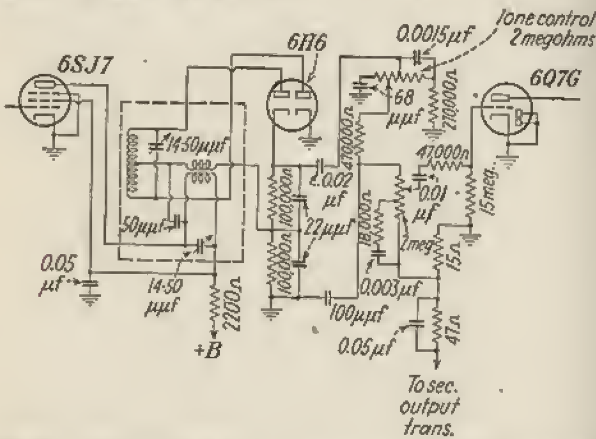


FIG. 34.—Balanced detector for frequency modulation receiver.

52. Direction Finders. The directional property of a loop antenna is utilized in direction finders to determine the plane in which the radio transmitter and the direction finder are located. The circuit diagram of a typical finder is shown in Fig. 35. The loop antenna in this receiver is enclosed in an electrostatic shield. The center tap on the loop is grounded. These precautions are taken to eliminate the electrostatic effect of the loop antenna. If this effect is present, a broad minimum is obtained as the loop antenna is rotated and it is impossible to obtain an accurate bearing. The diagram shows an arrangement for compensating for the effect of a near-by metal object which might distort the field around the loop. A small antenna is erected and connected through a resistor to the variometer shown in the diagram. By proper adjustment of the variometer the signals introduced by the near-by metal object and the compensating antenna and variometer arrangement are made to balance so that they produce no effect on the inherent directional properties of the loop antenna. The superheterodyne circuit is usually employed in direction finders. Both the loop antenna and oscillator

circuits are tuned through the use of a single control. Bearings can be determined to within about 1 deg.

53. Single-signal Receivers. Many of the receivers used by amateur radio operators are of the single-signal type which is characterized by its extreme selectivity. The high degree of selectivity is frequently obtained through the use of a quartz crystal as a coupling element in one of the i-f stages. The selectivity characteristic of a 460-kc quartz crystal may have band widths at 90 per cent and 10 per cent of 10 and 100 cycles, respectively. The limited frequency band required for code communication permits the use of receivers having such a selectivity characteristic.

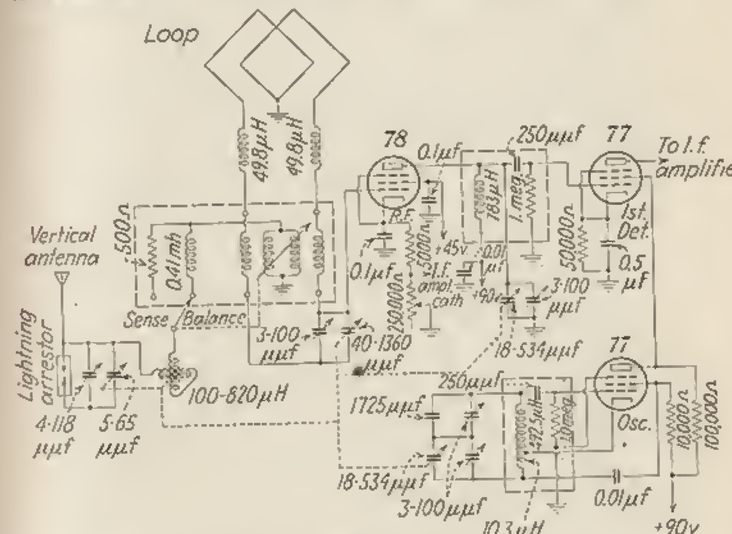


FIG. 35.—Radio direction finder.

Figure 36 shows the circuit diagram of a receiver employing a quartz crystal. As indicated by this diagram the crystal is used as a coupling element between the secondary of the first i-f transformer and the grid of the first i-f amplifier tube. In a number of receivers a parallel-resonant i-f circuit is also used on the grid side of the crystal filter. A neutralizing arrangement is employed to counteract the effect of the crystal holder capacity. This capacity limits the selectivity contributed by the crystal, and in conjunction with the inductance of the crystal, forms a parallel-resonant circuit which introduces considerable attenuation for a narrow band of frequencies near the frequency to which the crystal is resonant. A switch is provided for removing the crystal from the circuit when desired, thereby decreasing the receiver selectivity. A switch is also employed for rendering the a-v-c system inoperative when code signals are received. The receiver gain is then adjusted by means of a manual control. An i-f oscillator is used to heterodyne c-w signals.

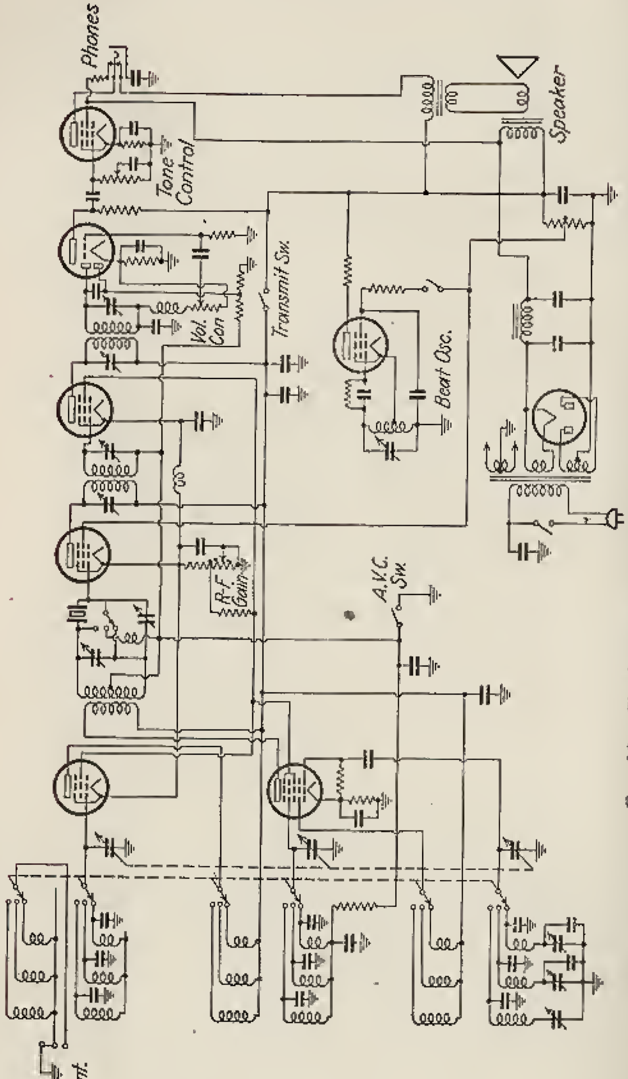


Fig. 36.—Single-signal receiver using quartz crystal.

SECTION 14

POWER SUPPLY SYSTEMS

By R. C. Hitchcock, Ed. D.[1]

1. Direct-current Power Requirement. The electrical power required for operating radio transmitters and receivers is usually "steady" d.c. for plate and grid circuits. Depending on conditions, either d.c. or a.c. is employed for heating tube filaments or cathodes. Figure 1 shows a variety of means that can be employed, by using suitable conversion apparatus, to deliver the desired d.c.

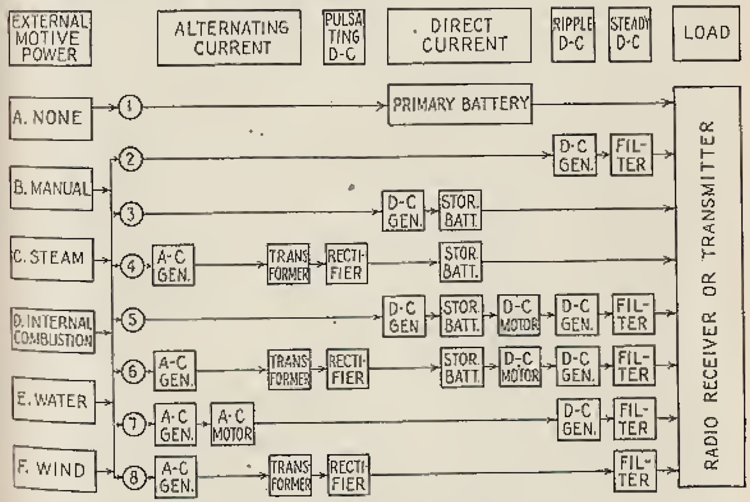


Fig. 1.—Types of power systems.

Five types of motive power, B to F, in the left column of Fig. 1 can be used with any one of seven of the eight numbered rows showing how d.c. is secured. The conditions under which each of these types of motive power is used will vary with the type of service which is desired.

2. Type of Service. The type of system required depends largely on the amount of power to be furnished. A portable receiver may operate for some time from self-contained dry-cell batteries, but a

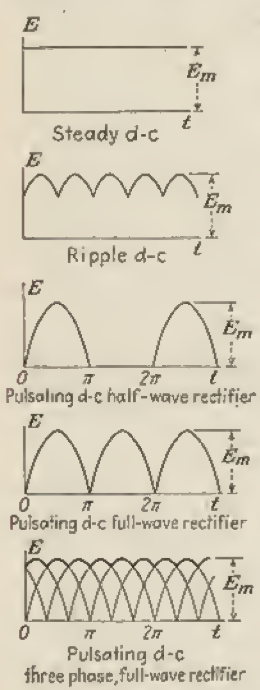
[1] Engineering Department, Westinghouse Electric and Manufacturing Company, Newark, N. J.

50-kw. broadcast transmitter on a regular schedule requires considerably more energy. An explorer can operate a hand-cranked generator of 50 or even 100 watts capacity for a short time, but for longer periods other devices are more applicable.

Transmitter power supplies are of two types. One is for regular use, and the other is the "emergency set" or "stand-by." The latter is ready on a few seconds notice, and is capable of supplying sufficient energy for regular operation. For multikilowatt stations preferably two independent sources of a-c supply are provided, on either of which the station can develop full rated power. A gasoline-electric set may serve the purpose, being independent of long wire lines.

CHARACTERISTICS AND MEASUREMENT OF D.C.

3. Indicating Instruments. Since d.c. is largely employed for radio transmitters and receivers, a brief analysis will be made of the various



	Minimum	Average	RMS (effective)	Form factor = RMS/average
Steady d-c	E_m	E_m	E_m	1.00
Ripple d-c	$0.65E_m$	$0.88E_m$	$0.89E_m$	1.01
Pulsating d-c half-wave rectifier	0	$0.32E_m$	$0.5E_m$	1.57
Pulsating d-c full-wave rectifier	0	$0.64E_m$	$0.71E_m$	1.11
Pulsating d-c three phase, full-wave rectifier	$0.87E_m$	$0.96E_m$	$0.96E_m$	1.00

FIG. 2.—Types and characteristics of d.c.

kinds of d.c. and their measurement. One reason for this analysis is that instruments of the repulsion-iron or dynamometer type will not

read the same as an "average" type on certain types of d.c. This difference in readings sometimes causes confusion.

If a d-c ammeter is specified, it usually refers to a d'Arsonval instrument (permanent magnet moving coil); one which reads "average" values.

Figure 2 shows five typical kinds of d.c., one or more of which are always present in any d-c power supply. Steady d.c. is the output from a primary or secondary battery, or from a suitable filter connected to a pulsating or ripple d-c source. Ripple d.c. is the usual output from a d-c generator, the ripple being caused by commutation. Three types of pulsating d.c. are as follows: (*a*) half-wave rectified single phase; (*b*) full-wave rectified single phase; and (*c*) three-phase full-wave rectified.

The ambiguity of the term *direct current* is readily apparent when considering Fig. 2 since all these wave forms fall into this classification. The figure shows the minimum voltage as a decimal part of the maximum voltage E_m; for example, the ripple d.c. shown has a minimum which is $0.65\ E_m$ or 65 per cent of its maximum.

The second column shows the average value of potential drop as a factor times the maximum E_m. These factors for pulsating d.c. vary from 0.32 for the half-wave rectified to 0.96 in the case of the full-wave three-phase rectified.

The r.m.s. or effective value of a current is such that the heating effect (I^2R) is the same for d.c. or a.c. For pulsating d.c. the watt reading found by the average voltage times the average current is not the same as the r-m-s voltage times the r-m-s current. The results of these average readings are sometimes called[1] *d-c watts.*

The readings of different types of instruments can be predicted from the value of the form factor. On the ripple d.c. from a d-c generator, when the form factor is 1.01 as shown, a d'Arsonval instrument would read 1 per cent lower than a dynamometer type of instrument. The d'Arsonval instrument reads average, and the dynamometer reads r.m.s. When there is a difference in readings, the r.m.s. instrument always reads higher. For the pulsating d-c output of the single-phase half-wave rectifier the form factor is 1.57, and an r.m.s. instrument (repulsion-iron, dynamometer, or thermocouple) would read 57 per cent higher than an average instrument (d'Arsonval).

TYPICAL POWER SUPPLY SYSTEMS

4. Television Receiver Power Supply. Two rectifier-filter systems characterize television receivers, as shown in Fig. 3.[2]

The kinescope second anode requires 7,500 volts d.c., which is supplied by a 2V3-G half-wave rectifier tube and a π filter comprising two 0.03-μf capacitors and a resistor of 470,000 ohms. As will be noted later, when considering filters, low-current circuits are adequately smoothed by this simple circuit.

The d.c. for the other parts of the receiver is supplied by a 5T4 full-wave rectifier, with a choke-capacitor double π filter, having 40, 80, and 10 μf, respectively, connected between the two choke coils.

5. Receivers Using Either Batteries or Utility Power. Figure 5 shows a combination battery- and socket-power receiver which has no relays,

[1] SMITH, L. R., Rectox Rectifier Testing, *Elec. Jour.*, August, 1938, p. 328.
[2] RCA Mfg. Co., Camden, N. J. Model TRK-12.

switches, or complicated change-over parts.[1] All battery connections are completed when the light socket plug is inserted in the chassis socket. When the plug is removed, the batteries are isolated so that the plug can then be inserted into a power-supply socket of 105 to 125 volts, a.c. or d.c. With this latter arrangement the set has high output power (117L7GT tube) while it has normal battery output (3Q5GT) on the battery connection.

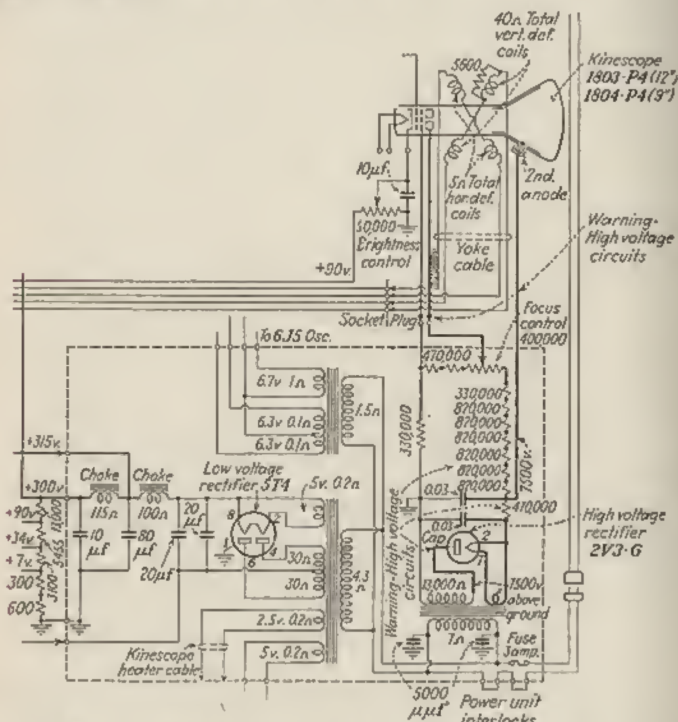


FIG. 3.—Typical television receiver power supply system.

On utility power the plug ZY is inserted in an a-c or d-c outlet of 115 volts, correctly poled if d.c. is used. When the line switch is turned on, the 117L7GT filament is heated across the line, and the rectifier section of this tube supplies half-wave energy to the filter choke. The other filaments in series are heated by the plate current (d.c.) of the output pentode section of the 117L7GT, which also provides bias potential, and this connection also provides bias potential for the control grid of this tube. With this connection the battery output tube (3Q5GT) is not used, and its filament is not heated since the A connection in the female receptacle is not completed.

[1] Emerson Radio and Phonograph Co., New York, N. Y., Model DJ-310.

When using battery power, the plug YZ is inserted in the female receptacle in the set, the Z connection bringing into the circuit the filament of the battery output tube by grounding the negative terminals of the 9-volt A battery. This Z connection also connects the negative terminal of the 90-volt B battery to ground. The filament of the 117L7GT is not lighted on battery power, and is connected through terminal Y to one side of the series filaments through a 1,000-ohm resistor.

Page 118 QST Dec 1939

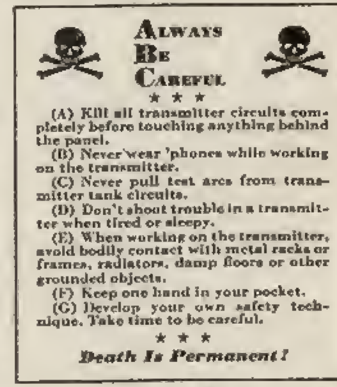


FIG. 4.—A good sign to be posted in high-voltage laboratory.

6. Battery-operated Receiver. Figure 6 is a completely battery-operated radio receiver which uses a 1.5-volt A battery for the five tube filaments which are connected in parallel, and two series-connected 45-volt B blocks.

Characteristic of both Figs. 5 and 6 is the permanent magnet moving-coil loud-speaker which does not require external power of any kind for energizing its magnetic circuit.

7. Receiver Power Unit for 115 to 230 Volts D.C. Figure 7 shows a vibrator-transformer-rectifier circuit which has two ranges of d-c power input, 105 to 125 and 210 to 250 volts. The link board shown in the figure changes the unit to suit the voltage available. These units provide an a-c heater voltage of 6.6 volts and rectified d-c plate potential of 360 to 400 volts.

8. Receiver Power Unit for A.C. and 6 Volts D.C. Figure 8 shows the two units required when a radio receiver is to be operated on either 6 volts d.c., or 5 ranges from 105/250 volts, 25 to 60 cycles.

The d-c power supply is a vibrator-transformer type, in which the vibrator also rectifies the high a-c voltage supplying d-c voltage.

The a-c power supply is a multitapped primary transformer, with a conventional full-wave rectifier tube. By placing the receiver power connector in the appropriate unit, the same receiver operates on either d.c. or a.c.

9. Transformer-rectifier Circuit for Transmitters. Figure 9 shows a typical high-voltage circuit for a transmitter, using six half-wave rectifier tubes. The tube filaments are all paralleled on a single-phase transformer, and the plate circuit comprises a three-phase transformer with a double Y secondary, which can be fed by a delta-connected primary.

Means are generally provided for placing a spare rectifier into the circuit. The spare tube filament is kept lighted, and its plate lead is connected to all six inactive jaws of a rack of 6 s.p.d.t. switches. The blade of each switch is the transformer lead, and the second jaw of each switch goes to the rectifier tube in use. If a rectifier fails, its s.p.d.t. switch is thrown (either automatically or by hand) from the regular tube plate to that of the spare tube—making the spare tube active and taking plate voltage off the regular tube so that it can be replaced when the s.p.d.t. switch is finally thrown back to the first position.

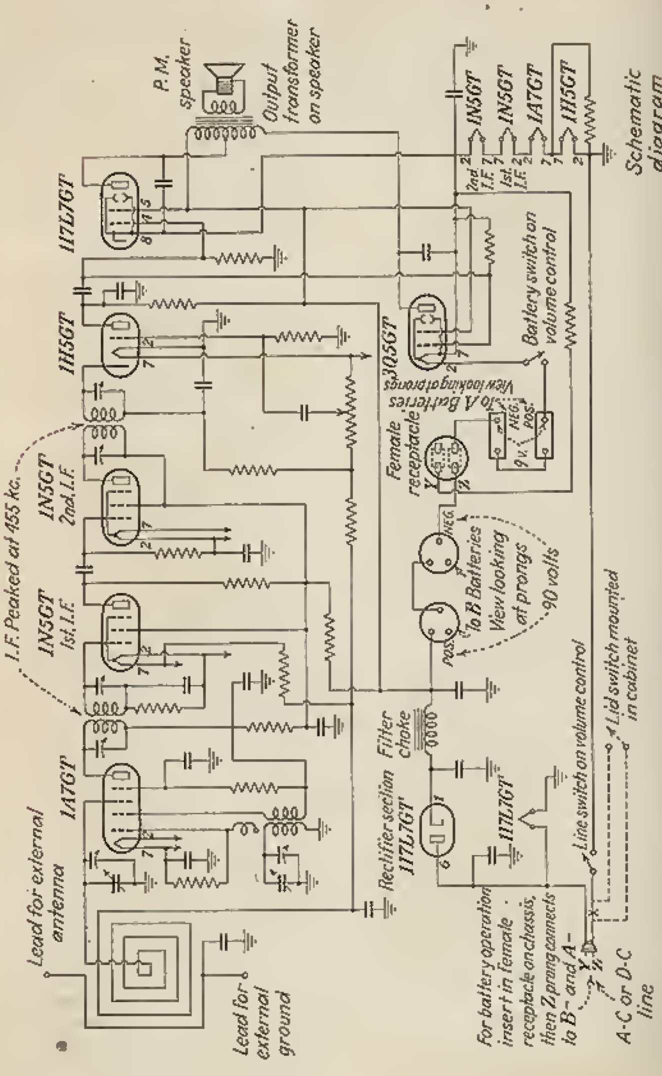


FIG. 5.—Circuit of receiver operable from either batteries or utility power.

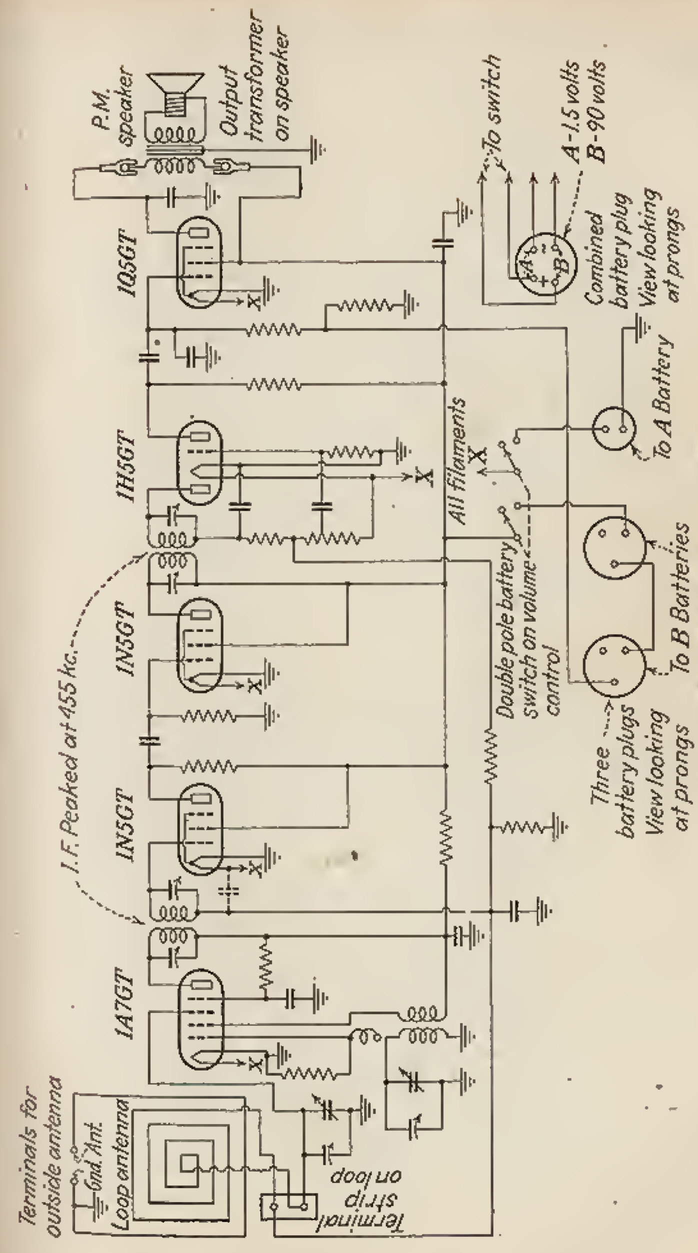


FIG. 6.—Battery-operated radio receiver.

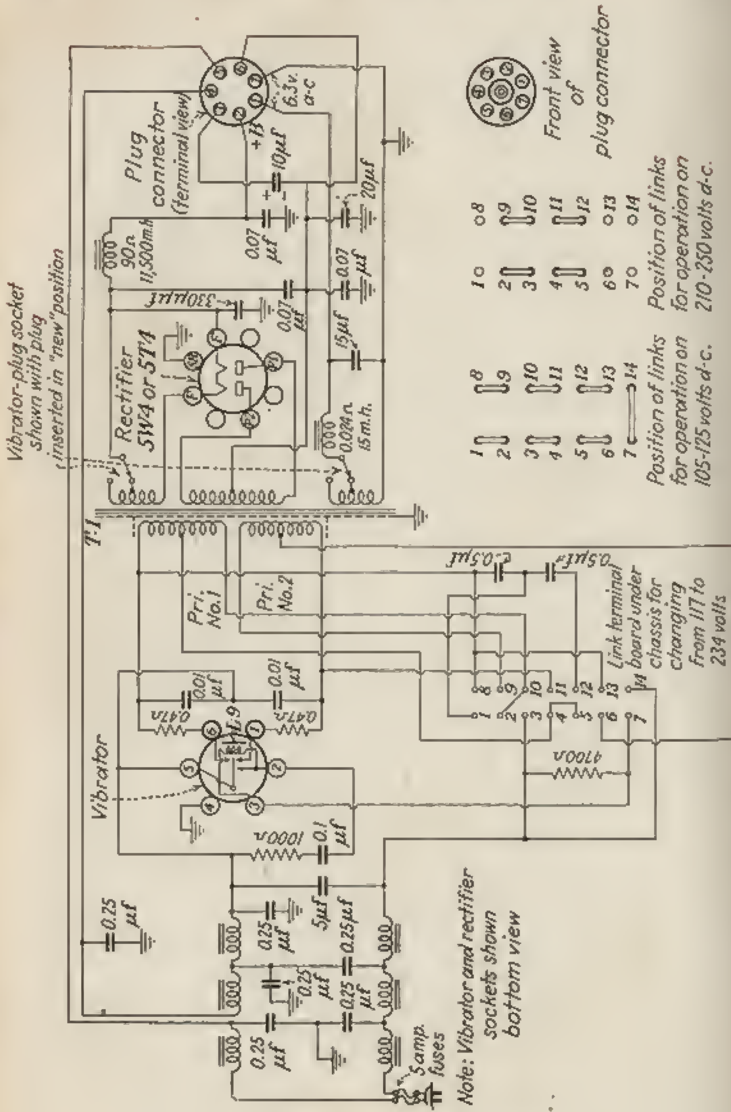


Fig. 7.—Power supply operable from 115 to 230 volts d.c.

The use of six rectifier tubes minimizes the possible trouble due to tubes, since the effect of losing one tube due to decreased emission is generally merely to introduce a hum into the rectified voltage supply, without decreasing the voltage very much.

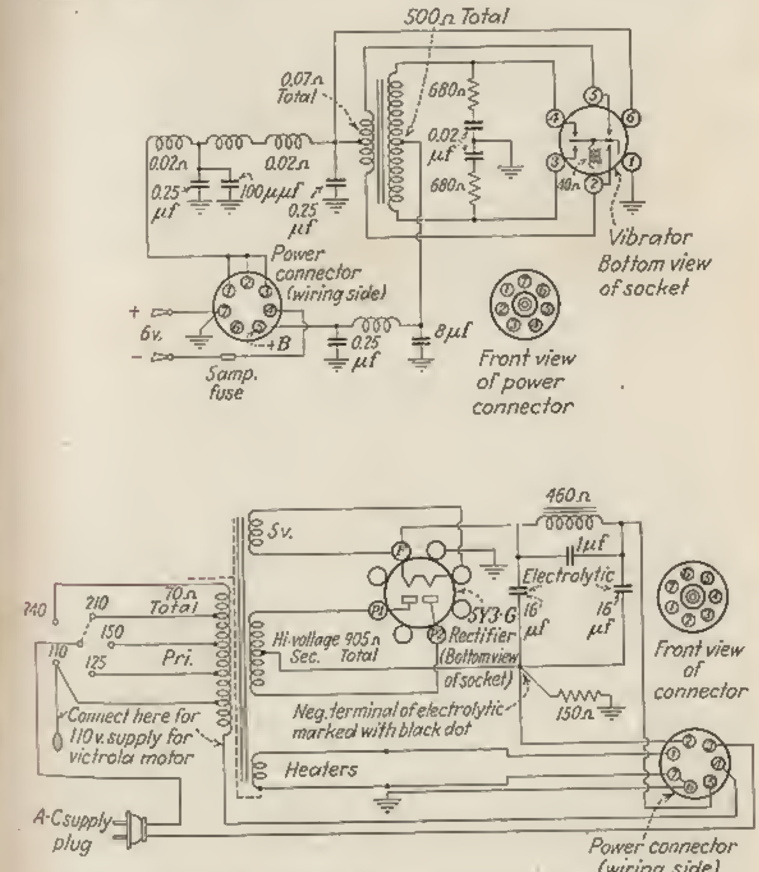
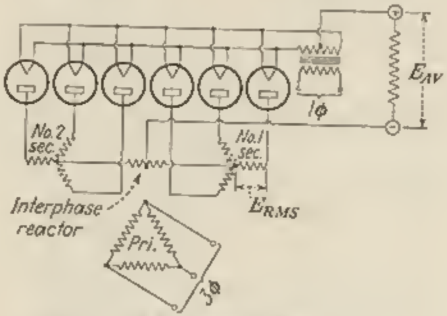


Fig. 8.—Power supplies from 6 volts d.c. and 110 to 240 volts a.c.

The interphase reactor between the common points of the double Y secondary is a center tapped transformer, carrying currents of six times the power-line frequency. The action of this device is like that of a choke used as an input filter. An iron core is used, with an air gap to limit saturation with the ensuing d.c. if a rectifier tube fails and thus destroys the balance of currents.

By using a suitably connected, three-phase filament transformer, some increase in the life of hot-cathode mercury-rectifier tubes can be obtained. The plate circuits are the same as shown in Fig. 9, but the filament supply of any one tube is connected so that it is 90 deg. out of phase with the plate



$$E_{RMS} = 0.855 E_{AV} = 0.408 E_{INV}$$
$$E_{AV} = 1.17 E_{RMS} = 0.478 E_{INV}$$
$$E_{INV} = 2.09 E_{AV} = 2.45 E_{RMS}$$

FIG. 9.—High-voltage supply for transmitter.

voltage supplied to it. This makes the filament an unipotential device at the moment that the crest plate current is drawn.

DRY-CELL PRIMARY BATTERIES

10. Ampere-hour Capacity.[1] The ampere-hour capacity obtainable from a dry cell of a given size depends on several factors, including formula, physical construction, current drain, hours of use per day, and cutoff voltage. For any specified discharge schedule there is an optimum value of current drain which will produce maximum capacity to a specified cutoff voltage. When the current is increased beyond this optimum value, the ampere-hour capacity decreases because of less efficient depolarization. When the current drain is less than the optimum, the ampere-hour capacity becomes less due to the subtractive effect of shelf deterioration.

Cells designed for heavy duty service will attain their peak capacity at higher current values than cells designed for light intermittent service.

It is not always practical to use the size and formula of cells which will operate at peak efficiency under an assumed set of service conditions, since size and portability may be the deciding factor on one hand, while the inconvenience and the cost of replacement may warrant the use of larger cells under certain conditions. Then too, the variety of service conditions to which battery-operated equipment may be subjected may indicate the use of a compromise size or formula of cell.

11. Cutoff Voltage. When cells are discharged at heavy current drains or for long continuous periods at more moderate current drains, a considerable increase in service life can be realized by using the cells to

[1] Articles 10 to 15 were supplied by Ralph E. Ramsay, Ray-O-Vac Co., Madison, Wis.

a lower cutoff voltage. The gain will depend on the size and formula of the cell. The lighter the current drain or service conditions, however, the higher the operating voltage and the flatter the discharge curve.

12. Shelf Life. For cells of a given formula and physical construction the shelf life will increase with the size of the cell. Loss of capacity in storage or during idle periods is due to local reactions, admission of oxygen, and loss of moisture. Certain cells designed for heavy duty industrial service achieve high initial service capacity by increasing the proportion of depolarizer to electrolyte, or by using more active oxides of manganese in the depolarizer, or both. In general these cells have a more rapid rate of deterioration on shelf than cells designed for light duty service. For C battery service, specifications call for a life of 18 months to 1.45 volts for the D size cell and 12 months for the B size.

TABLE I. DRY-CELL CAPACITY VERSUS DRAIN
For D-size B Battery Cells discharged 5 hr. per day, 5 days per week, to a cutoff voltage of 1.13; constant current discharge.

Current drain, milliamperes	To an end voltage of 1.13			Life to 0.80 volt in terms of life at 1.13 volts*
	Ampere-hours	Per cent of peak capacity	Weeks of service	
3	4.10	71.5	54.7	
5	4.90	87.5	39.2	104
7½	5.20	93	27.8	
10	5.60	100	22.4	108
15	5.20	93	13.9	114
20	4.90	87.5	9.8	119
30	3.25	58	4.3	129
40	2.40	42.9	2.4	138
50	1.70	30.4	1.4	

These values are for one size, one formula, and will not hold for other discharge schedules.

* Six hours per day, 5 days per week.

13. Effects of Temperature. Chemical reactions are accelerated by an increase in temperature. In the dry cell a temperature rise increases both the useful current-producing reaction and the parasitic local reaction during idle periods. The net effect on the total capacity delivered will depend on the balance between these two forces and will be different for various designs of cells.

Dry cells should be stored at low temperatures to minimize shelf reactions. This precaution is especially important for cells containing depolarizers of high activity. The high limits of temperature for dry-cell use are usually determined by the point at which seals flow or internal pressure expels the cell contents. This point may vary with the size, formula, and construction of the cell. Standard specifications call for a seal which will not flow at a temperature of 113°F. during a static test in which the sealed surface is held vertical for a period of 24 hr.

As the temperature decreases, the activity of the cell is lowered until finally it is unable to maintain a useful voltage while delivering current. The lowest temperature limit of use will depend on cell formula, cell size in relation to the current drain, and frequency of use. In general, dry cells will not furnish useful current when the actual cell temperature is less than $-15°$F. to $-20°$F. Freezing does not injure a dry cell, and capacity which cannot be realized at low temperatures will be available when the cell is returned to room temperature.

14. Amperage or Short-circuit Current. There is no relationship between the current delivered by a dry cell on a short-circuit amperage test and the service capacity of the cell. At best such a test is useful only in judging the uniformity of a particular lot of cells of a given formula.

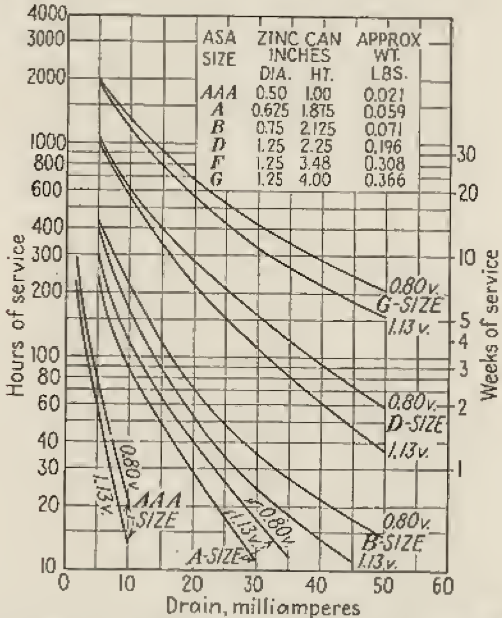


FIG. 10.—Radio-type dry cells; discharge *versus* time. Constant current, 6 hr. per day, 5 days per week to 1.13 volts and 0.80 volt per cell at 70°F.

15. Dry-cell Battery Standards. The American Standards Association in cooperation with the National Bureau of Standards issues a standard specification for dry cells and batteries. This standard sets forth various sizes of cells and batteries which are recognized as "standard" and also gives information on standard tests together with the corresponding performance requirements. There are many cells and batteries on the market which differ in size from those listed in the specification, and prospective users are advised to obtain current information from the dry-cell manufacturers.

16. Cost; Capacity; Weight; Life. It is fully realized that dry-cell characteristics will vary with the specific requirements of use, which in

turn influence the design. However, it may be of value to give some data which will allow the comparison of these cells as sources of electrical energy with other types of electrical equipment.

The cost of a dry cell varies from 0.6 to 11 cts. per watt-hour, and the weight of a bare cell ranges from 0.024 to 0.05 lb. per watt-hour. The volume ranges from 0.4 to 0.8 cu. in. per watt-hour. These values are based on an average of 1.3 volts per cell and optimum discharge rates.

Very small cells cost more in proportion to their capacity, since there are just as many manufacturing operations as in constructing a larger cell. Thus a premium is paid for increased portability.

The ampere-hour ratings of cells can also be computed, varying with the several factors which have been given, and are not proportional to the sizes of the zinc cans enclosing the cells. For discharge currents which are usually required from a dry cell, the variation in ampere-hour rating may cover a 3 to 1 range. A graph can be made from Fig. 10[1] with milliampere-hours as ordinate, and current drain as abscissa. From this the maximum milliampere-hour capacity at normal room temperatures under 6-hr. daily discharge for 5 days a week can be found, Table II being a typical schedule.

TABLE II. DRY-CELL CAPACITIES*

Cell	Maximum			
	Current drawn, milliamperes	Life: Hours	Life: Weeks	Capacity, milliampere-hours, to 1.13 volts
AAA	3	110	4	330
A	5	230	8	1,150
B	8	220	7	1,760
D	10	540	18	5,400
G	15	780	26	11,700

* Data from Burgess Battery Company and Ray-O-Vac Company.

The life of a dry cell is somewhat increased if the external current drawn is very low. In doing this, the capacity in milliampere-hours may be decreased by a few per cent. There is a more important factor, however, which is the time limit of shelf life, the case when no external current is drawn.

On expeditions which are to last several years, it is possible to carry along the separate ingredients from which dry cells can be made up. In this case the shelf life does not begin until the cells are assembled.

SECONDARY OR STORAGE BATTERIES[2]

17. Acid and Alkaline Cells. There are two general types of storage cells in use in the United States, the lead-acid and the nickel-iron-

[1] Data from Burgess Battery Company and Ray-O-Vac Company.
[2] General reference: VINAL, G. W., "Storage Batteries," 2d ed., John Wiley & Sons, Inc., New York, 1930. Helpful suggestions have been received from W. B. Manson, Thomas A. Edison, Inc.; A. E. Harrold, Willard Storage Battery Co.; H. H. Hudson, The Electric Storage Battery Co.; H. N. Stover, Philco Corporation, Storage Battery Division.

alkaline. Cells are usually combined in series to form batteries, but cells may also be used in parallel. Lead-acid cells form the major number of storage batteries in use. Lead is a very heavy metal which is not mechanically strong, and it is a tribute to the designers that such sturdy cells are available. The alkaline cell of nickel-iron is superior in its resistance to mechanical shock. The first cost of the alkaline cell is relatively higher, but its life expectancy is also longer than that of the acid cell.

18. Selecting a Storage Battery. To select a storage battery suitable for some particular use the following factors must be known:

1. Nominal circuit voltage.
2. Final permissible voltage.
3. Number of amperes required in use.
4. Hours of use before recharging.

The first two items determine the number of cells which are used in series, and the third and fourth items make definite the ampere-hour capacity of the battery. For example, if the lowest permissible value for a filament supply is 10.0 volts, the use of six acid cells at 1.75 volts each (total 10.5) or 10 alkaline cells of 1.0 volts each (total 10.0) are required. Further, since the initial voltage will be higher, unless the equipment is designed to use this higher potential, some means should be provided to reduce the value, *e.g.*, a series rheostat. The six acid cells will have an initial (charged) potential of $6 \times 2.05 = 12.3$ volts, an excess of 2.3 over 10.0, and the 10 alkaline cells will have an initial potential of $10 \times 1.45 = 14.5$, an excess of 4.5 volts over the required 10.0 volts.

19. Ampere-hour Ratings. The ampere-hour ratings are approximately determined by multiplying the hours of use before recharging, by the amperes drawn during that period. These ratings vary considerably with the length of discharge, as will be mentioned later. As a general rule, an acid storage-cell capacity should be adequate for at least 4 days of operation without discharging. The alkaline cells may be used on a 24-hr. charge-recharge schedule.

In specifying the capacity of a storage cell in ampere-hours, it is necessary also to give the rate of discharge and the permissible finish voltage because the ratings will vary over a wide range with changes in these two factors.

To specify "100 amp.-hr. at an 8-hr. discharge rate to 1.75 volts per cell" is quite definite and is one of the standard ratings for acid cells. These are usually termed "normal discharge rates."

20. Approximate Dimensions; Weight. Weight and volume of any battery can be approximated directly from Tables III and IV by multiplying the columnar values by the ampere-hour rating. A large range of cell capacities is included since the radio use for B voltage uses a low-capacity cell and for A voltages a high-capacity cell is required. For example, in Table IV since the usual range for stationary cells is 0.25 to 0.50 lb. per ampere-hour, a 100-amp.-hr. cell would be from

$$100 \times 0.25 = 25 \text{ lb.} \quad \text{to} \quad 100 \times 0.50 = 50 \text{ lb.}$$

in weight. Similarly the use of 100 amp.-hr. shows that the normal charge is $0.125 \times 100 = 12.5$ amp., which also is the normal 8-hr. discharge rate; and the trickle rate is $0.0025 \times 100 = 0.25$ amp.

TABLE III. LEAD-ACID STORAGE BATTERIES (10 TO 1,000 AMP.-HR.)*
All ampere-hours are for an 8-hr. discharge to 1.75 volts per cell at 77°F. to 80°F.

a. Weight, Capacity, Dimensions, per cell:

	Lb. per Amp.-hr.	Cu. In. per Amp.-hr.	Lb. per Cu. In.
Small capacity glass jar cells	1.1	25	0.05
Usual range of stationary cells	0.25–0.50	3.9–7.7	0.06–0.08
Starter batteries, high specific gravity	0.18	2.4	0.07

b. Charging Amperes:†

	Multiply Amp.-hr. by
Normal for 1.250 to 1.150 sp. gr. stationary cells	0.125
Trickle for 1.250 to 1.150 sp. gr. stationary cells	0.0025

	Per Cent	Multiply Amp.-hr. by
Normal current	100	0.125
Start	135	0.169
Finish	35	0.044
Trickle	2	0.0025

c. Charging Volts 2.5 at normal ampere rate, approximately 10 hr. to final specific gravity.

d. Discharging Amperes to Various End Voltages:

	Final Volts	Per Cent Rated Amp.	Factor by Which to Multiply Amp.-hr. to Get Amp. Dischg.	Resulting Factor by Which to Multiply Amp.-hr. Capacity
72 hr. discharge‡	1.75	16	0.020	1.50
8 hr. discharge	1.75	100	0.125	1.00
3 hr. discharge	1.72	200	0.250	0.80
1 hr. discharge	1.60	400	0.250	0.55
1 min. discharge	1.40	1,000	0.500	0.04

e. Volts versus Specific Gravity When Charged:

	Specific Gravity	Volts
Stationary cells	1.215–1.250	2.05 (charged)
Starter cells	1.280	2.10 (charged)
	1.300	2.20 (charged)

f. Freezing Points of Electrolyte:

Specific Gravity	Degrees Fahrenheit	Degrees Centigrade
1.280	−96	−71
1.250	−61	−52
1.200	−17	−27
1.180	− 6.5	−21
1.160	1.6	−17
1.150	5	−15
1.100	18	− 7.8

* Compiled from data furnished by Willard Storage Battery Co., Cleveland, Ohio; Gould Storage Battery Co., Depew, N. Y.; Philco Corporation, Storage Battery Division, Philadelphia, Pa., The Electric Storage Battery Co., Philadelphia, Pa.

† Normal charging amperes for 100 amp.-hr. cell is $0.125 \times 100 = 12.5$ amp.

‡ Suppose battery 8 hr. capacity is 200 amp.-hr. and the 8 hr. discharge rate is $200/8 = 25$ amp. For 72 hr. discharge 16 per cent of $25 = 0.020 \times 200 = 4$ amp., and the actual capacity now is $1.50 \times 200 = 300$ amp.-hr.

Based on an average voltage of 1.90, a lead-acid stationary cell of 100 amp.-hr. or greater capacity, weighs from 0.13 to 0.30 lb. per watt-hour and occupies from 2.1 to 4.1 cu. in. per watt-hour. Using an average voltage of 1.20, a nickel-iron alkaline cell of 100 amp.-hr. or greater weighs from 0.06 to 0.13 lb. per watt-hour, occupies from 1.4 to 1.9 cu. in. per watt-hour, and costs from 7½ to 12½ cts. per watt-hour for capacities from 150 to 900 amp.-hr.

21. Types of Charge and Discharge; Life and Cost.[1] The severest type of work for a storage battery is "cycling," meaning that the cells are run from complete charge to a complete discharge before recharging. As an approximate figure an average acid battery after 400 of such cycles

[1] H. H. Hudson, The Electric Storage Battery Co., New York City.

has about 75 per cent of its original ampere-hour capacity left, and the same rating for an alkaline battery is about 1,400 cycles.

The life guarantees of lead-acid storage cells vary with three major factors: the type of construction, the ampere-hour capacity, and the extent of discharge before recharging. This may be completely stated by assuming the extent of daily charging and discharging of cells over 60 amp.-hr. capacity. On "full float" service or not over 5 to 10 per cent daily charge-discharge, the life varies from 8 to 14 years. If the discharge and charge is 20 per cent daily, the life is 6 to 10 years, and, when discharged and charged 40 per cent daily, the life is 4½ to 7½ years. At the end of these times cells will have approximately 75 per cent of their new ampere-hour capacity.

The range of cost of a lead-acid cell depends both on the construction and the ampere-hour capacity. For low capacities 10 to 60 amp.-hr., the cost is about 10 cts. per watt-hour, and for capacities from 100 to 1,000 amp.-hr. the cost per watt-hour ranges from 3½ to 6 cts. Note that the nominal voltage is 2 volts and that the cost per ampere-hour per cell would be doubled.

For a given capacity in watt-hours the initial cost of a lead-acid cell is less than that of a similar nickel-iron-alkaline cell. The life of the latter, however, is definitely longer, and usually over a long period of time the cost is not appreciably different for the two types. It should be mentioned that stationary-type lead cells are higher in quality and also in first cost than cells used on automobiles. As a result the stationary cells have considerably longer life than the usual automotive type of cell.

Acid storage cells can be satisfactorily trickle charged and thus kept available for emergency service. Alkaline storage cells thrive best when charged at the normal rate, although they may be trickle charged if the service does not require high discharge rates. The alkaline cell when trickle charged will not deliver as good voltages at the higher discharge rates as when it is "cycled," but in actual operation the cell may be selected to meet such discharge requirements. However, a periodic complete discharge and charge at normal rate is recommended if full alkaline cell capacity is desired.

The regulation of a cell (the maintenance of terminal voltage under load) varies with the internal resistance. The lead cell will have better regulation than the nickel-iron cell.

Acid storage cells have a lower internal resistance than alkaline cells of the same ampere-hour ratings. In the case of stationary cells this factor is from ½ to ⅓. On momentary exceptionally heavy-load conditions the acid cell can deliver from two to three times the current that can be drawn from an alkaline cell. This may be both an advantage and a disadvantage. On a short circuit the damage to an acid cell is usually to the plate lugs and the top connectors, but the alkaline cell is not harmed. It is suggested for alkaline cells that periodic complete discharge be followed by an intentional short circuit and then completely charged. If an acid cell is short-circuited, it should be recharged immediately. Alkaline cells can be stored in a discharged and short-circuited condition indefinitely, but acid cells should be stored fully charged.

Acid storage cells are essentially low-resistance devices and therefore particularly useful for applications where very high currents are required. However, for each ampere-hour of electricity delivered, a definite weight of lead peroxide and sponge lead must be converted to lead sulphate by

electrolytic action. A very high current discharge rate causes a progressive slowing down of the electrolytic action until more active electrolyte can diffuse to and through the plates, and there is a reduction of the total capacity of electric energy stored at this particular high rate. To take an extreme case, a certain battery rated at 100 amp.-hr. on an 8-hr. discharge to 1.75 volts per cell will deliver 150 amp.-hr. at a 72-hr. rate to 1.75 volts per cell, but only 55 amp.-hr. when discharged in 1 hr. to 1.68 volts per cell, as shown in Table III.

TABLE IV. SUMMARY* OF NICKEL-IRON STORAGE BATTERIES (11 TO 900 AMP.-HR.)

All ampere-hours are for 3- to 5-hr. discharge to 1.00 volt per cell; above 60°F.

a. Weight—Capacity Dimensions, per cell:

	Lb. per Amp.-hr.	Cu. In. per Amp.-hr.	Lb. per Cu. In.
Small capacity (11 to 20 amp.-hr.)	0.28	6.6	0.043
Large capacity (100 amp.-hr. up)	0.11	1.7–2.3	0.048–0.060

b. Charging Amperes for 7 Hr.:

	Multiply Amp.-hr. by
Normal	0.200
Trickle	0.0066
Start	0.400
Finish	0.074

c. Charging Volts:

Hold at 1.7 volts per cell to obtain start and finish amperes of item *b*, above.

Charge at normal amperes to 1.8 or 1.9 volts per cell until voltage per cell remains constant for ½ hr.

d. Discharging Amperes:

Hours Discharge	Final Volts	Relative Ampere-hours
10	1.05	1.00
5	1.00	
2.5	0.91	
1	0.64	0.96

e. Freezing Point of Electrolyte:

Note that electrolyte density varies little, and does not show state of charge. At minimum sp. gr. of 1.160 at 60°F.:

Starts to freeze out at −4°F., −20°C.

Freezes to "slush" at −87°F., −60°C.

* Compiled from data furnished by Thomas A. Edison, Inc., West Orange, N. J.

While the capacity in ampere-hours of a lead cell decreases at increased discharge rates, this capacity in a nickel-iron cell does not decrease very much at high rates up to five times normal if no restrictions are placed on the useful voltage. At five times normal discharge rate the nickel iron cell delivers 96 per cent of its normal full rate ampere-hour capacity. This point is of theoretical rather than practical interest, however, since there is a lower limit below which the voltage cannot drop and be of actual use. At five times normal rate to get the 96 per cent amp.-hr. in a nickel-iron cell, the end voltage is 0.64, which is too low for most purposes.

22. Electrolyte Characteristics. The specific gravity of an acid cell varies with the amount of charge, being greatest at full charge and least when discharged. The freezing temperature of an acid cell varies with the amount of charge, a discharged cell at specific gravity 1.10 will freeze at 18°F., and a charged cell at sp. gr. 1.25 will freeze at −61°F. The electrolyte density of an acid battery is a readily tested indication of the extent of its charge.

On the other hand, the electrolyte specific gravity of an alkaline cell is no indication of its state of charge, since it remains practically constant during charge and discharge. However, owing to gradual deterioration, the specific gravity of the alkaline electrolyte in a cell ranges between 1.215 (new) to 1.160 and when it reaches this latter value it should be replaced. Electrolyte having a specific gravity of 1.160 at normal temperature will start to freeze out at −4°F. and will freeze into a slushy snow at −87°F. Higher gravity electrolyte starts freezing at correspondingly lower temperatures.

23. Effect of Temperature on Capacity. The capacity of a lead-acid cell decreases about 1.3 per cent for each degree C. the temperature is lowered (0.75 per cent per degree F.). The usual 100 per cent rating at 25°C., 77°F., is decreased to 50 per cent at −25°C., 10°F.

At about 5°C., 41°F., the alkaline cell has the minimum output at the normal rates of discharge. However, the higher internal resistance heats up in a cell which is either being charged or discharged, and in actual use the 5°C. temperature is only a rough approximation.

There are certain details which influence the construction of an acid storage cell for specified conditions. If the discharge rate must be high, the use of wood separators between the plates rather than perforated rubber separators are preferred, due to the lower internal resistance of cells having the former. Further, lead-plated copper connectors between cells can be used to reduce the resistance. An increase in life is obtained if the specific gravity of the electrolyte is reduced from the normal charged value of 1.250 to 1.210 or 1.220. (Note that stationary cells use lower values than those for portable or starting cells, where the specific gravity may be as high as 1.300.) However, the reduction of the maximum specific gravity of the electrolyte decreases the capacity of batteries at the 8-hr. discharge rate, or longer, but has no noticeable effect at higher rates or shorter times of discharge. One of the benefits resulting from the use of acid of lower specific gravity is the increased life of the wooden separators.

24. Charging Storage Batteries. There are several standard schemes of charging storage cells which are detailed in books on the subject and explained in booklets issued by battery manufacturers. Briefly the methods are termed the *step*, *constant voltage*, *modified constant voltage*, *constant current*, and in addition there are *equalizing* and *booster charges*. To charge fully a battery in a minimum time, the step method resembles the booster charge since both give a heavy current charge at the start. Since each of these methods requires a high-capacity charger, the constant-current method is often preferred. On the other hand the constant-voltage scheme automatically tapers the rate of current charge, since a discharged cell potential is low and a maximum difference between cell potential and charging potential is available to force through a heavy charging current at the start.

Under emergency conditions a battery may be prematurely and completely discharged. This means that in addition to a trickle charger, another rectifier should be available with sufficient capacity to charge the battery in from 10 to 24 hrs.

In charging, the positive potential lead of the charger is connected to the positive cell lead, forcing current to flow in the cell in the opposite direction to that of a cell which is furnishing current on discharge.

25. Relays for Charging. Relays to control automatic charging are actuated by the voltage of the storage cells. Ordinarily the increase

of the resistance of the relay exciting coil with increase in temperature will require a higher voltage to initiate the relay action.

In one relay[1] this problem is met by using a bimetallic strip on which the actuating contacts are placed. This strip bends with temperature change and intentionally overcompensates for temperature, so that the voltage temperature requirements of the battery are followed faithfully.

When storage batteries are used, indicating instruments are of great value in showing the instantaneous charging or discharging current and voltage across the battery when charging or the battery voltage when discharging.

26. Integrating Meters. An integrating ampere-hour meter is frequently very desirable. The mercury-motor d-c ampere-hour meter[2] is admirably suited for use with a storage battery. The use of a mercury-floated rotating disk eliminates the use and maintenance of a commutator

Table V. Fuel Consumption per Kilowatt-hour, Engine-generator Sets

	Gasoline, gal. per kw-hr.	Fuel oil, gal. per kw-hr.	Gas, 800 B.t.u. per cu. ft., cu. ft. per kw-hr.	Diesel engines, gal. per kw-hr.
Full load	0.21 = 100%	0.17 = 100%	28 = 100%	0.13 = 100%
¾ load	119%	118%	118%	104%
½ load	138%	135%	137%	108%
¼ load	172%	171%	172%	112%

and brushes. The cumulative ampere-hour dials are very desirable since they show at a glance the extent of charge in the battery. Standard dials are available for many ranges from 150 to 1,200 amp.-hrs., and the current ranges of standard meters cover practically all possible charging rates. In one model two-rate charging is provided by a switch which, when a preset number of ampere-hours, say 20 per cent of the battery capacity, has been supplied at the first high rate, initiates a suitable circuit breaker giving the second lower rate of charge. In addition, when the full number of ampere-hours have been supplied to the battery, another switch operates to discontinue charging. Other models are available without switches, so that charging rates are manually adjusted as required and shown by the dial indication.

Since all storage batteries require more ampere-hours when charging than discharging, in all meter models an ingenious mechanism actuated by the direction of disk, which runs forward when charging and backward when discharging, requires more ampere-hours when charging than discharging to show a given amount on the dial.

The amount of excess in charging over discharging can be adjusted up to a maximum of 35 per cent to suit both the type of battery used and the normal discharge rate which is desired.

The main precaution in installing mercury-motor d-c ampere-hour meters is to avoid excessive heat and continuous vibration. A clever design of the mercury chamber prevents damage if the meter is turned over during shipment or prior to installation.

[1] TVR Relay, U. S. Patent 1960198. The Electric Storage Battery Co.
[2] Type N, Sangamo Electric Co., Springfield, Ill.

TABLE VI. GASOLINE-ELECTRIC GENERATING SETS

Model and cylinders	Engine		Generator		Floor space, inches	Height, inches	Type starting	Total weight, pounds	Cooling
	R.p.m.	Horsepower	Volts	Watts					
2B6, 1*	2,250	1/4	6†	200	15 × 12	13	Self-cranking	43	Air
2B12, 1	2,250	1/3	12	200	15 × 12	13	Self-cranking	43	Air
5B3, 1	2,100	3/4	32	500	19 × 15	18	Self-cranking	97	Air
8B3, 1	1,800	1 3/4	32	800	20 × 15	23	Self-cranking	165	Air
10DH1, 1	1,800	2	110	1,000	20 × 17	23	Rope	195	Air
10AA1, 1	1,800	2	120	1.0 kva 60 cycle	20 × 17	23	Automatic‡	265	Air
15AB3, 2	1,650	3	32	1,500	35 × 19	23	Automatic§	375	Water Radiator
50AA1, 4	1,200	18	110	5 kva 60 cycle	62 × 20	37	Automatic	1,100	Water Radiator
60AD1, 4	1,200	18	120	6,000	58 × 20	37	Automatic	1,070	Water Radiator
Petrel, 6‖	1,200	115	...	50 kw	36 × 118	49	Electric	8,400¶	
Dolphin, 6	1,200	180	...	100 kw	33 × 144	57	Electric	9,900	
Viking, 18	1,200	600	...	350	51 × 204	82	Electric	26,750	

* Delco Appliance Division, General Motors Sales Corporation, Rochester, N. Y.
† Volts are nominal, 6-volt generator will charge a 6-volt battery.
‡ A-c plants may be started by hand crank, remote control, or automatically.
§ D-c plants may be started by a voltage control on battery, by load, or by remote push button.
‖ Sterling Engine Co., Buffalo, N. Y. All generators are three-phase 60-cycle direct-converted exciters. Other d-c and a-c models are available.
¶ Including generator, bedplate, radiator, and auxiliaries. Radiators are based on full hp. and 100°F. ambient.

TABLE VII. FUEL-DRIVEN ELECTRIC GENERATING SETS*

Model	Cylinders	Displacement, cu. in.	Watts or kva	Volts	Phase	R.p.m.	Floor space, inches	Height, inches	Starting	Weight, pounds	Fuel	Cooling
ES	1‖	...	800 watts	32	d.c.	1,250	24 × 17	29	Semi-automatic	292	Kerosene	Air
IB	2	34	1,500 watts	32	d.c.	1,400	35 × 17	26	Automatic	420	Gasoline or gas	Air
23	4	69	3,500 watts	110/125†	d.c.	1,500	47 × 19	38	Various§	850	Various‡	Water Radiator
13	4	69	3 1/4 kva	120/240‡	1	1,200	57 × 19	38	Various	950	Various	Water Radiator
215	6	218	15,000 watts	110/125	d.c.	1,200	76 × 24	49	Various	1,900	Various	Water Radiator
316	6	288	20 kva	120/240	3	1,200	98 × 29	64	Various	2,500	Various	Water Radiator
240	6	501	40,000 watts	110/125	d.c.	1,200	103 × 35	67	Various	4,800	Various	Water Radiator
336	6	501	45 kva	120/240	3	1,200	117 × 35	67	Various	4,600	Various	Water Radiator

* Continental Motors Corporation, Muskegon, Mich.
† Direct current 110/125 volts, two wire; 220/250 volts, two wire.
‡ Three phase can be 120/240 volts, three wire, or 120/208 or 220/380 volts, four wire. One phase can be 120/240 volts, two wire, or 120/240 volts, three wire.
§ Starting can be hand crank, electric, remote, emergency automatic, full automatic. Fuel can be gas or gasoline, derate 11 per cent for natural gas. Models with other ratings also available.
‖ 1 1/2 hp.

27. Precautions to Observe in Using Lead-acid Storage Batteries.

1. Keep the level of the electrolyte covering the plates and insulation.
2. Use distilled or approved water to replace loss by evaporation.
3. Do not allow the cells to stand for any great length of time after the specific gravity has reached the lower limit. They should be given a charge to bring up the specific gravity.
4. Charge for the proper length of time at the proper rate.
5. Do not add acid or electrolyte to the cells.
6. Do not allow the temperature of the electrolyte to rise above 110°F, or 43°C.
7. Keep the battery and the battery compartment clean and dry.
8. Keep the terminals clean, tight, and well covered with vaseline.
9. Keep away from flames when charging with vents open.

28. Precautions to Observe in Using Nickel-iron Alkaline Storage Batteries.

1. Keep the plates covered with electrolyte.
2. Use distilled or approved water to replace losses by charging and evaporation.
3. An occasional short circuit is not detrimental, in fact this is suggested after complete discharge, before recharging.
4. Hydrometer readings mean little; charge to 1.8 to 1.9 volts per cell until the cell voltage remains constant for ½ hr.
5. Renew the electrolyte by that supplied by the manufacturer when the specific gravity decreases to 1.160. Do not pour off old electrolyte until the new is ready to put in.
6. Do not put acid in the cells.
7. Do not allow the cell temperature to rise above 115°F.
8. Keep away flames while charging.

FUEL-DRIVEN ENGINE-GENERATOR SETS

29. Types of Fuel. Gasoline is widely used, but kerosene and natural and manufactured gas, as well as No. 1 fuel oil and commercial Diesel fuel, are also in general use. Gasoline-driven generators give from 3½ to 5 kw-hr. per gallon, and Diesel-driven generators give 7.7 kw-hr. per gallon. Proper engine adjustment and good grades of fuel will improve these outputs.

TABLE VIII. DIESEL-ELECTRIC GENERATING SETS*

Name	Cylinders	1,200 r.p.m. engine horsepower	A.c., kilowatt, 80 per cent P.F.	Floor space, inches	Height, inches	Total weight, pounds
DG3C	1†	5	3	22 × 56	44	1,475
DGH10C	4‡	19	10	23 × 72	44	2,000
DGH25C	6	42	25	24 × 90	44	2,850

* John Reiner & Co., Inc., New York. Cooling is either by radiator and fan or by cooling tank or tower. D-c models available with same ratings as above.
† Single cylinder models are of Stover type, hand-crank starting.
‡ Four- and six-cylinder models are Hercules four-cycle full Diesel operation.

The Diesel engines of Tables VIII and IX are characterized by a high compression ratio, about 16:1, and have good efficiency at various load percentages, but the speed range is limited by that of its particular design. This may be a disadvantage if different output voltages are to be secured by varying the engine speed, but for a-c generation, where the frequency depends directly on speed, this is an advantage. The Diesel is being steadily improved in design and is increasing in its popularity.

The gasoline engine, characterized by a lower compression ratio (about 5:1) and spark-plug ignition, is further along in its state of perfection, and this means less maintenance trouble and also ease of starting. For large Diesel engines a small gasoline engine is supplied for starting.

Kerosene is often used as a fuel, in which case auxiliary means are generally provided for starting the engine with a fuel of higher volatility, such as gasoline.

30. Cooling. For cooling, the small-sized engines use air which is forced past the cylinder blocks by a fan. Large sizes use water cooling, and in general this can be supplied either by a water tower or by a fan-radiator-pump system. The latter is usually recommended since the use of proper water minimizes the formation of scale in the jacket passages of the engine block.

31. Emergency Service. Fuel-driven generators are often employed as an emergency source of electrical energy. Automatic switches may be used to start up the engine when the regular power has been discontinued for 15 sec. This requires an electric starting motor and a storage battery as part of the equipment. For emergency service a high-speed engine permits the rated capacity to be obtained in a minimum of space, and for a reduced first cost. For regular service the heavy-duty, slow-speed engines are generally recommended, although they require more space and have a higher initial cost.

Diesel-powered electric plants are not usually available for "full-automatic" starting. Semiautomatic operation is sometimes employed, in which the Diesel is manually started when the load is expected and operated at low idle continuously. With the plant running, if the normal power supply fails, relays operate to cause the Diesel to speed up and take full load.

TABLE IX. DIESEL-POWERED ELECTRIC GENERATING SETS*

Model	Cylinders	Displacement, cubic inches	Kilowatts†	R.p.m., 60 cycle	Floor space, inches	Height, inches	Starting	Approximate net weight, pounds
D17000	8	1,662	85	900	163 × 41	78	27-hp. gasoline engine or 32-volt electric motor	16,000
D8800	4	831	41	900	131 × 50	72	17½-hp. gasoline engine	10,500
D4600	6	468	30	1,200	124 × 41	59	15-hp. gasoline engine or 24-volt electric motor	6,200
D-3400	4	221	15	1,200	96 × 26	49	10-hp. gasoline engine or 12-volt electric motor	4,200

* Caterpillar Tractor Co., Peoria, Ill. Cooled by water in radiator, plus fan and water pump. Other ratings also available.
† D-c generators: 125, 250, or 600 volts, two wire; 125/250 volt, three wire.
A-c generators: three phase, 50 or 60 cycle, 110, 220, 440, 120/208, 120/240, 240/480, and 2,300 volts available; two phase, 50 or 60 cycle, 240, 480, and 2,300; one phase, 50 or 60 cycle, 120 and 240 volt. All a-c ratings are at 80 to 100 per cent power factor.

32. Regular Service A-c Generators. Fuel-driven generators are also employed as a primary source of electrical power, and for this type of work may be divided into two classes, a.c. and d.c. For a-c generators, where no storage of electrical energy is possible (unless a converter and storage battery are also provided), it is desirable to supply means for stopping and starting the engine, so that it runs only when a load is connected. This is sometimes called *full automatic* control and suitable relays and switches operate to start up the engine generator when a load of 1 per cent or more of the rated capacity is connected to the line. While running, any load can be supplied up to the rated capacity. When the load is decreased to less than 1 per cent of the rating, the engine shuts down. Push-button stopping and starting is also employed on a-c plants because it is less expensive than the full automatic control.

33. Regular Service D-c Generators. When the generator supplies d.c., the problem may be somewhat different than that just described. A d-c generator is almost always used to charge a storage battery, so that, when the battery is fully charged, it is seldom necessary to run the generator continuously. A voltage-controlled relay may be used to start up the generator when the battery voltage drops to a certain amount. This arrangement automatically charges the battery when its voltage decreases and shuts down the generator when the proper point of charge has been reached. Light loads are carried by the battery alone, but heavy loads pull down the battery voltage so that the generator starts up and helps to carry the load.

Push-button stopping and starting is often used for d-c as well as for a-c plants. When a d-c plant is to furnish fairly steady loads of over half its rated capacity, sometimes the storage battery is not used at all. In this case the automatic starting on battery voltage is not used, but it is possible to employ an automatic device to start up the engine when 1 per cent or more of the rated load is connected. Hand-crank starting and the use of a rope starter are also employed, mainly for small-sized plants used to charge storage batteries.

A convenient formula for determining wire sizes requires the knowledge of the amount of current supplied, the distance, and the maximum allowable voltage drop along the set of two wires. This maximum voltage drop is 2 volts for a 32-volt system, and 10 volts for a 110-volt system. The formula is as follows:

$$\frac{22 \times \text{load amperes} \times \text{distance in feet (one way)}}{\text{Allowable voltage drop}} = \text{wire size in circular mils}^1$$

While this formula was primarily derived for use with various voltages on two-wire d-c systems, it is also of value in single-phase two-wire a-c systems of 110 volts.

34. Cost; Capacity; Weight. Gasoline-engine driven generators of 200 to 6,000 watts output capacity range in weight from 0.18 to 0.22 lb. per watt, occupy 7 to 12 cu. in. per watt, and cost 17 to 35 cts. per watt.

For capacities of 50 to 350 kw the weights range from 0.08 to 0.17 lb. per watt, occupy 2.4 to 4.2 cu. in. per watt, and cost 3.5 to 6 cts. per watt.

Diesel-driven generators of 3 to 25 kw weigh from 0.11 to 0.49 lb. per watt, occupy 4 to 18 cu. in. per watt, and cost from 11 to 33 cts. per watt.

[1] Delco Appliance Division, General Motors Sales Corp., Rochester, N. Y.

For ranges of 15 to 85 kw, Diesel generators range from 0.19 to 0.29 lb. per watt, occupy 9 to 16 cu. in. per watt, and cost from 7½ to 13½ cts. per watt.

TABLE X.* DIRECT-CURRENT TO ALTERNATING-CURRENT CONVERTERS; OUTPUT A.C. = 110 VOLTS 60 CYCLES, 1 PHASE

Code No.	Input, volts d.c.	Input, amperes d.c.	Output, volt-amperes a.c.	
1010 K	110	1.6	110	Heavy-duty 1,800 r.p.m. 4-pole ball bearings, 7 × 10¼ × 8 in., 35 lb.
1020 K	110	2.7	200	
1075 K	110	10	750†	
3215 K	32	7.8	150	
3250 K	32	25	500‡	
A-080	6	19	80	2-pole converter, 4½ × 5 × 8½ in., 13½ lb.
B-1215§	12	21	150	
C-3250	32	7.6	150	
D-1015	110	2.6	150	

* Carter Motor Co. Chicago, Ill.
† 75 lb., 7 × 16 × 8 in., with starter box.
‡ 55 lb., 7 × 12½ × 8 in.
§ Models with 40-volt-amp. output and weight of 8 lb. are available, operating from 6, 12, 32, and 110 volts d.c.

TABLE XI. DIRECT-CURRENT TO ALTERNATING-CURRENT CONVERTERS; OUTPUT A.C. = 110 VOLTS, 60 CYCLES, 1 PHASE

Type No.	Input, volts, d.c.	Input, amperes, d.c.	Output, volt-amperes, a.c.	Length, inches	Width, inches	Height, inches	R.p.m.	Weight, pounds
640*	6	13.3	40	10	6	10	.....	21
1216	12	20	160	10	6	10	.....	26
3230	32	15	300	11	6	10	.....	30
1130	110	3.9	300	11	6	10	.....	30
11T100	110	....	1,000	14	8	9	.....	73
2R61†	6	....	40	..	..	..	3,600	37
2R121	12	....	80	..	..	..	3,600	46
2R319	32	....	80	..	..	..	3,600	39
4R328	32	....	2,000	..	..	..	1,800	245
2R151	115	....	90	..	..	..	3,600	33
4R158	115	....	2,500	..	..	..	1,800	315

* Pioneer Gen-E-Motor Corp., Chicago, Ill.
† Electric Specialty Co., Stamford, Conn.
Other power ratings are available.

MOTOR-GENERATOR SETS

35. Direct-current to Alternating-current Converters. Converters permit a storage battery to be used as a source of energy to operate a-c devices, even when the charging d-c generator is not running. Since many radio units are designed to operate on 110 volts 60 cycles, the converter is often a valuable piece of equipment. Change-over switches can be provided to connect the radio unit normally to utility a-c power and in an emergency to switch to the storage-battery-converter equipment for continued operation.

36. Dynamotors. When low voltage d.c. is available, a dynamotor can be used to supply d-c plate voltage. Marine, aircraft, police, sound systems, and amateur use are among the types of service which employ dynamotors. The single armature, two bearings, and general compactness are features of the designs. Dynamotors are usually designed to have a high efficiency in order to conserve the limited capacity of the storage-battery power source. High speed is usually employed to get a maximum output with restricted space and weight limitations.

37. Motor Generators. A motor generator usually comprises two distinct units, each with two bearings, generally coupled together mechanically to run at the same speed. Motor generators are used for both emergency supply and regular operation. In the latter cases low-speed units are preferred, as they have longer life and require less maintenance.

Part of Table XII gives typical motor-generator ranges, including double-current generators. The latter have two distinct d-c supplies, one being high voltage for the plate circuits and the other low voltage for the filament circuits.

38. Price and Weight Ranges. There are several factors which cause variations in price and weight for a given output. For d-c to a-c converters, lightweight high-speed devices are available for low outputs, ranging in price from 15 to 46 cts. per watt and weighing from 0.07 to

TABLE XII.* PLATE-VOLTAGE GENERATORS, DOUBLE-CURRENT GENERATORS, A-C AND D-C MOTORS, AND DYNAMOTORS

Item No.	Motor	Generator				D.C. and 60 cycle, r.p.m.‡	
		Volts	Watts,† a.c. and 60 cycle	Filaments Volts	Filaments Amperes		
2	D.c., 50, 60 cycle	350	40	....	.....	3,500	
17	D.c., 50, 60 cycle	1,000	1,000	....	.....	1,750	
372	D.c., 50, 60 cycle	5,000	9,000	....	.....	1,750	
28	D.c., 50, 60 cycle	400	50	6-12	40	3,500	Double-current generators
405	D.c., 50, 60 cycle	4,000	4,000	8-16	1,200	1,200	
45	32-230 d.c.	350	40	....	.....	3,500	Dynamotors 32-230 volts d.c.
62	32-230 d.c.	1,000	800	....	.....	2,000	
			Volts	Weight, pounds			
100	12 volts, 3 amp.	150/250	15	6.75		....	Special high efficiency lightweight ball-bearing
105	12 volts, 26.5 amp.	250/800	100	20			
110	12 volts, 50.5 amp.	1,200/1,500	400	36			

* Electric Specialty Co., Stamford, Conn.

† 50-cycle output, approximately 75 per cent of that when using d.c. and 60-cycle motors.

‡ 50-cycle r.p.m., approximately five-sixths that of d.c. and 60 cycle.

TABLE XIII. DIRECT-CURRENT PLATE-VOLTAGE DYNAMOTORS

Type No.	Input Volts	Input Amperes	Output Volts	Output Milliamperes	Approximate over-all dimensions, inches	Weight, pounds
E1W272*	6	4.7	250	50	6 × 5 × 5	7½
E1W339	6	7.5	250	100	6 × 5 × 5	7½
E3W413	6	15	500	100	7 × 5 × 5	11
RA1W549	6	25	750	125	9 × 6 × 6	17¼
RA3W534	12	32	1,000	250	10 × 6 × 6	23¼
A420†	6	23.5	400	200	All 4½ × 5 × 8½ in., weight 13½ lb. Ratings are for continuous duty, will carry 75 per cent overload on intermittent use	
B420	12	12	400	200		
B1150	12	19.8	1,000	150		
C420	32	4	400	200		
C1150	32	8.3	1,000	150		
B415	12	24	400	450	All 4½ × 5 × 10½ in., weight 18½ lb.	
C1250	32	14	1,000	250		
D1250	110	4.2	1,000	250		

* Pioneer Gen-E Motor Corp., Chicago. Other ratings are available.

† Carter Motor Co., Chicago, Ill. Other ratings are available.

0.52 lb. per watt capacity. Heavy-duty devices usually run at lower speeds and are available in ranges up to the highest outputs, varying in price from $0.16 to $1.50 per watt and weighing from 0.14 to 1.00 lb. per watt, the lesser weights and lower prices applying to higher output devices.

Similarly for the motor-generator sets listed in Tables XII and XIII, the high-speed devices range in price from 9 to 27 cts. per watt and weigh from 0.04 to 0.7 lb. per watt. The heavy-duty and high-capacity units vary from 25 to 30 cts. per watt and weigh from 0.16 to 0.40 lb. per watt.

WIND-DRIVEN CHARGERS

39. Data on Wind Conditions in the United States. Properly installed, a wind-operated generator may be used, as an auxiliary means of charging a storage battery, over a larger portion of the United States than is generally supposed. The Weather Bureau, under the U. S. Department of Agriculture in Washington, maintains various stations throughout the country, and data on the average velocity of the wind, prevailing direction, and the date and maximum velocity of wind are published yearly in several pamphlets, entitled the Annual Meteorological Summary, each of which gives data for one of the various stations.

Most wind-driven chargers commercially available are characterized by a two- or three-bladed propeller, some type of governor preventing speeds rising much above those encountered in a 20-mile per hour wind, and a d-c generator either driven "direct" at the speed of the propeller or by a "step-up" gear drive. The governors are of several types. The centrifugal "air-flap" device comprises two vanes, held concentric with the propeller axle at low speeds by springs; but at high speeds the flaps become radial, acting as an air brake.[1]

[1] Wincharger Corp., Sioux City, Iowa.

A dual tail vane is also used, to change the relation of the propeller to the wind direction. Here the propeller axle is always parallel with the ground.[1] A third method is the "tilt back," in which the propeller axle is moved toward a vertical position thus decreasing the effect of the wind.[1]

A two-bladed propeller permits ready adjustment in the field of the "tracking" of the blades and the governor is essential to prevent damage in high winds. The propellers can be of wood, with either stainless-steel or copper-lined leading edges, or constructed entirely of metal. The generator is usually of the three-brush constant-current variety, giving

FIG. 11.—Average hourly velocity of the wind for an elevation of 100 ft. (*Courtesy of Weather Bureau, U. S. Department of Agriculture.*)

maximum output at low speeds; in conjunction with the governor it keeps a reasonably stable current output at high wind speeds (see Fig. 12).

In general the approximate hour of the greatest wind movement in the United States is at 3 P.M. local standard time, and over considerably more than half the country this average is above 12 m.p.h. at an elevation of 100 ft. For New York City the minimum hourly wind velocity is 8 m.p.h. in July and August, at which time the lowest of the maximum hourly velocities is 10 m.p.h.

The attractiveness of the absence of fuel costs and the maintenance of the accessories required by a fuel-driven engine are to some extent offset by the uncertainty of the supply of wind. In any case the primary

[1] LeJay Mfg. Co., Minneapolis, Minn.

need is for a suitably high location of the generator and propeller. Local conditions vary so widely that only general suggestions can be made. The installation should be high enough to secure the maximum effect of the wind and should therefore be away from, or higher than, obstructions such as trees and buildings. One instruction book specifies that

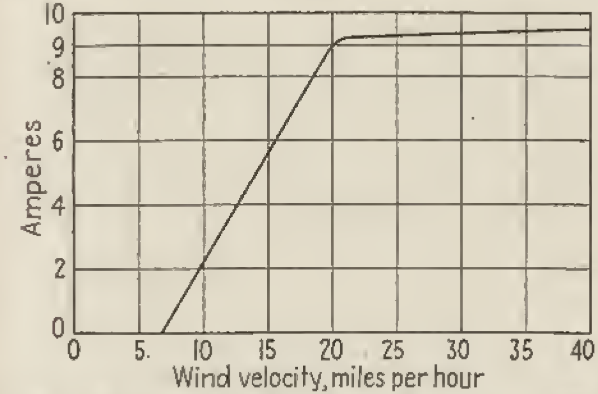


FIG. 12.—Characteristics, 110-volt, 1,200-watt direct-drive Wincharger generator, 11¼-ft. propeller.

the wind charger should be 15 ft. above any obstruction more than 400 ft away.[1]

40. Reliability. The amount of wind suitable as motive power can hardly be guaranteed for any location. If the preliminary survey of local wind conditions indicates a reasonable average wind velocity, the observance of proper installation procedure will generally produce satisfactory results. In general, the maximum winds occur in the spring and the minimum in midsummer, varying, of course, with the section of the country. The steadiness of output from a wind charger cannot be compared with that, for example, of a fuel-driven engine. The wind charger operates at variable output unless the wind velocity varies above a certain minimum speed.

For reliable results the wind-driven generator should begin charging the battery at the lowest possible wind speeds, since in midsummer the hourly wind velocities in many localities do not reach very high values. Here the use of the word "auxiliary generator" becomes of importance. A wind generator can be relied upon for a great deal of the time, but an auxiliary source of power such as a fuel-driven generator is recommended where absolute continuity of power supply is required.

A one-year curve[2] of the performance of a 1,200-watt wind-driven generator, in Sioux City, Iowa, shows a cyclic output, ranging from a maximum of 215 kw-hr. for April to a minimum of 120 kw-hr. for August.

[1] Montgomery Ward & Co.
[2] Wincharger Corp.

For four wind-charger plants used on farms in Minnesota, all having over 1,000 watts capacity, the average monthly consumption was 60 kw-hr. and the lowest consumption was 18 kw-hr. in August.[1]

The wind chargers of Table XIV develop sufficient voltage to start charging at 5½- to 7½-m.p.h. wind velocities, and their outputs are substantially constant and at their maxima if the wind speed is above 20 m.p.h.

41. Direct Current Only. Since the speed of a wind charger will vary, the generators are made to furnish d.c. only. A storage battery is invariably required to store energy for use when there is no wind. As a general rule, a two-week "calm" is the maximum, which requires battery capacity to carry the load for that length of time.

Some wind chargers are mounted on towers above roofs, and in this case mention should be made of the necessity of a suitable means for absorbing the vibration, such as the use of rubber pads. Special towers or poles for mounting the chargers are frequently employed, in which case vibration prevention is of little importance, rigid mounting being preferred.

42. Length of D-c Leads. For low-voltage chargers the leads to the battery have to be low in resistance. It is generally advisable to have less than 200 ft. between the battery and the generator and preferably less than 50 ft. As a rule the voltage drop at full-current rating should be less than 20 per cent of the nominal voltage rating. For example, a 6-volt charger, 200 ft. away, with a capacity of 120 watts will furnish 20 amp. maximum. If the IR drop is to be 20 per cent total, each lead wire has 10 per cent of 6 volts or 0.6 volt. Hence the allowable resistance $R = E/I = 0.6/20 = 0.03$ ohm, which is the maximum resistance for each of the two conductors, *i.e.*, the wires must be No. 2 B. & S. gage or larger. Thus two copper wires each 200 ft. long, having a minimum size

TABLE XIV. WIND-DRIVEN GENERATORS*

Name	Diameter, feet, two-bladed propeller	Rated output in 20 m.p.h. wind		Rated 10 m.p.h., amperes	Minimum m.p.h. to charge	Approximate generator r.p.m., 20 m.p.h.	Pound weight less tower	Type of generator drive
		Watts	Amperes					
6 volt†	6	120	17‡	4	7½	1,100	58	Direct
Heavy duty 6 volt	7½	200	25	8	5¼	800	96	Direct
Heavy duty:								
12 volt	7½	225	14	5	6	800	96	Direct
32 volt	8	600	15	4	7¼	800	140	Direct
32 volt	11	1,200	30	6	6½	1,600	270	Gear 4:1
Streamliner:								
110 volt	11¼	1,200	9	2.3	6½	400	681	Direct

* From data furnished by the Wincharger Corp., Sioux City, Iowa.

† The volt ratings are nominal, a 6-volt charger will charge a three-cell (6-volt) lead-acid battery, a 110-volt battery, etc.

‡ The third brush on the generator, in conjunction with the wind governor, keeps the current substantially at this value for wind speeds above 20 m.p.h.

[1] University of Minnesota, *Agr. Eng. News Letter*, No. 23, February, 1934.

of No. 2 B. & S. are required. For higher voltages correspondingly smaller wires may be used for a given watt rating. In any case short leads decrease the first cost by requiring smaller cables, but the IR drop should be lower than 20 per cent if possible, preferably 5 per cent, or 0.12 volt on a 6-volt system.

43. Capacities and Types. Wind-driven generators are available for outputs of 6, 12, 32, and 110 volts. Standard gear-driven models are available, allowing the use of smaller sized, higher speed generators, as well as direct-driven generators which have the same speed as the propeller.

In addition to, or in conjunction with, the governors, brakes and other devices for stopping the propeller are provided. Some of these turn the tail vane at 90 deg. from its operating position, so that wind does not affect the propeller; others are of conventional brake-shoe construction.

44. Radio Interference. Since the generator is high up in the air, it is essential to prevent radio waves from emanating, and the commutator ripple is usually minimized by built-in condensers. In addition, however, it is suggested that the metal parts of the tower be solidly grounded at the base and that one of the lead-in wires to the battery also be grounded.

45. Maintenance and Depreciation. Generators are permanently lubricated, but provision is made for removing a cover to lubricate the collector rings, about which the propeller-generator unit moves as an axis. The small number of moving parts reduces maintenance to a minimum. On the larger models the design is arranged so that the units are individually assembled. For example, the governor can be removed without disturbing the balance of the propeller, and the collector ring can be removed without disturbing the generator assembly. With reasonable care, depreciation ranges from 5 to 20 years and thus can be figured at from 20 to 5 per cent per year.

RECTIFIERS AND CHARGERS

The general types of rectifiers which are mainly used for supplying d.c. for radio power from various a-c supplies are as follows:

Vacuum tubes, with filament or indirectly heated cathodes.
Gas-filled tubes with filament cathodes.
Igniter-type gas-filled tubes.
Dry-contact metal rectifiers.

There are two general types of rectifier tubes, vacuum and gas filled. Both types have certain maximum inverse voltage ratings, *i.e.*, the largest safe voltage which can be applied to the tube in the non-rectifying direction. Both types have maximum current ratings, either in r.m.s. or peak values, or both. Indirect heated vacuum rectifiers are used generally on voltage doublers and have a maximum safe voltage which can be applied between heater and cathode.

Hot-cathode mercury-vapor rectifiers usually require cathodes to be heated for periods up to 1 min. before plate voltage is connected. These tubes must be used in circuits which limit the current, since mercury-vapor rectifiers operate at practically constant voltage drop and thus would be ruined by too high a current demand.

46. Characteristics of Rectifiers for Receivers. Figure 13 gives average plate characteristics for receiver-type rectifiers. The vacuum rectifiers have a varying voltage drop (vertical coordinate) with current

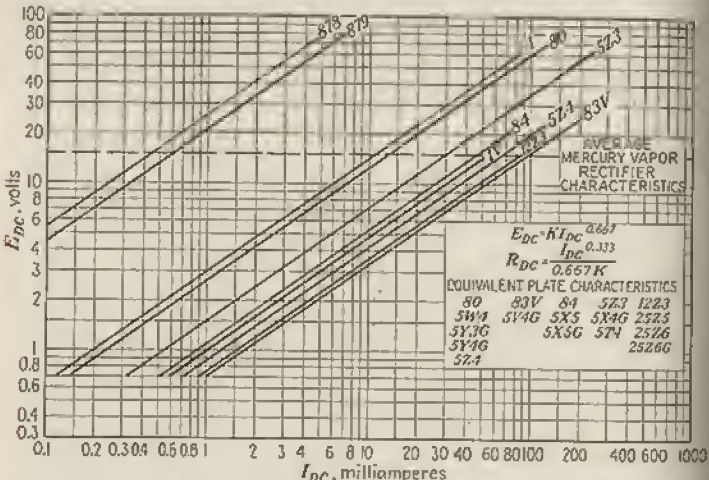


Fig. 13.—Average plate characteristics of receiver-type rectifiers.

drain (horizontal coordinate). The equivalent types of tubes are shown in the figure, *e.g.*, 80, 5W4, 5Y3G, 5Y4G, 5Z4, each of which is represented by the line labeled "80." On this figure the average mercury-

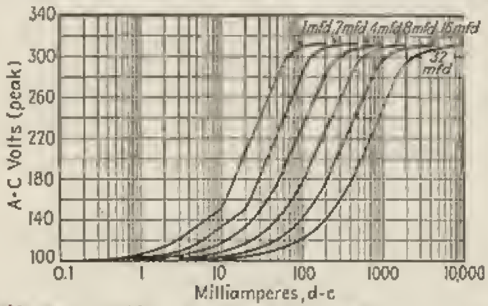


Fig. 14.—Half-wave rectifier characteristics, useful for receiver circuits.

vapor rectifier drop is indicated by a horizontal dot-and-dash line[1] at 15 volts, and does not vary with current drain.

Half-wave rectifiers, typified by the high-voltage circuit of Fig. 3, have output characteristics shown by Fig. 14, and the peak inverse

[1] Figures 13, 14, 15, and 17 are reprinted by courtesy of Aerovox Corp., from the *Aerovox Research Worker*, August-September, 1937.

voltage is 2.83 times the transformer voltage. A full-wave rectifier, also shown in Fig. 3 has output characteristics shown by Fig. 15, the peak inverse voltage is 2.83 times half the secondary transformer voltage.

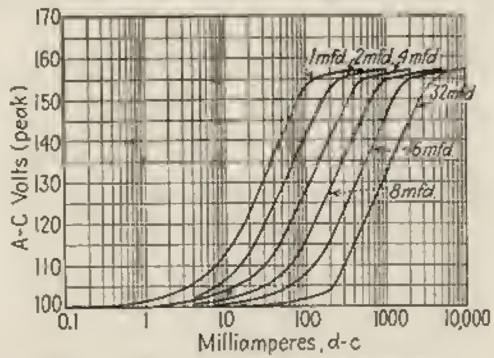


Fig. 15.—Full-wave rectifier characteristics.

47. Voltage-doubler Circuits. A typical voltage-doubler circuit shown by Fig. 16 has output characteristics shown by Fig. 17. For receiver circuits there are several tubes comprising two rectifier elements in one envelope, among which are 25Z5, 25Z6, 25Z6G, 25Z6GT.

If electrolytic condensers are used in voltage-doubler circuits, the positive lead of one condenser must be connected to the negative of the other. *It is not possible to use a dual electrolytic condenser for voltage doubling if the negative lead is common.*

Voltage-doubler tubes are often used in a-c receivers with both plates and both cathodes connected in parallel. This connection does not permit voltage doubling, reduces the internal drop below that of a single element tube and increases the rectifier output from a low voltage a-c supply line. On d.c. the tube acts as a resistance and also forces the user to plug in the set with the proper polarity.

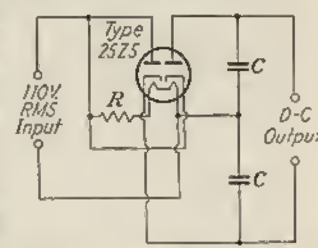


Fig. 16.—Voltage-doubler circuit.

As a voltage doubler the two half-wave rectifiers operate on consecutive half cycles of input power to charge their respective condensers. The load is connected across the two condensers in series. The ripple frequency is twice the line frequency. The maximum inverse voltage applied is 2.83 times the applied a.c.

48. Vacuum-tube Rectifier with Filament Cathode. This type of rectifier is used in nearly all a-c powered radio receivers and has numerous applications in higher powered circuits. Oscillograms[1] showing the

[1] Wise, Roger, *Radio Broadcast*, April, 1929, pp. 394–395.

effect of different load circuits are given in Figs. 18 and 19. In both figures the letters *a* to *e* refer to similar load circuits, *a* being a simple resistor load, *b* a 4-μf condenser across the resistance, *c* a 20-henry choke in series with the resistor, *d* a standard three-condenser, two-choke filter with load resistance, *e* the same as *d* with the first condenser omitted.

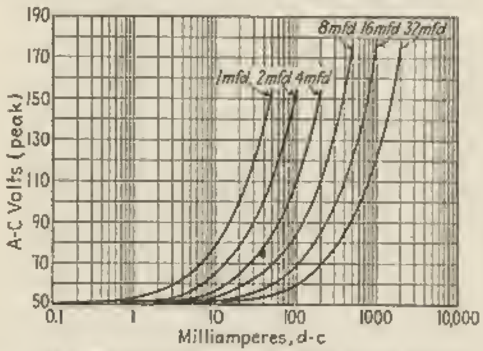


FIG. 17.—Voltage-doubler characteristics for receiver circuits.

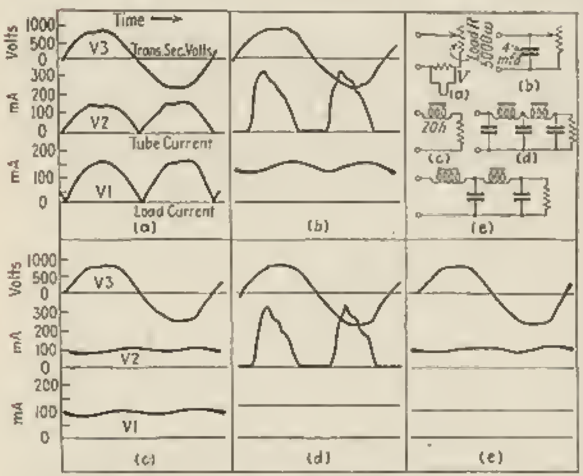


FIG. 18.—Half-wave rectifier, different load circuits.

For each load three factors are shown, the *V* letters denoting the oscillograph vibrators, the transformer secondary voltage being V_3, the tube current V_2, and the load current V_1. The curves of special interest are those of *d* and *e* in Figs. 18, and 19. In both figures *d* shows a severe load current being drawn from the rectifier tube, the peak current from the half-wave

tube being 540 ma. and the output current 102 ma, a ratio of 5.3:1. The full-wave tube peak current is 290 ma, while the output current is 118 ma, a ratio of 2.5:1. For the *e* section of these two figures the half-wave tube peak current is 130 ma and the load 45 ma, a ratio of 2.9:1; while the full-wave

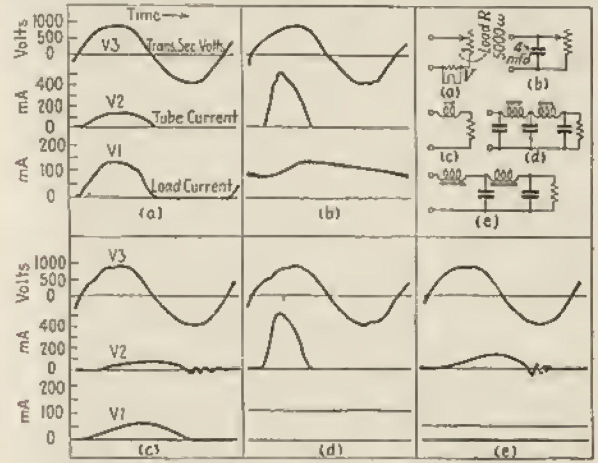


FIG. 19.—Full-wave rectifier, different load circuits.

peak current is 110 ma and the load 96 ma, a ratio of 1.5:1. In all these curves the power transformer was the same, and an idea of the relative output voltages and currents can be secured by comparing the desired circuits of Figs. 18 and 19.

From the standpoint of the rectifier tube, these figures show that the omission of the first filter condenser will decrease the high periodic loads which are required by the standard filter having an input condenser. By referring to Fig. 34 it will be seen that the omission of C_1 decreases the available voltage, and this is verified by the curves in Figs. 18 and 19, as the same transformer supplied the voltages to both *d* and *e* circuits in turn.

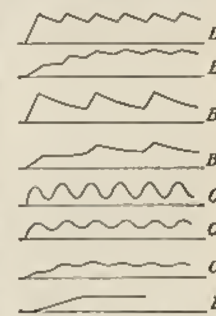


FIG. 20.—Load currents for several forms of filter.

Figure 20[1] gives the load current, through several cycles, for several forms of filter. The letters are made the same as for Figs. 18 and 19 wherever possible. Curve *B* of Fig. 20 corresponds to the *b* curve of the full-wave rectifier of Fig. 19 while *B'* is the same as *b* with the condenser capacity approximately six times as large. *B''* is the same as *B'*, for a half-wave rectifier, and *B'''* has about six times as much capacity as *B''* but is otherwise the same. Curve *C* corresponds to the regular *c* of the former figures, and *C'* is the same as *c* with 2.13 μf across the rectifier side of the choke. Curve *C''* is like *C'* with the condenser increased to nearly six times its original value. Curve *D* resembles the *d* of the former figures, except that it comprises only one filter section instead of two as in Fig. 19.

[1] KUHLMAN and BARTON, *Jour. A.I.E.E.*, January, 1928, p. 17.

49. Hot-cathode Mercury-vapor Rectifier. The hot-cathode mercury-vapor rectifier[1] differs from the mercury-arc tube in two respects: (1) It operates at a relatively low temperature, so that the vapor pressure is low. This low mercury pressure gives a useful characteristic, a high breakdown voltage in the inverse direction. (2) The electrons are emitted from the filament and not from a pool of mercury. In the second respect this tube resembles the vacuum-tube rectifier, but the difference lies in the much lower potential drop due to the neutralizing of the filament space charge by the positively charged mercury ions.

The filament-to-plate drop of the mercury-vapor tube is about 15 volts and is practically independent of the load current. This low drop improves regulation and increases the available d-c output. This tube is self-igniting and does not require the starting mechanism of the mercury-arc rectifier.

50. Battery Chargers. For low-voltage high-current rectification the argon-filled, tungsten-filament Rectigon and Tungar bulbs fill the need. The use is largely that of charging storage batteries, and no filter is needed for this application.

Filters have been designed for use with these rectifiers, so that the output can be fed directly to the filaments of d-c radio tubes. To design a proper low-pass filter for d-c tube filament currents, Eqs. (1) and (2) Art. 62 should be used, as the chart of Fig. 30 does not cover this range. The condenser has to have a large capacity, and low-voltage dry electrolytic condensers are often used. In using these condensers it is important to connect the correct polarity to the rectifier.

TABLE XV. BATTERY-CHARGER TUBE CHARACTERISTICS
(All half-wave single phase)

Number*	Filament		Maximum d-c anode ratings		Maximum inverse voltage	Over-all length, inches
	Volts	Amperes	Volts	Amperes		
289415	2	12	75	2	275	5¾
289416	2.2	18	90	6	375	6¾
766776	2.5	27	60	15	225	8⅝

* Style number of Westinghouse Electric & Manufacturing Co., Rectigon tubes. General Electric manufactures similar tubes under the name Tungar.

51. Mercury-arc Rectifiers. Formerly, mercury-arc rectifiers were most used[2] in the field lying between the argon tube and the filament-vacuum rectifier. With the introduction of the mercury-vapor tubes many of the advantages of the mercury-arc rectifier—low voltage drop, high efficiency—were duplicated. The mercury-arc tube requires a starting electrode, and usually a mechanical tilting device for starting.

52. Igniter-type Mercury Rectifier. A mercury-pool cathode provided with a thyratron igniter is called the *ignitron*.[3] This rectifier

[1] PIKE and MASER, QST, February, 1929, p. 20.
[2] PRINCE and VOGDES, "Principles of Mercury Arc Rectifiers and Their Circuits," p. 23, McGraw-Hill Book Company, Inc., New York, 1927.
[3] Westinghouse Electric & Manufacturing Co., East Pittsburgh, Pa.

is characterized by an efficiency of about 90 per cent from 10 to 125 per cent rated load for a typical 300-kw 275-volt d-c output from 2,300 volts, three phase, 60 cycles. This is a higher efficiency over a wider range of loads than is possible from a synchronous converter or a synchronous motor-generator set.

The high vacuum required is maintained by a mercury-vapor pump plus a rotary pump. The latter, automatically controlled, is used only a short time each day.

Ignitrons are particularly useful for supplying high-current direct voltage below 600.

53. Dry-contact Rectifiers.[1] At present the dry rectifier field in the United States is divided among three different types of rectifiers, the copper sulphide,[2] the copper oxide,[3] and the selenium.[4]

54. Copper Sulphide Rectifiers. The copper sulphide rectifier is assembled from disks of copper sulphide and magnesium, with or without radiating fins, mounted on bolts and clamped together under high pressure. The rectifier is characterized by small size and weight, by ability to operate with a high temperature rise (100°C.), by its low initial cost, definitely limited life, good voltage regulation, and poor efficiency. The thermal capacity of the element is relatively low. There is also some difficulty in operating units in parallel, generally requiring separate transformer secondaries for each rectifying element. There is also a limitation as to the number of elements that can be operated in series, so that the rectifier is found only in the low-voltage fields. An additional limitation exists in the range of sizes of the rectifying elements due to the high pressure required. They cannot be used in the small-power classification, *e.g.*, in instruments.

The proper field of application for the copper sulphide rectifier is for intermittent duty, where the definitely limited life can be stretched out over a satisfactorily long time and where in addition the initial cost is important. Among such applications are power units for operation of various types of electromagnetic loads, such as circuit-breaker solenoids, and in intermittently operated battery chargers such as used around the home. The low efficiency and definite life limitations appear to bar this rectifier from fields where long life and good efficiency are paramount.

55. Copper Oxide Rectifiers. The copper oxide rectifier is assembled from oxidized copper disks with lead washers, with or without radiating fins, clamped on a bolt under high pressure, or from large area low-thermal-capacity plates furnished with sprayed or plated collecting surfaces, which can be assembled under little or no pressure at all. This rectifier is characterized by its large size and weight, its limitation to low temperature rises in operation (15°C.), good efficiency, poor voltage-regulating characteristics, indefinitely long life, high thermal capacity, and high initial cost. The rectifier is flexible as to size of element of any dry rectifier, elements in production today ranging from $\frac{1}{32}$-in. diameter to 50 sq. in. in area. Figure 21 shows a typical *E-I* curve for a 1½-in. disk.

[1] MAIER, K., "Trockengleichrichter," Oldenbourg, Berlin, 1938.
[2] Manufactured by the B-L Corp., St. Louis, Mo., and by the P. R. Mallory Co., Indianapolis, Ind.
[3] Manufactured by Westinghouse Electric & Manufacturing Co. and General Electric Co.
[4] Manufactured by the International Telephone Development Corp., New York.

Applications for this rectifier are almost unlimited. It is best fitted, of course, for those purposes which require long life and fairly good efficiency and where the matter of first cost is less important than is the operating cost. Because of its volt-ampere characteristics this rectifier can be used for almost any application requiring d.c., provided it can be economically justified. Its high thermal capacity, particularly in the disk type, and the ability of the elements to withstand high voltages for short periods of time have made the rectifier very useful in intermittently loaded applications, such as the operation of circuit-breaker solenoids.

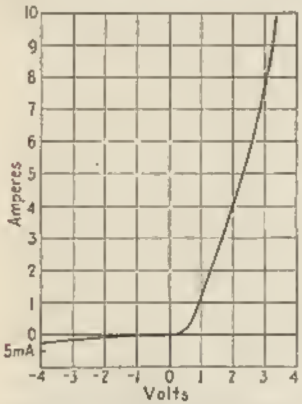


FIG. 21.—Typical Rectox current-voltage curve.

56. Selenium Rectifiers. The selenium rectifier is assembled from plated iron disks coated with selenium and sprayed with metal. High assembly pressure is not required for this rectifier because each disk is an integral rectifying unit with a low-resistance contact surface on either side. Assemblies are available without radiating fins and also with or without increased disk spacing. This rectifier is noted for its small size and weight compared to power-handling ability.

The efficiency is comparable with copper oxide, and the permissible operating temperature is higher. Voltage overloads are not permissible, but the rated voltage per disk is higher than copper oxide and short-period current overloads are permissible. Fields of application and flexibility are roughly the same for selenium rectifiers as for copper oxide except that selenium rectifiers are not available for very low capacity uses, such as indicating instruments.

Information as to the life of this type of rectifier comes from Europe where it has been in use for about 11 years. Long-term service tests have not been completed in this country since the American product has been available for less than 2 years.

57. General. From the above brief descriptions, it may be seen that copper oxide and selenium are nearly alike on the basis of performance with selenium having the advantage as to size and weight, and copper oxide having the advantage in range in element sizes and use on instrument applications.

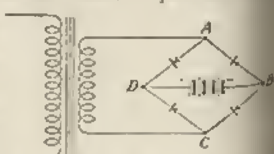


FIG. 22.—Bridge circuit for Rectox rectifier.

Neither of these rectifiers is competitive with the copper sulphide rectifier in those fields where efficiency is unimportant, operation is intermittent, and first cost is significant. Copper sulphide rectifiers reach into fields, which are also supplied by copper oxide and selenium, such as railway battery charging and electroplating.

A typical rectifier circuit employing a transformer without a center tap is shown in Fig. 22. When A is positive, the path of the current is $ADBC$; when C is positive, the current path is $CDBA$. During the time that A is positive, nearly the entire transformer secondary voltage is applied across the bridge arms DC and AB.

LOW-POWER TRANSFORMERS FOR RECEIVERS

58. Transformers. The design of a reliable power transformer, having high efficiency, requires fairly elaborate calculations. To take into account the d.c. which flows in a transformer secondary when a half-wave rectifier is used, some interesting equations have been derived.[1]

A simple approximate-design method will be given here, for the construction of single-phase low-powered transformers up to 180 volt-amp., or 180 watts for approximately unity power factors. This design is especially suited to transformers which supply a full-wave rectifier and filament energy to an a-c powered radio receiver, three factors making it possible to secure a satisfactory transformer without complicated design methods. These factors are as follows:

1. There is no urgent need for high efficiency. An 80 per cent efficient transformer which takes 60 watts to supply 48 output watts is fairly satisfactory if it can radiate the heat which it generates.
2. These transformers are operated at a fairly constant load. This improves the maintenance of the various output voltages as each secondary winding will have a constant IR drop.
3. The load on the transformer secondary is nearly of unity power factor. The filament power load is essentially a resistance load with unity power factor. The current supplied to the filter has slightly less than unity power factor, but this can be disregarded in low-powered transformers. The indirect heated receiving tubes, such as the 56, require less than half as much d-c power in their plate and grid circuits as that which is needed to heat their cathodes. This would mean a unity power-factor heater supply and (assuming a series voltage divider) less than half as many additional watts for plate and grid supply at a lower power factor. It is true that a power tube, such as 6L6 at its maximum rating, uses slightly over four times the wattage in its $B + C$ circuit as it does in its filament. It is rare, however, to have more than two power tubes in a receiver, and the assumption that the power factor of the secondary is unity is usually not over 20 per cent off. This means that the wire of the high-voltage secondary and of the primary should be increased to allow for this added current.

59. Small Transformer Details. Economy in a transformer is secured when the winding encloses a maximum of core area with a minimum of wire, and the magnetic path should be as short as possible.

The core form of a small transformer can be of several shapes, but it is usual to use standard punchings shaped like capital letter E's. As a rule, two punchings are used, one having longer legs than the other so that the magnetic circuit "breaks joints" in stacking the iron. Another convention usually followed in small transformers is the use of a single-winding form, all secondaries and primary being on the middle leg of the E core.

The spool form is usually an insulating tube, and side pieces may be fitted, on which terminals are placed; or, if the coil is to be machine-wound with interwoven cotton, the side pieces can be omitted, and flexible leads provided.

[1] HARDER, E. I., *Elec. Jour.*, October, 1930, p. 601.

60. Ten Steps in Designing a Small Power Transformer. 1. *Determine the Volts and Amperes Needed for Each Secondary.*

a. Find the total maximum secondary watts $= W_s = E_1I_1 + E_2I_2 + \cdots$

b. Find the total watts needed for primary $= W_p$

Assuming 90 per cent efficiency $W_p = W_s/0.9$

c. Find primary amperes assuming 90 per cent power factor

$$I_p = \frac{W_p}{E_p \times 0.9} = \frac{W_s}{0.81E_p}$$

and for $E_p = 110$ volts, $I_p = W_s/89.1$ amp.

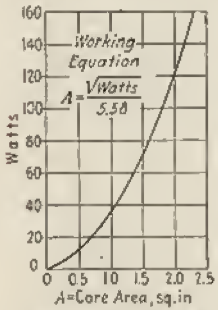


FIG. 23.—Small power transformer core area as a function of watts.

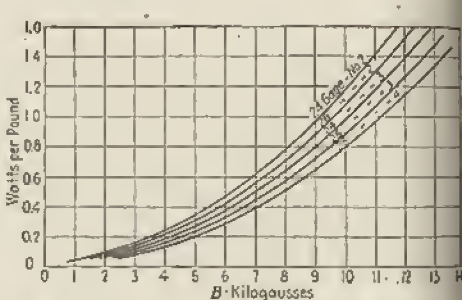


FIG. 24.—Core-loss curves Armco Radio grades (60 cycles).

2. *Size of Wire.* Knowing the current for each winding, the wire size is determined by the circular mils per ampere which it is desired to use. A safe rule is to use 1,000 cir. mils per ampere for transformers under 50 watts and 1,500 cir. mils per ampere for higher powers.

3. *Core Considerations.* A curve showing core areas for different powers is Fig. 23 which shows the area for 40 watts to be 1 sq. in.; 70 watts, 1.5 sq. in.; and 120 watts, 2 sq. in. The area of the core is the same as the inside dimensions of the spool, making a 10 per cent allowance for stacking; for example, a spool 1 by 2 in. inside would enclose 2 sq. in., but, allowing for a 10 per cent loss, only 90 per cent or $0.9 \times 2 = 1.8$ sq. in. is the net core area. The core area is needed to determine the turns per volt.

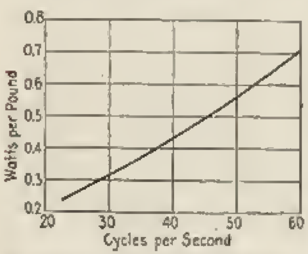


FIG. 25.—Core loss *versus* frequency $B = 10{,}000$.

4. *Core Loss and Induction.* The flux density at which the core is to be worked determines the iron (core) loss. Figure 24 gives several curves of different core materials, watts per pound being plotted against flux densities in kilolines per square inch. Sixty-five kilolines per square inch is an average value of the induction. The making of a curve such as Fig. 24 depends largely on experimental data, not directly on a theoretical basis. For this reason, no definite value of the core loss can be given; it depends on the quality of core material which is available. It should be noted that better and better

core material is constantly being made, having lower loss per pound, so that the use of higher flux densities is becoming possible. Up to 15 kilolines is not uncommon, but unusual for this application. The core loss increases with frequency, a typical curve being Fig. 25.

5. *Induced-voltage Equation, Turns per Volt.* The elementary definition, that 10^8 magnetic lines cut per second will induce one volt pressure, is the basis of the equation

$$E = \frac{BANf}{10^8} \times 4.44$$

where E = voltage
A = area of the core
B = flux density in the same units as A
f = cycles per second
N = number of turns.

A more useful working equation for small power transformers is obtained by solving for N/E in turns per volt:

$$\frac{N}{E} = \frac{10^8}{BAf4.44}$$

Figure 26 is an alignment chart of this equation. The left column is the flux density B, in both kilolines per square inch and kilogausses (kilolines per square centimeter); the center column is the net core area in both square inches and square centimeters; the right column giving the turns per volt for both 25 and 60 cycles per second.

Using a flux density of 65 kilolines per square inch and the net core area mentioned in step 3 (1.8 sq. in.), the turns per volt for 60 cycles are found to be 3.1 turns per volt. Thus, for each volt on the transformer, there must be 3.1 turns. It is customary to change the turns per volt to an even number so that center taps can be provided. In this case, by using 4 turns per volt, with the same core area, the induction will be lower, with a corresponding lower core loss. It is also quite possible, and sometimes advisable, to change the core area so that an even number of turns per volt is given. For example, by increasing the core area to 2.8 sq. in., 2 turns per volt could be used; by decreasing to 1.4 sq. in., 4 turns per volt would be used. The reason for desiring the even numbers of turns per volt is to supply the ½-volt steps for receiving tubes, such as 6.3 volts, which would require an integral number of turns when the turns per volt are used.

The voltage drop in the transformer winding should be mentioned here, and it will be again taken up in detail in the example. For instance, the load voltage at a tube filament is lower than the no-load voltage by the amount of IR drop in the winding and the connecting wires to the tube. Thus it may be that to secure 6.3 volts at the tube filament, the transformer no-load voltage will have to be 7. In this case any integral number of turns per volt, either odd or even, will suit the design.

6. *Turns for Each Winding.* In step 1 the desired voltages were given, E_1, E_2, etc. Using the value of turns per volt in step 5, the total turns for each winding are found. For example, with 4 turns per volt, a 110-volt winding should have $4 \times 110 = 440$ turns.

7. *Winding Space Required.* From the total turns for each winding, and the wire size, the total area of winding space is calculated. Different wires and insulations have definite turns per square inch. The method of insulation, however, may have these values vary by factors of as much as three to one. That is, a 900-turn coil wound in layers with enamel wire may take up 1 sq. in. of cross-section area. By interleaving thin insulating paper between layers, only 600 turns can be wound on a square-inch area; and by using a certain size of cotton interwoven between turns, only 400 turns can be wound in a square inch. Thus the space of winding depends to a large degree on the kind and thickness of insulation. Double cotton-covered wire takes up

considerably more space than enameled wire. Yet, if the extra-needed insulating space for the interlayer protection is considered, the space ratio may not be so great.

After adding up the winding space of all the windings, the area should be compared with that of the core. If the winding will go in the core space, this part of the design is finished.

If the wires will not go in the available space, the winding may be redesigned, or the core area increased. Using thinner coverings for wire, fewer secondaries or fewer circular mils per ampere will decrease the space needed for the wire. A larger iron size or a thicker stack of the same sized iron will

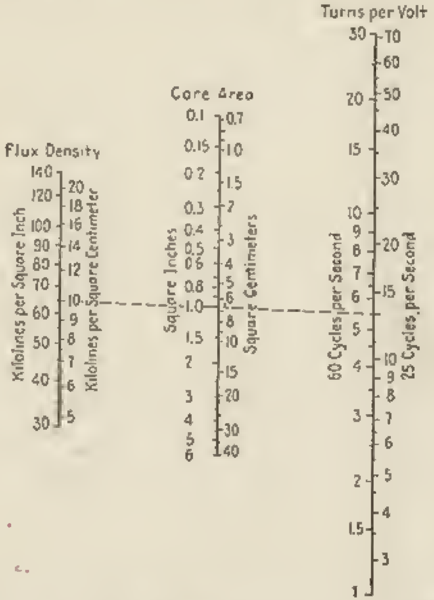


FIG. 26.—Transformer design chart based on $E = \dfrac{BANf \times 4.44}{10^8}$.

increase the core area and allow a smaller number of turns per volt, thus decreasing the cross section of the winding.

8. *Copper Loss.* *a.* Find the length of the mean (average) turn in feet.

b. Find the length of each winding in feet by multiplying the number of turns by the mean turn length.

c. From wire tables find the ohms per 1,000 ft. for the size wire used, and then from 8-*b* the actual ohms for this length.

d. Multiply the current squared for each winding by the ohms for that winding.

e. Add the I^2R's for each winding to get the copper loss L_1.

9. *Core Loss.* The core loss in watts L_2 is found from the weight of the core and flux density and kind of core used in step 4. A useful factor is that 4 per cent silicon steel weighs 0.27 lb. per cubic inch.

10. *The approximate percentage efficiency is* $\dfrac{W_s \times 100}{W_s + L_1 + L_2}$, W_s being the secondary watts (see step 1).

NOTE. If step 10 shows about 90 per cent efficiency, the design is complete. If much loss than 90 per cent, step 1*a* must be modified, a new larger value of I_p being used in finding a larger primary wire. This will not change the efficiency but will prevent overloading the primary winding due to its carrying a greater current than that for which it was designed.

It is desirable, as a rule, to keep the efficiency above 90 per cent, and this can be done by reducing L_1 and L_2 by using larger wires or larger cores.

61. Typical Small Transformer Design. This transformer gives a full-wave rectifier supply, filament supply for rectifier and receiver, and works on a primary voltage of 110, at a frequency of 60 cycles.

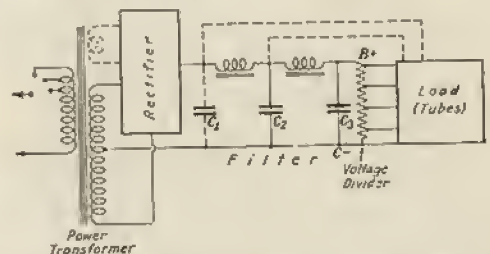


FIG. 27.—Typical a-c powered unit for furnishing B and C voltages.

1. The desired secondary voltages and currents are as follows:

E, volts	I, amperes	Use	Watts = EI
330	0.075	B and C supply	24.75
330	0.075	B and C supply	24.75
5.0	2.0	Rectifier filament	10.0
6.3	1.8	Filament	11.34
2.5	3.0	Filament	7.5

a. Total secondary watts W_s 78.34
b. Primary watts $W_p = W_s/0.9 = 78.34/0.9 = 87.0$
c. Primary amperes $I_p = W_s/89.1 = 78.34/89.1 = 0.88$

2. This transformer is over 50 watts, so 1,500 cir. mils per ampere is the current density to use in finding the proper-sized wire. The wire sizes, with

Volts	Amperes	Size wire
110	.88	18
330	0.075	29
330	0.075	29
5	2.0	14
6.3	1.8	15
1.5	3.0	12

the identifying current and voltages, are listed in the foregoing table. The use of larger wires of even numbers keeps the IR drop lower than when using a smaller wire. However, if the use of these larger wires makes too large a winding cross section, smaller wires must be used.

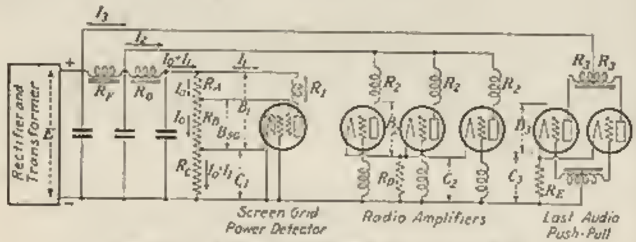


FIG. 28.—Voltage divider with graded filter.

3. The core area available is 1⅜ × 2 in., the net area being 1⅜ × 2.0 × 0.9 = 2.48 in. This is larger than necessary as shown by Fig. 26, but allows the design, in this case, of a transformer with good efficiency and good regulation.

4. The flux density used is 65 kilolines per square inch and 4 per cent silicon iron with a loss of 0.6 watt per pound.

5. The turns per volt for 65 kilolines per square inch and core area of 2.48 sq. in. give three turns per volt.

6. The turns for each winding are as follows:

Volts	Turns
110	330
330	990
330	990
5	15
6.3	18.9(20)*
1.5	4.5(5)

* It is usual to add ½ to 1 turn to filament windings to allow for the IR drop in the winding and leads to the tube filaments.

7. Winding space, in square inches, using enamel wire, follows:

Turns	Feet	Ohms per 1,000 ft.	Actual ohms	IR volts drop	I^2R watts
330	320	6.51	2.08	.....	1.61
990	906	83.4	75.6	.....	0.43
990	906	83.4	75.6	.....	0.43
15	13.7	2.6	0.035	0.07	0.14
20	18.3	3.25	0.064	0.115	0.37
5	4.6	1.8	0.008	0.024	0.07
Total.........	.....	.....	......	.....	3.05

a. The mean turn is 11 in. = 11/12 ft.

b. The space needed is 1.6 sq. in. and the space available is 1 × 2 = 2 sq. in., so the extra space can be used for the spool and for insulation between windings and layers.

Turns	Size wire	Turns per square inch	Actual space, square inch
330	18	400	0.82
990	29	3,500	0.28
990	29	3,500	0.28
15	14	175	0.09
20	15	220	0.09
5	12	120	0.04
Total..............	..	.....	1.60

c. The copper loss L_1 is 3.05 watts

8. The core weighs approximately 5 lb., which at 0.6 watt per pound gives $5 \times 0.6 = 3.0$ watts $= L_2$.

9. Watts output $= 78.34 = W$.

$$\text{Losses} = L_1 + L_2 = 3.0 + 3.05 = 6.05 \text{ watts}$$

$$\text{Per cent efficiency} = \frac{78.34 \times 100}{78.34 + 6.05} = \frac{7{,}834}{84.39} = 93 \text{ per cent}$$

NOTE. The copper losses are approximately the same as the iron loss, which is generally an indication of good design.

10. *Volts Drop.* It is seen by the IR column that the drop in the winding is not serious.

The core weighs 5 lb., and the copper winding is 2.8 lb.; allowing no weight for insulation, this transformer gives 78 watts output for 7.8 lb., or 0.1 lb. per watt of output.

FILTERS FOR SMOOTHING RECTIFIED AND GENERATED D.C.

62. Low-pass Filters. The filters used to give d.c. from rectified a.c. are known as low-pass filters.[1] Low-pass filters are divided into two

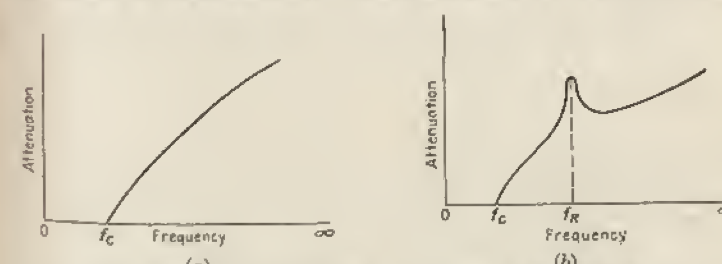


FIG. 29.—(a) Low-pass filter. (b) Tuned low-pass filter.

classes, tuned and untuned filters. The tuned filter offers a maximum impedance or attenuation to the frequency of the supply, but the impedance at near-by higher or lower frequencies is not quite so great (see

[1] The theory of filters is admirably covered in the following books: K. S. Johnson and T. S. Shea, "Transmission Circuits for Telephone Circuits"; G. W. Pierce, "Electric Waves and Oscillations."

Fig. 29b), although the general trend of the curve is a rising attenuation as the frequency increases.

The usual form of untuned low-pass filter is that of Figs. 27 and 28, using three condensers and two chokes. This filter (Fig. 29a) has a continuously rising curve of impedance as the frequency increases. To obtain good filtering with this filter, it is desirable to choose f_c, the frequency at which attenuation begins, as low as possible. The equations for determining the proper inductance and capacity for this filter are as follows:

$$C = \frac{1}{\pi f_c R} = \frac{0.3183}{f_c R} \text{ farads} \tag{1}$$

$$L = \frac{R}{\pi f_c} = \frac{0.3183 R}{f_c} \text{ henrys} \tag{2}$$

where f_c = frequency at which attenuation begins
C = capacity in farads
R = resistance in ohms
L = inductance in henrys.

As this is an often-used type of filter, Fig. 30 is devised to give the data of Eqs. (1) and (2) in a convenient chart form. The four columns from

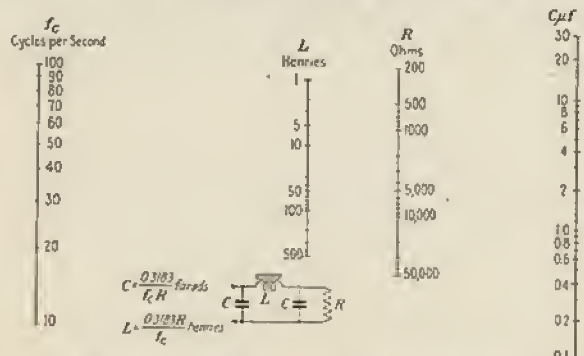


FIG. 30.—Low-pass filter design chart, π section.

left to right are f_c in cycles per second; L in henrys; R in load ohms; and C in microfarads. Thus with any two of the factors fixed, the corresponding two are determined from this chart by a straightedge across the two known factors. For use on 60-cycle half-wave rectification, it is necessary that f_c be below 60, and for the double-wave rectifier f_c should be below 120 cycles; the lower the f_c the better will be the filtering at the desired frequency, as shown by the rising attenuation curve of Fig. 29a.

The third column R is the usual starting place for finding the filter values when the voltage divider and tube load have been calculated first. When the point on the R column is fixed and f_c is, say, 50 cycles per second, the values of L and C are quickly determined. It is seen from Fig. 30 that, for a given cutoff frequency f_c, as the load resistance increases the L increases, while the C value goes down. Very high-load

resistances require chokes of large inductance values, but as high-resistance loads mean small currents, the use of large inductances is feasible.

63. Ripple per Stage. By assuming that the load resistance R does not affect the values of L or C, a useful approximation[1] can be secured concerning the amount of filtering needed in each stage for the circuit shown in Fig. 28. Suppose the output stage is supplied with plate power which is filtered x per cent, so that its hum is reduced to x per cent of its unfiltered value, and at this value it gives no noticeable hum in the loud-speaker. Suppose further that the amplification between the plate of this last tube and the preceding tube plate is A. Then the preceding stage must have its power supply filtered x/A per cent. This means that the ripple in the plate supply of the next to the output stage must be $1/A$ as much as the output stage, because of its amplification. Figure 31 gives this relation in useful graphic form. If a stage of amplification has a gain of 25, it is essential that the preceding tube be supplied with plate power with one twenty-fifth the ripple, or 4 per cent. An LC product of 56 will give this degree of filtering at 100 cycles, according to Fig. 31, and this means a 28-henry choke and a 2-μf condenser which are close to standard values.

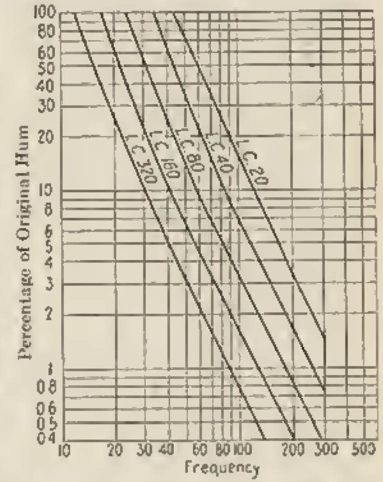


FIG. 31.—Smoothing effected by various products of inductance (henrys) and capacity (microfarads).

64. Resistor-capacitor Filter. A similar circuit to Fig. 28, using resistors instead of chokes, is frequently used to provide an extra degree of filtering for stages preceding a power stage (see Fig. 32). This is especially useful when the output stage requires a high voltage and when the voltage for the other stages must be materially reduced. The reason chokes are used is that they have high impedance to the unwanted rectified a.c., but low resistance to the desired d.c. Now, if the amount of d.c. is no great object, a resistance of as great a value as the impedance can be employed, and this is quite useful in some cases where the voltage is to be reduced. If, as in Fig. 32, two stages of choke and condenser filtering are used, the additional resistance and condenser filter stages simply increase the amount of filtering without the extra cost of chokes which are more expensive than

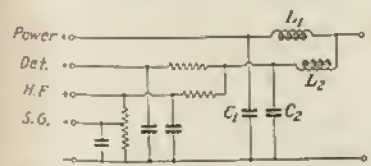


FIG. 32.—Circuit which minimizes feedback.

[1] COCKING, W. T., *Wireless World*, Nov. 19, 1930, pp. 565–568.

resistors. The RC values and the degree of filtering are given in Fig. 33, and the use is the same as that of Fig. 31. The circuit of Fig. 32 is quite similar to Fig. 28, in eliminating the undesired feedback effects.

The use of the chart (Fig. 30), based on Eqs. (1) and (2), gives very satisfactory results, but the experimental curves[1] showing the effects of load and different condenser values are quite interesting and will give a clearer idea of the validity of the chart.[2]

65. First Filter Condenser. The effect of the first filter condenser, shown dotted in Fig. 27, is to raise the available output voltage. Figure 34 gives the output voltage available as the first condenser C_1 is changed, as a function of the load current.

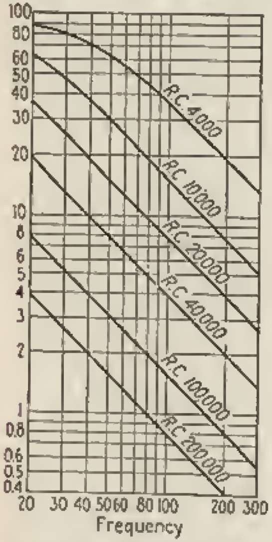


FIG. 33.—Filtering effected by resistance (ohms)-capacity (microfarads) circuit.

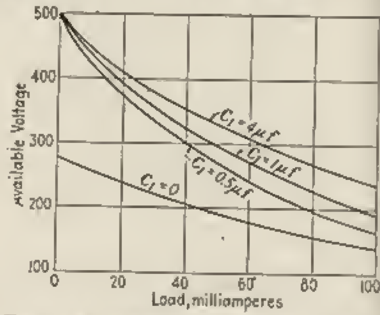


FIG. 34.—Effect of C_1 on voltage available.

Figure 35 gives the per cent ripple in output as the capacity of C_1 is varied. This curve shows that the use of a single condenser C_1 can never reduce the ripple much below 10 per cent with a reasonable value of capacity. Much less than one-half of 1 per cent is needed in a good filter, and, as at least two condensers must be used to provide a single filter section, Fig. 35 agrees with the theory.

66. Second and Third Filter Condensers. Figure 36 gives the per cent ripple as a function of C_2 and C_3 for a given current drain. It will be seen that, when $C_2 = C_3$, the most economical filter results. For example, suppose the ripple permissible to be 0.1 per cent. This can be supplied with $C_2 = 0$ if $C_3 = 5\ \mu f$, a total of $5\ \mu f$. But this can also be met with $C_2 = 2\ \mu f$, and $C_3 = 2\ \mu f$, a total of only $4\ \mu f$. The dotted line gives the ripple value where C_2 and C_3 are equal. The per cent ripple figures, of course, apply only to a specific filter, but the relations between the condenser values hold for similar filter circuits.

[1] The curves (Figs. 34 to 38) are experimental curves taken from the *Aerovox Research Worker*, articles by Sidney Fishberg, research engineer.

[2] A theoretical calculation of the effects of C_1, C_2, and C_3 on the output voltage is given in *Gen. Elec. Rev.*, 19, 177, 1916.

Figure 37 gives the percentage hum as a function of the current drain. This shows that the higher the values of C_2 and C_3 the lower the percentage hum. It should be remembered that increasing current means a decreasing load resistance. From Fig. 30, assuming f_c is constant, the capacity should increase and the inductance decrease as the load resistance decreases. Thus, as Fig. 37 was taken using the same

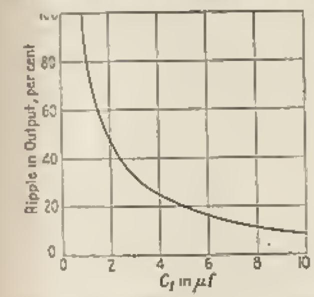


FIG. 35.—Effect of C_1 on ripple in output.

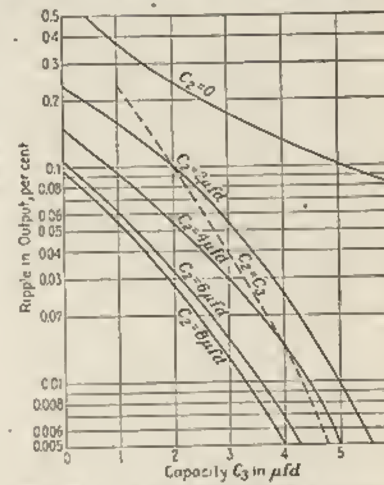


FIG. 36.—Percentage ripple as a function of C_2 and C_3.

inductance coils throughout, larger values for C_2 and C_3 are needed as the current drain increases. It is almost certain that the inductance values of

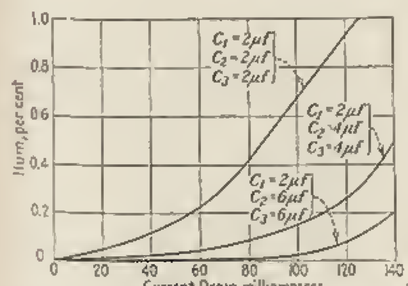


FIG. 37.—Percentage hum as a function of current drain.

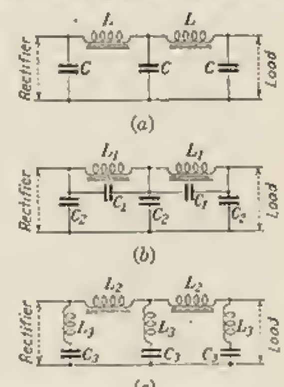


FIG. 38.—(a) Low-pass filter. (b) and (c) Tuned low-pass filters.

the chokes decreased as the current through them increased. To a certain extent this inductance decrease does not interfere with the filtering, especially if the capacity is increased, as, referring again to Fig. 30, when the resistance decreases to half a certain value, the capacity should be doubled, while the inductance need be only half its former value, if f_c be kept the same. Thus in Fig. 37 as in the other figures, the experimental facts agree with the theoretical chart (Fig. 30) and Eqs. (1) and (2) for this type of filter.

67. Swinging Choke. A "swinging" choke is often used in transmitter circuits as the first choke following the rectifier. The inductance of this type of choke varies from a maximum greater than that shown to be needed by calculation, at no load, to a minimum equal to that shown by the calculation.

68. Tuned Low-pass Filter. Two tuned low-pass filter circuits are given in Fig. 38, *b* and *c*, whose attenuation characteristics were given in Fig. 29*b*. For comparison, Fig. 38*a* gives the ordinary low-pass filter.

For the tuned filter of Fig. 38*a*, having the series chokes shunted by small condensers, the equations are

$$C_1 = \frac{1}{4\pi f_c R a\sqrt{a^2 - 1}} = \frac{0.07858}{f_c R a\sqrt{a^2 - 1}} \text{ farads} \qquad (3)$$

$$C_2 = 4C_1(a^2 - 1) \text{ farads} \qquad (4)$$

$$L = R^2 C \text{ henrys} \qquad (5)$$

$$a = \frac{f_R}{f_c} \qquad (6)$$

For the tuned filter of Fig. 38*c*, having small chokes in series with the condensers, the equations are

$$C_3 = \frac{\sqrt{a^2 - 1}}{f_c R a} = \frac{0.3183\sqrt{a^2 - 1}}{f_c R a} \text{ farads} \qquad (7)$$

$$L_2 = R^2 C_3 \text{ henrys} \qquad (8)$$

$$L_3 = \frac{L_2}{4(a^2 - 1)} \text{ farads} \qquad (9)$$

$$a = \frac{f_R}{f_c} \qquad (10)$$

If wide variations in the supply frequency were likely to occur, this type of filter would not be advisable. As a rule, the frequency of most power companies is now kept constant enough to run synchronous electric clocks, and this is quite good enough for this type of tuned circuit. However, the values of C_1, L_1, and C_3, L_3 have to be accurately maintained in order fully to secure the advantages of the tuned filter. Owing to these closer manufacturing limits, the use of the tuned filter is not so wide in large production as its advantages would seem to warrant. A combination of tuned low-pass filter and the regular-type filter is sometimes used with very good results.

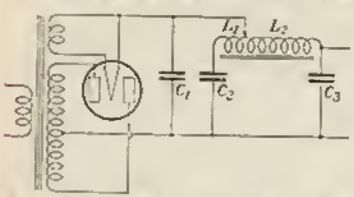


Fig. 39.—Tapped choke-filter circuit.

69. Filter Chokes Having Mutual Inductance. An interesting type of filter is one in which the first and second choke are magnetically

coupled. Figure 39 shows a tap on the first choke[1] to which the positive rectifier lead and a filter condenser are connected. The a-c component, flowing through the L_1 section of the choke, neutralizes to a large degree the a-c component of L_2, so that the output ripple is reduced. Figure 40 shows the relative a-c output ripple with a variable C_2 as the tap on the choke is changed, so that L_1 uses from 10 to 40 per cent of the total turns of the choke.

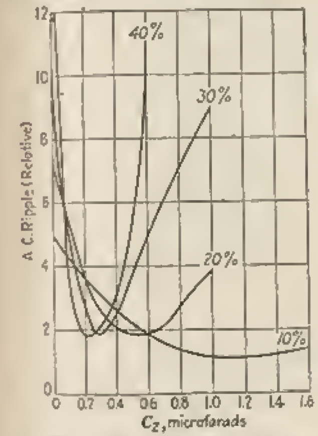


Fig. 40.—Tapped choke. Percentage of total turns (Fig. 39) used in L_1.

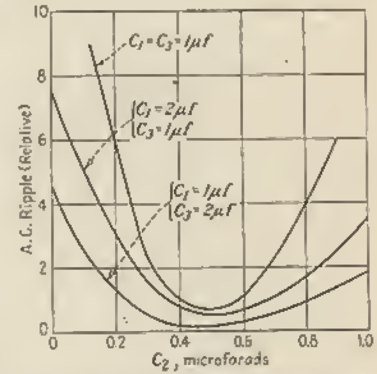


Fig. 41.—Condenser values for tapped choke filter (Fig. 39).

Figure 41 shows how the values of C_1 and C_3 affect the relative a-c ripple as a function of C_2. These curves indicate that the best C_2 value is fairly independent of C_1 and C_3.

70. Design of Filter Chokes. It is important that the filter choke be designed to carry the desired d.c. and at the same time to offer the necessary reactance to the a-c component. A direct method of design[2] has been derived using both the normal and incremental permeability curves for the core material.

The derivation gives the two following working equations:

$$\frac{LI^2}{V} = \frac{B^2\left(\frac{1}{\mu} + \frac{a}{l}\right)^2 \times 10^{-8}}{0.4\left(\frac{1}{\mu\Delta} + \frac{a}{l}\right)} \qquad (11)$$

$$\frac{NI}{l} = \frac{B}{0.4\pi}\left(\frac{1}{\mu} + \frac{a}{l}\right) \qquad (12)$$

where L = henrys
I = d-c amperes
V = core volume in cubic centimeters
N = turns

[1] *Proc. I.R.E.*, January, 1930, p. 161, from which Figs. 39 to 41 are taken.
[2] Hanna, C. R., *Jour. A.I.E.E.*, **46**, 128, February, 1927.

l = magnetic path in centimeters
a = air gap in centimeters
B = steady flux density on iron and air gap in gausses
μ = normal permeability B/H
$\mu\Delta$ = incremental permeability $\Delta B/\Delta H$ for a minor hysteresis loop.

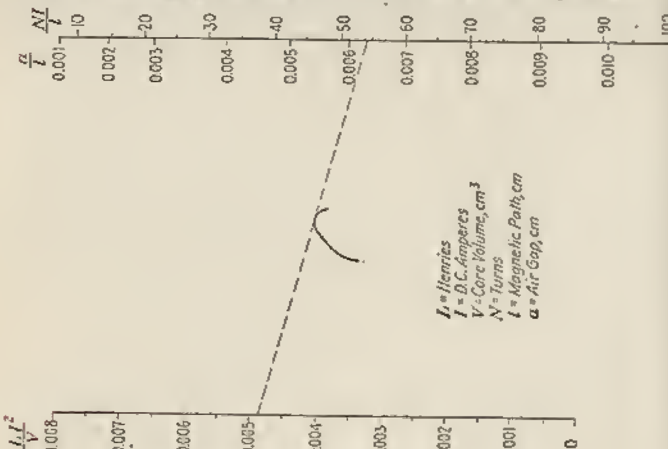


Fig. 43.—Choke design; hipernik core.

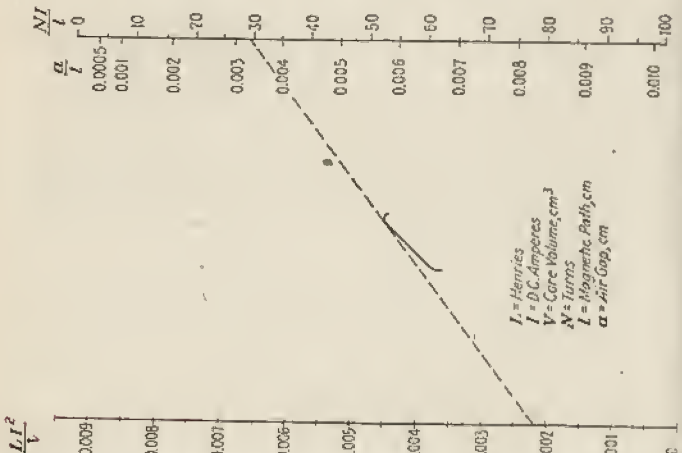


Fig. 42.—Choke design; 4 per cent silicon core.

The original curves were plotted with a/l as a parameter, LI^2/V being the ordinate, and NI/l as the abscissa for both 4 per cent silicon steel and hipernik. Figures 42 and 43 are alignment charts which include the data of the original curves. LI^2/V is the left column, and NI/l and a/l are on the right column. A straightedge passing through a given LI^2/V and tangent to the curve in

the central part of the chart will cut the right column at the corresponding value of NI/l and a/l. The reverse procedure, beginning with NI/l, is also possible.

Figure 44 gives typical permeability curves for three grades of magnetic material which is commercially available.[1] A chart for calculating

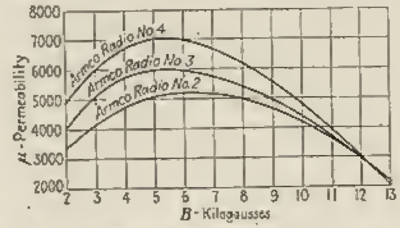


Fig. 44.—Typical permeability curves of radio grades of Armco iron.

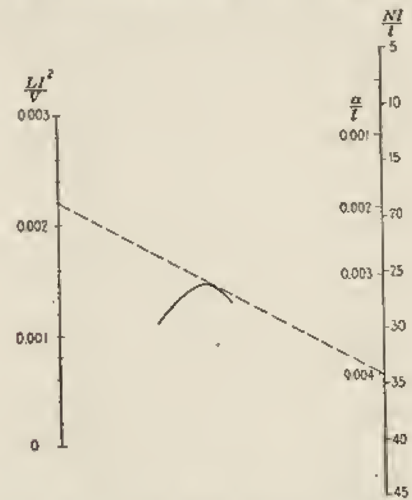


Fig. 45.—Choke design; Armco Radio 4.

chokes, using Armco Radio 4 is Fig. 45, the values of LI^2/V and NI/l being the same as for Figs. 42 and 43. In Fig. 45 either the desired value of LI^2/V is followed over the curve and then down to NI/l or the reverse procedure can be followed. The gap ratio a/l shown opposite the curve has exactly the same significance as before.

71. Designing a Choke to Carry D-c. A small choke to carry 80 ma and have 14 henrys is desired. The left column of Fig. 42 is LI^2/V, and this is

[1] These curves were supplied by C. W. Rust, electrical engineer, American Rolling Mill Co., Middletown, Ohio.

calculated first. L is 14 henrys. I is 0.08 amp.; I^2 is 64×10^{-4} amp.² V is the volume of the core, which was calculated to be 83.6 cc.

$$\frac{LI^2}{V} = \frac{14 \times 64 \times 10^{-4}}{83.6} = 10.7 \times 10^{-4} = 0.00107$$

Lining up this value with a straightedge which is tangent to the central curve (Fig. 42), the value of NI/l is found to be 18. The core used has l = 14 cm, so $N = 18 \times l/I = 18 \times 14/0.08 = 3{,}150$ turns. Thus to get 14 henrys, 3,150 turns are wound on the core given. To have this inductance at 80 ma, an air gap is needed, as shown in Fig. 42, the a/l (gap ratio) being 0.0021. As l is 14 cm, $a = l \times 0.0021$ or $14 \times 0.0021 = 0.029$ cm (equivalent to $0.029/2.54 = 0.011$ in.). This required air gap is made by inserting paper sheets of the proper thickness between the punchings, and then clamping them firmly in position.

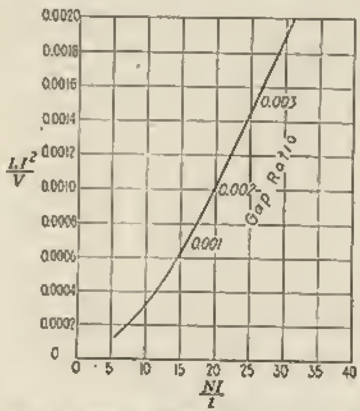


FIG. 46.—Typical reactor design curve; V in cubic centimeters and l in centimeters. Armco Radio 4.

The inductance of a choke depends to some degree on the frequency. For use with low frequencies in a filter circuit the inductance remains practically constant. Both the hysteresis loss and eddy-current loss are of importance in choosing a core material for chokes and transformers. The hysteresis loss is directly proportional to the frequency if the maximum flux density remains constant, and to the 1.6 power of the maximum flux density if the frequency remains constant.

The eddy-current loss can be kept low by using thin sheets of core material. A usual standard thickness is 0.014 in., and this is quite satisfactory for filter choke and transformers for 60 cycles. The insulation between laminations does not need to be very thick, the usual oxide layer on the sheet being sufficient.

72. Filter-condenser Ratings. Some rectifiers begin supplying rectified voltage before the tubes in the load heat up sufficiently to take their rated currents. (This is especially true of the slow indirect-heated tubes.) For this reason it is often desirable, especially from a safety factor viewpoint, to use peak voltages in calculating all condenser ratings.

The first condenser should, then, be able to stand the peak voltage of the power-transformer secondary. For a 400-volt secondary the peak is 564 volts. For reliable continuous use, the rating of the first filter condenser should be 564 volts. If no current flows, the voltage on both the second and third condensers will also be within a few per cent of the peak value 564 volts.

Assuming that an appreciable percentage of the total load current flows in the voltage divider as a "waste" or "circulating" current, the second and third condenser ratings do not have to be so high as that of the first condenser by the amount of the voltage drop in the chokes. This drop is figured by the usual $E = IR$ formula, where the circulating current is I and the resistance is that of the respective chokes.

If an appreciable load of resistors, or fast-heating tubes, is always in the circuit, the IR drop through the chokes can be subtracted from the voltage applied to the first condenser. For instance, if a current of 60 ma flows through the first choke having R = 400 ohms, the voltage drop is $0.06 \times 400 = 24$ volts. Assuming the r-m-s voltage (neglecting the tube drop) at the first condenser is 400 volts, the steady voltage component at the second condenser is $400 - 24 = 376$ volts. To this should be added 10 per cent to allow for the ripple, so that $376 \times 1.1 = 413.6$ volts should be the d-c rating for the second condenser.

It is true that a good filter condenser will stand, for a time, voltages greater than its d-c rating, but the practice of applying these higher voltages is seldom advisable.

73. Filter Condensers. The advent of reliable electrolytic condensers has greatly simplified the construction of adequate filter systems at a low cost. The untuned low-pass filter[1] can readily be made by using a very few high-capacity electrolytic units. The one thing to note carefully is that electrolytic condensers are suitable only for d.c.[2] and must be connected correctly. Condensers mounted in metal cans usually have the can negative. Cardboard units always have colored tracers, the red lead usually being plus. Wrongly connecting electrolytic condensers not only will ruin the condenser but are likely to damage the transformer and the rectifier tube. Mercury-vapor rectifiers, with their low voltage drop, are especially susceptible to such wrong connections.

Electrolytic condensers may be considered as comprising two general classes: low voltage, high capacity, from 1.0 to 50 µf, for grid filtering; and high voltage (up to 450 volts d.c.) for plate filtering, from 1 to 16 µf. Higher voltages are available by using several 450-volt units in series.

While electrolytic condensers are constantly being improved, there probably will always be a characteristic current which flows through the condenser. This current rises with temperature and may be a considerable fraction of a milliampere per microfarad at the temperature of radio-set operation.

There is a lower limit of temperature at which electrolytic condensers lose an appreciable portion of their capacity to store electrical energy. This limit is invariably below freezing (32°F.) and is usually of little consequence for indoor radio-set use.

[1] See Art. 64.
[2] For capacitor-type a-c motors, two electrolytic condensers are used in series, so that at all times one is connected correctly.

SECTION 15

HIGH-FREQUENCY TRANSMISSION AND RECEPTION

By Dale Pollack, Sc. D.

PROPERTIES OF H-F WAVES

1. Classification into Frequency Ranges. For purposes of analysis it is convenient to divide radio waves into bands of frequencies within which propagation effects are similar. Any such classification must, in part, be arbitrary, since changes in the properties of waves with frequency are not sharply defined and are dependent upon time. In this section the classification of Table I will be followed in general, but minor changes

Table I.—Classification of Frequency Ranges

Range		Nomenclature	Approximate useful communication radius
Kilocycles	Meters		
Below 550........	Above 545	Low frequencies	Ground wave: 0–1,000 miles Sky wave: 500–8,000 miles
550–1,600........	545–187	Broadcast band	Ground wave: 0–100 miles Sky wave: 100–1,500 miles
1,600–30,000......	187–10	High frequencies	Ground wave: 0–15 miles Sky wave: 15–8,000 miles
Above 30,000.....	Below 10	Ultra-high frequencies	0–150 miles

in the dividing frequency between ranges will be made when desirable. The column Approximate Useful Communication Radius is relatively approximate, the exact radius being dependent upon the power, time, earth properties, and other conditions.

2. General Characteristics of H-f Waves. The signal intercepted by a radio-receiver antenna may have been propagated either by the ground wave, which travels along the earth's surface, or by the sky wave, which travels in the air, including the ionized layers above the earth. The received signal is, in the general case, made up of components of both types, but it is often convenient to investigate transmission by each of the modes independently. Frequently one or the other predominates, facilitating such independent analysis.

The *ground wave* is attenuated by losses in the earth and falls off at distances from the transmitter, the exact manner in which it decreases being dependent upon the frequency and the conductivity and dielectric constant of the earth. The ground wave is useful only for medium

distance communication at low frequencies and for short distances at low and broadcast frequencies. At high frequencies it can be employed only for local communication.

The *sky wave* may be employed for communication over greater distances since it travels through the air, in which the attenuation is relatively small. It can be refracted and reflected by the ionized layers in the upper atmosphere (the ionosphere) and by the earth. Thus the wave may return to earth at distances remote from the transmitter. All frequencies, up to the ultra high, may be reflected or refracted to earth by the ionosphere, and their sky waves may, therefore, be useful.

If the ground wave is to render good service at a distant receiving point, it must be strong compared with (1) the noise and interference level and (2) the sky wave. The limit to the useful service range of the ground wave may be fixed by one or both of these factors. The ground wave may be increased with respect to the noise and interference level either by increasing the transmitter power or by employing a directional antenna, but with respect to the sky wave the only effective aid is the use of an antenna system in which high angle radiation is minimized. In the region in which the sky wave and the ground wave are nearly equal in magnitude (within, perhaps, 2 to 1 of each other) fading, particularly selective fading, is excessive. At locations beyond the region in which the ground and sky waves are of similar magnitude, the sky wave predominates and is most useful for communication.

Ultra-high frequencies are very rapidly attenuated over the earth's surface, and their sky wave is not normally returned to earth by the ionosphere. Consequently, the transmission properties of u-h-f waves must depend upon the direct ray from the antenna and upon reflections of the direct ray from the earth's surface, as assisted by diffraction around the curved surface of the earth and refraction in the lower atmosphere. Their usefulness is limited to short distances although greater than the optical line-of-sight limitation which is too frequently assumed. As the frequency is increased further, however, diffraction and refraction phenomena become less useful in extending the service radius. Consequently communication by means of centimeter waves is limited to paths that are only a little longer than optical.

3. Propagation of the Ground Wave. The distance for which the ground wave is useful decreases as the frequency is raised. The ground wave of low frequencies is useful up to medium distances, and in the broadcast band its usefulness is limited to short distances. Above the broadcast band it can be used only for local transmission.

Sommerfeld[1] has computed the propagation of waves over a plane earth, i.e., for distances short enough that the earth's curvature may be neglected. The results may be further simplified if the dielectric constant of the earth may be neglected, which is true within an error smaller than 2 to 1, if the frequency is less than

$$f_e = \frac{7 \times 10^{15}\sigma}{\epsilon} \tag{1}$$

where f_e = frequency below which dielectric constant may be neglected, in kilocycles

σ = soil conductivity in e.m.u.

ϵ = dielectric constant.

[1] Sommerfeld, A., The Propagation of Waves in Wireless Telegraphy, *Ann. Phys.*, **28**, 665–736, Mar. 16, 1909; **81**, 1135–1153, Dec. 11, 1926.

Typical values of soil conductivity and dielectric constant are as follows:

Type of soil	σ (e.m.u.)	ϵ	f_c (kc)
Dry, sandy or rocky, ground	10^{-14}	5	1,400
Eastern and far Western United States (approximately)	3×10^{-14}		
Average ground, Central United States and Europe (approximately)	10×10^{-14}		
Moist ground	30×10^{-14}	30	7,000
Salt water	4×10^{-11}	80	350,000

Detailed data on soil conductivity throughout the United States have been published by the FCC.[1]

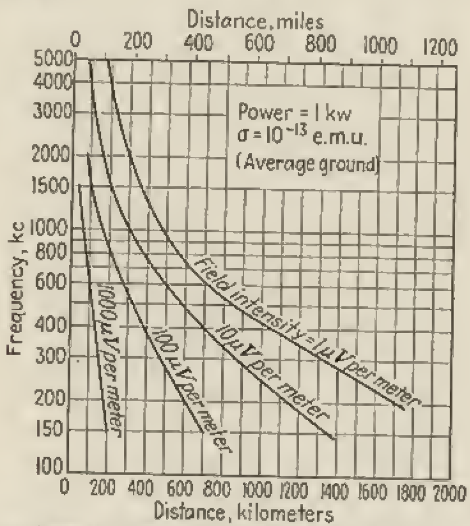


FIG. 1.—Ground-wave propagation over average ground; 1 kw, vertical $\lambda/4$ antenna. (*Report of Committee on Radio Wave Propagation, Proc. I.R.E.*, **26**, 1193, *October*, 1938.)

Making use of the above assumptions, the ground-wave field intensity, according to van der Pol's[2] simplification of Sommerfeld's work, is given by

$$\mathcal{E} = \frac{k\sqrt{PA}}{d} \tag{2}$$

where $\mathcal{E}$ = field intensity in millivolts per meter
k = antenna constant (= 195 for quarter-wave antenna or 270 for half-wave antenna)

[1] Federal Communications Commission, Standards of Good Engineering Practice concerning Standard Broadcast Stations, U. S. Government Printing Office, 1940.
[2] VAN DER POL, B., Propagation of Electric Waves, *Zeit. Hochfrequenz*, **37**, 152–156, April, 1931. For an excellent discussion of propagation formulas, see K. A. Norton,

d = distance in miles
P = power radiated in kilowatts
$A = \dfrac{2 + 0.3\rho}{2 + \rho + 0.6\rho^2}$ (Sommerfeld's "reduction factor")
$\rho = \dfrac{9.38 \times 10^{-21} f^2 d}{\sigma}$ (Sommerfeld's "numerical distance")
f = frequency in kilocycles
σ = soil conductivity in e.m.u.

When large numbers of such data must be computed, the slide rule of J. F. Morrison[1] is convenient. In practical computations of wave propagation, account must be taken of changes in the soil conductivity within the transmission range. A procedure for accomplishing this is outlined by P. P. Eckersley.[2]

More recent analyses have extended propagation calculations to account for the earth's curvature. Curves giving the results of such

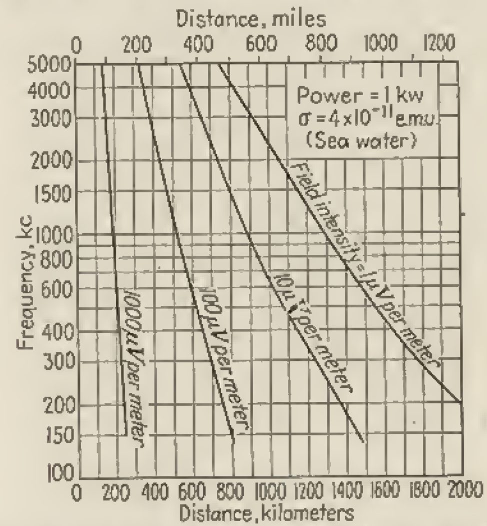


FIG. 2.—Ground-wave propagation over sea water; 1 kw radiated power, vertical $\lambda/4$ antenna. (*Report of Committee on Radio Wave Propagation, Proc. I.R.E.*, **26**, 1193, *October*, 1938.)

calculations will be found in the Report of Committee on Radio Wave Propagation.[3] Two such sets of curves for two typical earth conductivities (those of average ground and sea water, respectively) are shown in Figs. 1 and 2. For radiated powers different from 1 kw, the field

The Propagation of Radio Waves over the Surface of the Earth and in the Upper Atmosphere, Part I (Ground Wave), *Proc. I.R.E.*, **24**, 1367–1387, October, 1936.
[1] For sale by Keuffel & Esser Co., Hoboken, N. J.
[2] The Calculation of the Service Area of Broadcast Stations, *Proc. I.R.E.*, **18**, 1160–1193, July, 1930.
[3] *Proc. I.R.E.*, **26**, 1193–1234, October, 1938.

intensities of these figures should be multiplied by $\sqrt{P}$, where P is the radiated power in kilowatts. For other antenna structures the field should be multiplied by the appropriate factor. For a half-wave antenna, for example, this factor is 1.4.

4. Propagation of the Sky Wave. Energy which leaves the antenna at angles greater than zero constitutes the sky wave, in contrast to the ground wave, which leaves the antenna at a zero angle to the earth's surface. The sky wave travels through the atmosphere until it reaches the ionized layers 100 to 500 km above the earth, called the *ionosphere* or the Kennelly-Heaviside layer. It may then be bent back to the earth immediately, it may travel in the ionosphere for some distance before being returned to earth, or it may pass through the ionosphere and never return to earth. Long-distance radio communication is almost always accomplished by means of the portion of the sky wave refracted to earth. After the wave returns to earth, it may be reflected from the earth's surface into the ionosphere once more and to the earth at a more distant point.

The medium through which the ground wave is propagated, the surface of the earth, changes little with time. The sky-wave propagation, however, depends upon the ionosphere characteristics, which vary widely with time.

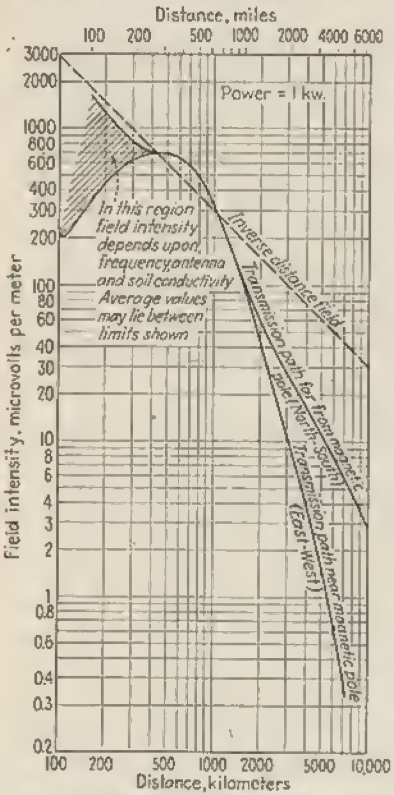


FIG. 3.—Sky-wave propagation, quasi-maximum field intensity for frequencies up to 1,600 kc. (*Report of Committee on Radio Wave Propagation*, *Proc. I.R.E.*, **26**, 1193, *October*, 1938.)

Sky-wave propagation is conveniently subdivided into low, broadcast, and high frequencies. The sky wave at low frequencies is usually considerably stronger at night than during the day. The day field may be estimated from the modified Austin-Cohen formula[1]

$$\mathcal{E} = 3 \times 10^5 \frac{\sqrt{P}}{d} \sqrt{\frac{\theta}{\sin \theta}}\, e^{-46\times 10^{-6} f^{0.4} d} \qquad (3)$$

[1] AUSTIN, L. W., Preliminary Note on Proposed Changes in the Constants of the Austin-Cohen Transmission Formula, *Proc. I.R.E.*, **14**, 377–380, June, 1926.

where $\mathcal{E}$ = field intensity in microvolts per meter
P = radiated power in kilowatts
d = distance in kilometers
θ = angle at center of earth subtended by transmission path in radians
f = frequency in kilocycles.

The quasi-maximum field (the field intensity which is exceeded only 5 per cent of the time) for a propagation path completely in darkness is given by Fig. 3.[1] For long distances a distinction must be made between transmission

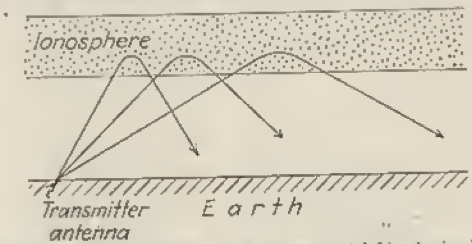


FIG. 4.—Probable paths traversed by sky wave at l.f. A single ionized layer is shown.

paths near the magnetic pole (corresponding to a north-south or south-north transmission, such as between Europe and South America, or between North and South America) and transmission paths far from the earth's magnetic pole (corresponding in general to east and west transmission, such as between northern United States and northern and central Europe, or between northern and central Europe and Siberia).[1] The median value of the field (the field exceeded 50 per cent of the time) is about 35 per cent of the quasi-maximum value. For powers other than 1 kw multiply the field intensity by $\sqrt{P}$ where P is in kilowatts. P should also include the antenna power gain if a directional antenna is employed.

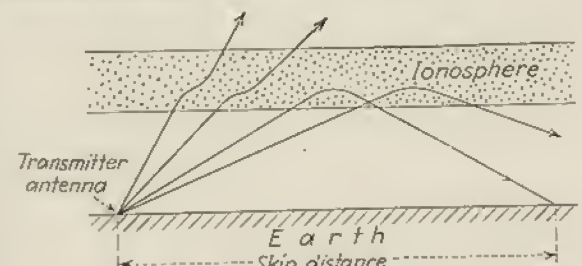


FIG. 5.—Probable paths traversed by sky wave at h.f. A single ionized layer is shown.

When part of the transmission path is in twilight, l-f propagation suffers. The field strength is likely to be lower than during either night or day transmission, for long-distance communication, and its value is not easily predicted.[2]

[1] Report of Committee on Radio Wave Propagation, *Proc. I.R.E.*, **26**, 1193–1234, October, 1938.

[2] ESPENSCHIED, L., C. N. ANDERSON, and A. BAILEY, Trans-Atlantic Radio Telephone Transmission, *Proc. I.R.E.*, **14**, 7–57, February, 1926.

At broadcast frequencies the sky wave does not return to earth during the day. Sky-wave propagation, therefore, need be considered only during the night. Figure 3 represents propagation of broadcast frequencies at night.

High-frequency sky-wave propagation is a complex phenomenon. For each of the ionosphere layers there exists a critical frequency, below which radiation from any angle from the antenna is returned to earth. Above this critical frequency high-angle radiation passes through the ionized layer, while radiation from lower angles is still returned to earth. This is illustrated by Figs. 4 and 5. In Fig. 4, for a frequency lower than the critical value, the entire sky wave is returned to earth. In Fig. 5, however, the frequency is higher than critical, and the high-angle radiation is not returned to earth. The distance between the transmitter and the point at which the highest angle radiation returns to earth is the *skip distance* for that frequency. Conversely, for this smallest skip distance, the corresponding frequency is called the *maximum usable frequency*. In general the higher the frequency the greater is the skip distance.

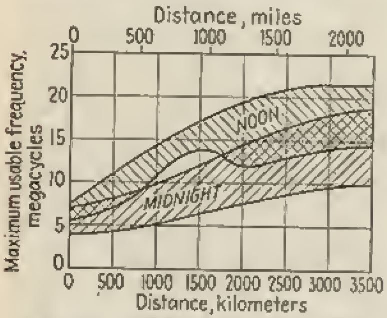


FIG. 6.—Average maximum usable frequencies during summer at latitude 39°N. Time refers to place where wave is reflected. The upper edge at each range is for the sunspot maximum, 1938–1939, the lower edge for the sunspot minimum, 1933–1934. (*Bureau of Standards Ionosphere Reports, Monthly in Proc. I.R.E.; Smith, Gilliland, and Kirby, Natl. Bur. Standards, Jour. Research*, **21**, 835, *December*, 1938; *Gilliland, Kirby, Smith, and Reymer, Proc. I.R.E.*, **26**, 1347, *November*, 1938.)

Within the skip distance the signal strength is usually too weak to be useful, which explains the necessity for employing frequencies lower than the maximum usable frequency. In traversing nonionized air, h-f waves suffer little attenuation, other than that resulting from spreading of the wave front. In passing through ionized air, however, the attenuation is greater. An operating frequency should be chosen for which as little of the path as possible is in the ionosphere. This will be the case if a frequency slightly lower than the maximum usable frequency is employed.

The maximum usable frequency is dependent upon the time of day (or longitude), month, year, and latitude, as well as upon the distance between transmitter and receiver. The curves of Figs. 6 and 7 give the values over which the maximum usable frequency has ranged during the period between 1933 and 1939 for summer and for winter transmission, respectively. The time (noon and midnight are shown separately in the figures) and latitude (39° north) refer to the place at which the ionosphere reflection takes place, usually halfway between transmitter and receiver for transmission paths less than 3,500 km long. The curves represent average conditions only; during the relatively infrequent periods of ionosphere disturbances, the maximum usable frequencies may be considerably changed. While the measurements were made at latitude 39° north, the results may probably be used with insignificant error between 30° and 50° north.

The curves are shown for distances up to 3,500 km. For greater distances the maximum usable frequency is substantially the same as for 3,500 km.

The higher values within the ranges illustrated in Figs. 6 and 7 were measured during the period of high sunspot activity, 1938–1939, while the lower limits are for the inactive portion of the sunspot cycle, 1933–1934. The cycle is expected to repeat itself, with the next minimum about 1944.[1]

Curves giving the maximum usable frequency in greater detail will be found in the Report of Committee on Radio Wave Propagation.[2] Monthly reports are published in the *Proc. I.R.E.* (since September, 1937) and in the *Bull. U.R.S.I.*

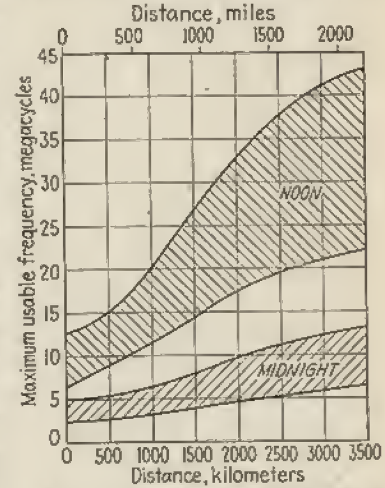


FIG. 7.—Average maximum usable frequencies during winter, at latitude 39°N. Time refers to place at which wave is reflected. The upper edge of each range is for the sunspot maximum, 1938–1939. The lower edge is for the sunspot minimum, 1933–1934. (*Bureau of Standards Ionosphere Reports, Monthly in Proc. I.R.E.; Smith, Gilliland, and Kirby, Natl. Bur. Standards, Jour. Research*, **21**, 835, *December*, 1938; *Gilliland, Kirby, Smith, and Reymer, Proc. I.R.E.*, **26**, 1347, *November*, 1938.)

From a knowledge of the ionosphere characteristics over the earth, it is now possible to compute the propagation of h-f waves.[3] During the past few years, sufficient ionosphere data have been accumulated to permit such calculations to be made. Propagation maps, for frequencies of 8.6 and 18.8 Mc, calculated entirely from ionosphere data are given in the Report of Committee on Radio Wave Propagation.[4]

The refraction of ultra-high frequencies by the ionosphere is erratic. In the winters of 1936 through 1939, however—the peak in the sunspot cycle was in 1938–1939—maximum usable frequencies for long-distance daytime transmission exceeded 40 Mc, as is indicated in Fig. 7. Except for such instances of extreme maximum usable frequencies, ultra-high frequencies are rarely employed for dependable ionosphere transmission.

[1] SMITH, N., T. R. GILLILAND, and S. S. KIRBY, Trends of Characteristics of the Ionosphere for Half a Sunspot Cycle, *Natl. Bur. Standards, Jour. Research*, **21**, 835–846, December, 1938; GILLILAND, T. R., S. S. KIRBY, N. SMITH, and S. E. REYMER, Maximum Usable Frequencies for Radio Sky-wave Transmission, 1933–1937, *Proc. I.R.E.*, **26**, 1347–1359, November, 1938.

[2] *Loc. cit.* Also SMITH, GILLILAND, and KIRBY, *loc. cit.*; GILLILAND, KIRBY, SMITH, and REYMER, *loc. cit.*

[3] An early attempt at such computation was by S. Namba and T. Tsukada, A Method of Calculation of Field Strengths in High Frequency Radio Transmission, *Proc. I.R.E.*, **21**, 1003–1028, July, 1933. More recent and more satisfactory methods are presented in the following papers: SMITH, N., Application of Vertical-incidence Ionosphere Measurements to Oblique-incidence Radio Transmission, *Natl. Bur. Standards, Jour. Research*, **20**, 683–705, May, 1938; MILLINGTON, G., The Relation between Ionosphere Transmission Phenomena at Oblique Incidence and Those at Vertical Incidence. *Proc. Phys. Soc. (London)*, **50**, 801–825, September, 1938.

[4] *Loc. cit.*

5. Propagation of Ultra-high Frequencies. Ultra-high-frequency waves, as intercepted at the receiver antenna, are made up of two components, one received directly from the transmitter antenna, the second reflected by the earth.

When the distance between transmitter and receiver is small enough so that the earth's curvature may be neglected, the received field strength may be computed from

$$\mathcal{E} = \frac{12.6 \times 10^6 \mathcal{E}_0 f h_t h_r}{d^2} \ \mu\text{v per meter} \tag{4}$$

where h_t and h_r = heights of the transmitting and receiving antennas, respectively, in kilometers
$\mathcal{E}_0$ = field intensity 1 km from transmitter in microvolts per meter, in the direction of maximum field strength. (For a half-wave dipole, $\mathcal{E}_0 = 220\sqrt{P}$)
P = transmitter power in kilowatts
d = distance between transmitter and receiver in kilometers
f = frequency in megacycles.

This equation assumes the earth to be a perfect plane reflector and both antennas are at least several wave lengths above the earth.

More recently the theory of u-h-f propagation has been improved to take account of diffraction of waves around the curved surface of an earth of finite conductivity.[1] The results, from Eckersley's report, are plotted in Figs. 8 and 9 for frequencies of 50 and 150 Mc. For other frequencies the original paper or the Committee Report on Radio Wave Propagation[2] should be referred to. The curves are plotted for several values of H, which represents either the receiver or transmitter antenna height, assuming the the other antenna height to be zero. If both antennas are above the surface of the earth, the correction may be obtained from the curves in the following manner: Assume first (the roles of transmitter and receiver antenna may be interchanged if desired) that H = transmitter antenna height and that the receiver antenna height is zero. Read the appropriate field strength from the curves. To this a correction for the actual receiver antenna height is to be added. This correction is the vertical distance, in decibels, as read from the right-hand ordinate, between the $H = 0$ curve and the curve for H = receiver antenna height at the appropriate distance. To illustrate, find the field strength for $h_t = 500$ meters, $h_r = 50$ meters, $f = 50$ Mc, $d = 100$ km, and $P = 1$ kw (if the antenna were at the earth's surface). The field for $H = 500$ at $d = 100$ km is 12 μv per meter. The correction between $H = 0$ and $H = 50$ is $34 - 5 = 29$ db, corresponding to a ratio of 28.2:1. The corrected field is thus $12 \times 28.2 = 338$ μv per meter. This method of correction applies accurately to transmission distances beyond the line of sight. For low antenna heights the method of correction of Eckersley's paper should be referred to.

Line-of-sight distances are plotted in Fig. 10 and are also indicated in Figs. 8 and 9. Note in Figs. 8 and 9 that the field strength at the line-of-sight distance is nearly independent of the antenna height. For a 1-kw antenna power at frequencies between 50 and 150 Mc, the field is between 20 and 40 μv per meter for antenna heights between 100 and 2,000 meters. While attenuation is more rapid beyond the line-of-sight distance, nevertheless reliable u-h-f communication is perfectly possible to distances far beyond the optical horizon, as is evident from Figs. 8 and 9.

[1] ECKERSLEY, T. L., Ultra-short Wave Refraction and Diffraction, *Jour. I.E.E.*, **80**, 286–304, March, 1937; VAN DER POL, B., and H. BREMMER, The Diffraction of Electromagnetic Waves from an Electrical Point Source round a Finitely Conducting Sphere, *Phil. Mag.*, **24**, 141–176, 826–864, July and November, 1937.
[2] *Loc. cit.*

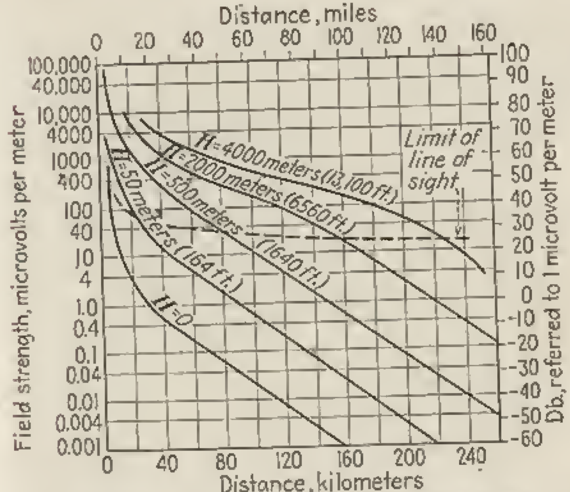


FIG. 8.—Ultra-high-frequency propagation over average land ($\epsilon = 5$, $\sigma = 10^{-13}$ e.m.u.) at 50 Mc. Antenna current equals that produced by 1 kw in antenna at surface of earth. (*Report of Committee on Radio Wave Propagation, Proc. I.R.E.*, **26**, 1193, *October*, 1938.)

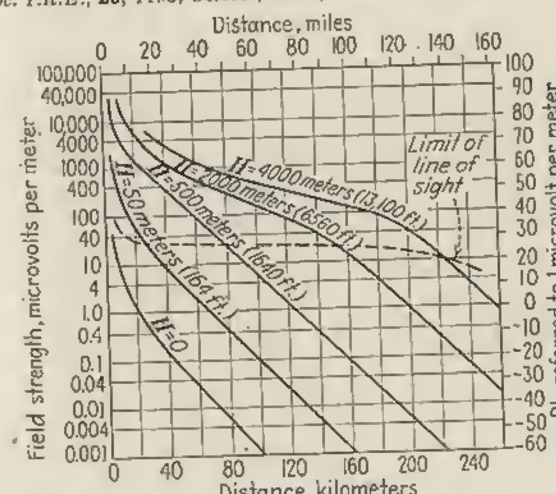


FIG. 9.—Ultra-high-frequency propagation over land ($\epsilon = 5$, $\sigma = 10^{-13}$ e.m.u.) at 150 Mc. (*Report of Committee on Radio Wave Propagation, Proc. I.R.E.* **26**, 1193, *October*, 1938.)

The rate of attenuation beyond the horizon, however, increases with frequency, and for very short centimeter waves transmission is not possible very far beyond the line of sight. Within the horizon the attenuation may be expressed, roughly, by an inverse square relationship, Eq. (4), but beyond the horizon the attenuation is better expressed by

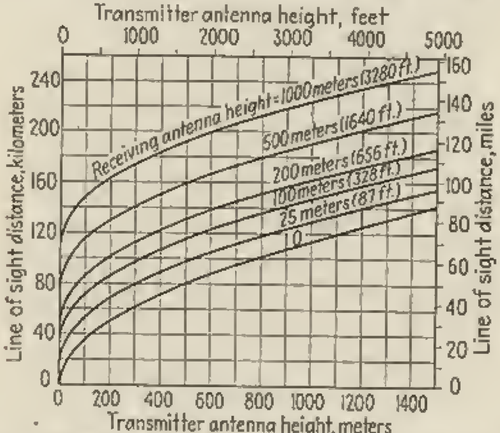


Fig. 10.—Length of line-of-sight path for various antenna heights.

an exponential equation of the form

$$\mathcal{E} = ke^{-\alpha d} \tag{5}$$

in which k and α are constants independent of the distance. If $\mathcal{E}$ is measured in microvolts per meter and d in kilometers, the value of α may be obtained from Fig. 11.

Figures 8 and 9 take account of diffraction around the curved surface of the earth, but they do not include the effect of refraction caused by the variation in the density of the air near the earth's surface. Refraction

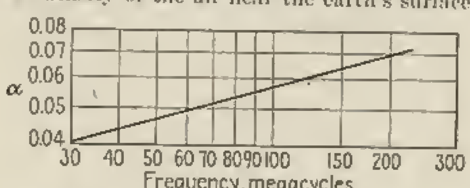


Fig. 11.—Factor α in Eq. (5).

has two effects. It increases the received field, and it causes it to vary or fade as the temperature gradient varies. The effects may be expressed in terms of an increase in the effective radius of the earth.[1] The observed

[1] Smith-Rose, R. L., and J. S. McPetrie, Ultra-short Waves: Refraction in the Lower Atmosphere, *Wireless Eng.*, **11**, 3–11, January, 1934; Englund, C. R., A. B. Crawford, and W. W. Mumford, Further Results of a Study of Ultra-short-wave

fields are in consequence higher than the theory represented by Figs. 8 and 9 indicates, and are equivalent to an increase of perhaps 15 to 35 per cent in the earth's radius, with a corresponding increase in the line-of-sight transmission distance. Furthermore, the received field is probably higher in summer than in winter and higher at night than during the day. Some evidence on these points is conflicting, and insufficient data prevent the drawing of any more specific conclusions.[1]

The propagation formulas assume that the terrain is flat. In practice such conditions rarely occur. This is especially true in cities among tall

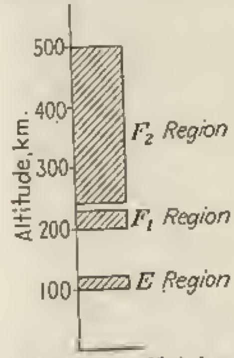


Fig. 12.—Heights over which ionosphere layers range.

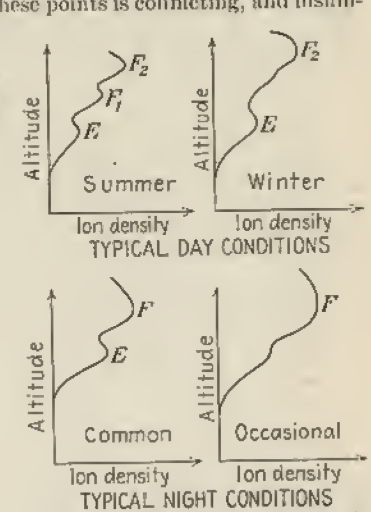


Fig. 13.—Probable variation in atmospheric ion density with altitude.

buildings, where field strengths considerably lower than the values predicted by the curves, may be observed.[2]

6. Ionosphere Characteristics.[3] All long-distance radio communication takes place through refraction of waves in the ionized layers above the earth. Ionosphere research has been accelerated during recent years, and knowledge of ionosphere characteristics has increased accordingly.

Transmission Phenomena, *Bell System Tech. Jour.*, **14**, 369–387, July, 1935; Hull, R. A., Air-wave Bending of Ultra-high-frequency Waves, *QST*, **21**, 16–18, May, 1937.

[1] Burrows, C. R., A. Decino, and L. E. Hunt, Ultra-short-wave Propagation over Land, *Proc. I.R.E.*, **23**, 1507–1535, December, 1935; Stability of Two-meter Waves, *Proc. I.R.E.*, **26**, 516–528, May, 1938.

[2] Burrows, C. R., L. E. Hunt, and A. Decino, Ultra-short-waves in Urban Territory, *Elec. Eng.*, **45**, 115–124, January, 1935.

[3] For a general introduction to the subject see P. O. Pederson, "The Propagation of Radio Waves," published by G. E. C. Gad, Copenhagen (in English) or K. K. Darrow, The Ionosphere, *Elec. Eng.*, **59**, 272–283, July, 1940. A historical survey and résumé of information is contained in S. S. Kirby, L. V. Berkner, and D. M. Stuart, Studies of the Ionosphere and Their Application to Radio Transmission, *Proc. I.R.E.*, **22**, 481–521, April, 1934, and is brought up to date in Report of Commission II, Radio Wave Propagation, International Scientific Radio Union, *Proc. I.R.E.*, **27**, 645–649, October, 1939; and in T. R. Gilliland, S. S. Kirby, N. Smith, and S. E. Reymer, Characteristics of the Ionosphere and Their Application to Radio Transmission, *Proc. I.R.E.*, **25**, 823–840, July, 1937.

The ionization is apparently caused by ultraviolet radiation from the sun, and it varies with the 11-year sunspot cycle, the time of day, the time of year, and the longitude. At the earth's surface the ion density is very small, increasing to maximum values at altitudes between 100 and 500 km. While the ions are distributed continuously in the atmosphere, the concentrations vary and several maximums are reached at various altitudes. The regions near these maximums are called *layers*. For each such layer there is a *critical frequency* above which an electromagnetic wave, directed vertically, will not be returned to earth. Ionosphere characteristics are frequently studied from measurements of critical frequencies.[1]

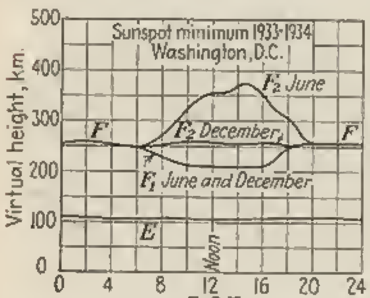


Fig. 14.

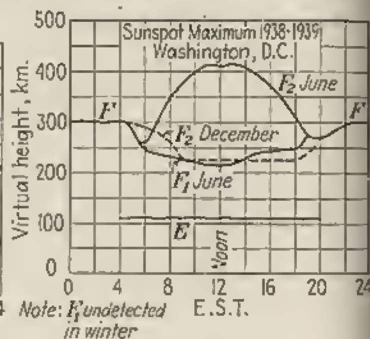


Fig. 15.

Fig. 14.—Average diurnal variation in virtual heights of ionosphere layers during sunspot minimum. (*Bureau of Standards Ionosphere Reports, Monthly in Proc. I.R.E.; Smith, Gilliland, and Kirby, Natl. Bur. Standards, Jour. Research*, **21**, 835, *December*, 1938; *Gilliland, Kirby, Smith, and Reymer, Proc. I.R.E.*, **26**, 1347, *November*, 1938.)

Fig. 15.—Average diurnal variation in virtual heights of ionosphere layers during sunspot maximum. (*Bureau of Standards Ionosphere Reports, Monthly in Proc. I.R.E.; Smith, Gilliland, and Kirby, Natl. Bur. Standards, Jour. Research*, **21**, 835, *December*, 1938; *Gilliland, Kirby, Smith, and Reymer, Proc. I.R.E.*, **26**, 1347, *November*, 1938.)

The three ionosphere regions of greatest importance, so far as radio communication is concerned, are denoted the E, F_1, and F_2 layers. The range of altitudes over which these layers may vary is shown in Fig. 12. At night the F_1 and F_2 layers merge, forming the F region. It is also possible that a C layer, lower than the E, and a G layer, higher than the two F layers, exist.[2]

The ion density in the upper atmosphere probably varies in some such manner as shown in Fig. 13. Only the virtual heights of the maximum points and the apparent densities at these heights are known. For intermediate altitudes it is usually assumed that the ion density goes through minimum points as shown.

[1] Gilliland, T. R., Multifrequency Ionosphere Recording and Its Significance, *Proc. I.R.E.*, **23**, 1076–1101, September, 1935.

[2] Kirby, S. S., and E. B. Judson, Recent Studies of the Ionosphere, *Proc. I.R.E.*, **23**, 733–751, July, 1935; Cowell, R. C., and A. W. Friend, The Lower Ionosphere, *Phys. Rev.*, **50**, 632–635, Oct. 1, 1936.

The E layer is usually constant in height at 100 to 120 km, except during the relatively infrequent appearance of the "sporadic E," whose course is not predictable. At such times the E layer height goes through extreme variations. The E layer, furthermore, is not always detectable at night.

The F layer, which exists only at night, exhibits little diurnal or seasonal change in height. Its height does, however, change during the sunspot cycle. The F layer divides into the F_1 and F_2 layers during the day. The F_2 layer varies symmetrically about noon and is considerably higher in summer than in winter. The F_2 layer appears to be the daytime continuation of the F layer, while the F_1 layer is non-existent during the night. After 1933 the F_1 layer gradually became less clearly defined and between 1936 and 1939 disappeared entirely during winter. It will probably reappear in winter before 1943.

A wave directed toward the ionosphere is split into two components, one in the direction of the earth's magnetic field, the other at right angles to it. The two rays are called the *ordinary* and the *extraordinary* rays, and, since the magnetic field acts upon them differently, their propagation in the ionosphere is different.[1] One consequence of this is that two critical frequencies exist for each ionized layer, the critical frequency of the ordinary ray, usually denoted by a superscript o, and the critical frequency of the extraordinary ray, denoted by the superscript x. Subscripts refer to the layer. For example, $f_{F_2}{}^x$ refers to the critical frequency of the extraordinary ray for the F_2 layer. For frequencies greater than 2.5 Mc the ordinary and extraordinary critical frequencies at Washington, D.C., are related approximately by

$$f^x = f^o + 0.8 \qquad (6)$$

where the frequencies are in Mc.[2] Detailed consideration of critical frequencies and their significance in radio communication is described in the literature.[3]

Normal ionosphere properties, *i.e.*, those whose variations may now be predicted with reasonable precision, have been emphasized above. In addition to these variations there are those resulting from less easily predictable ionosphere disturbances, *viz.*, ionosphere storms, sudden ionosphere disturbances resulting in fade outs and lengthy periods of absorption below the E layer. Considerable study is being devoted to such disturbances, but various investigators disagree as to the predictability of such effects.[4]

The U. S. Bureau of Standards and the International Scientific Radio Union (U.R.S.I.) have broadcast and published comprehensive ionosphere data for several years. For details of maximum usable frequencies, virtual

[1] Appleton, E. V., and G. Builder, Ionosphere as a Doubly refracting Medium, *Proc. Phys. Soc.*, **45**, 208–220, Mar. 1, 1933.

[2] Gilliland, T. R., S. S. Kirby, N. Smith, and S. E. Reymer, Characteristics of the Ionosphere and Their Appliction to Radio Transmission, *Proc. I.R.E.*, **25**, 823–840, July, 1937.

[3] Smith, Gilliland, and Kirby, *loc. cit.*; Gilliland, T. R., S. S. Kirby, N. Smith, and S. E. Reymer, Maximum Usable Frequencies for Radio Sky-wave Transmission, 1933–1937, *Proc. I.R.E.*, **26**, 1347–1359, November, 1938; Pederson, *op. cit.*; Kirby, Berkner, and Stuart, *loc. cit.*; Report of Commission II, *loc. cit.*; Gilliland, T. R., S. S. Kirby, N. Smith, and S. E. Reymer, Characteristics of the Ionosphere and Their Application to Radio Transmission, *Proc. I.R.E.*, **25**, 823–840, July, 1937.

[4] Kenrick, S. W., A. M. Braaten, and J. General, The Relation between Radio-transmission Path and Magnetic Storm Effects, *Proc. I.R.E.*, **26**, 831–847, July, 1938; Dellinger, J. H., Sudden Disturbances of the Ionosphere, *Proc. I.R.E.*, **25**, 1253–1290, October, 1937.

heights, critical frequencies, and ionosphere disturbances, reference should be made to the publications of these organizations.[1]

7. Fading. Fading, according to the standards of the I.R.E., "is the variation in intensity of radio signals resulting from changes in the transmission medium." Fading may conveniently be divided into two types. In the first the radio wave reaches the receiver antenna over a single path. In the second the received wave is made up of two or more components traveling over paths of different lengths.

In *single-path* transmission the received field depends directly upon the properties of the transmission medium. In such transmission, fading is likely to consist of slow variations, and the distortion in the received wave is negligible.

In *multipath* transmission small changes in the length of one of the transmission paths may have a considerable effect on the strength of the received signal. Such fading, therefore, is usually more rapid and the range of the fading greater than that in single-path transmission. In addition, as will be indicated below, the frequency and phase characteristics of the medium may be imperfect and the wave is distorted during transmission.

When the propagation is dependent upon frequency, the fading is called *selective*. The effects of *selective fading* on reception will be shown by means of the following simplified analysis:

The voltage at the receiver antenna is assumed to be made up of two components, received over different paths:

$$v = V_1 \sin \omega t + V_2 \sin (\omega t - \psi) \tag{7}$$

where v = received voltage
V_1 = amplitude of one component
V_2 = amplitude of second component
ω = angular velocity of signal $= 2\pi f$
ψ = phase angle by which second component is delayed with respect to first.

This equation may be manipulated to give

$$v = V_2\sqrt{2\frac{V_1}{V_2}\cos\psi + \left(\frac{V_1}{V_2}\right)^2 + 1}\quad \sin\left(\omega t - \tan^{-1}\frac{\sin\psi}{\frac{V_1}{V_2} + \cos\psi}\right) \tag{8}$$

If the difference in lengths between the two paths is Δs, then the phase angle ψ is

$$\psi = \frac{\omega\,\Delta s}{3 \times 10^8} \tag{9}$$

where the velocity of propagation has been assumed to be 3×10^8 meters per second. Equations (8) and (9) show that the amplitude, represented by the square root factor in Eq. (8), and the phase angle, represented by the arctangent term, of the received signal, both depend upon the frequency of the signal and the ratio of the magnitudes of the two received components.

If an a-m signal is transmitted, the phase and amplitude relationships between carrier and side bands may be seriously disturbed if the distance between successive minimum points in Fig. 16 is comparable in magnitude

[1] The Bureau of Standards has published Washington, D. C., ionosphere data monthly in the *Proc. I.R.E.* since September, 1937. The monthly *Bull. U.R.S.I.* contains a summary of data taken throughout the world.

to the width of the frequency band transmitted. This is particularly true when the signals transmitted over the two paths are of similar magnitude, *i.e.*, V_1 approximately equal to V_2. If the difference between the path lengths is known, the delay time and the frequency difference

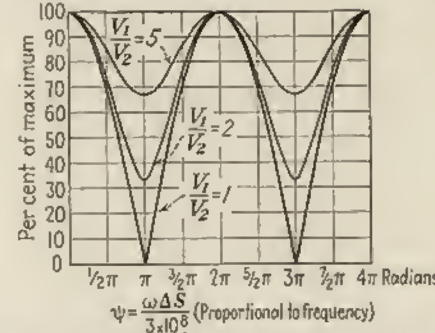


FIG. 16.—Variation in amplitude against frequency in selective fading. V_1 and V_2 are the amplitudes of the two received components.

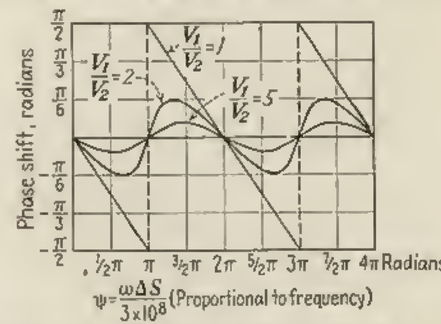


FIG. 17.—Variation in phase against frequency in selective fading. V_1 and V_2 are as in Fig. 16.

between successive minimums of Fig. 16 may be calculated from the following relations:

$$\Delta f = \frac{1}{\Delta t} = \frac{3 \times 10^8}{\Delta s} \tag{10}$$

where Δf = frequency difference between successive minimums of Fig. 16 in cycles per second
Δt = time by which one ray is delayed compared to the other in seconds
Δs = difference in path lengths in meters.

The rapidity of fading in multipath transmission depends upon the rate at which the path lengths change. A change in the difference between path lengths of a half wave length may bring about a change from a minimum to a maximum in a fade.

The selectivity or non-selectivity of fading depends upon the numerical value of Δf and the width of the band transmitted. When one of the transmission paths is by way of the ionosphere and another is the ground wave, the difference in path lengths is of the order of hundreds of kilometers, making Δf of the order of kilocycles. This is a type of selective fading common in the broadcast band at night, and it limits the primary service range of high-power broadcast stations to something between 100 and 200 km, at which distances the sky wave and ground wave have similar amplitudes. The presence of small amounts of undesired f.m. causes a particularly pernicious type of audible distortion when selective fading occurs.[1] With modern broadcast transmitters, however, the amount of residual f.m. is small. The periodicity of broadcast band fading is usually large, of the order of minutes. Such slow changes may be accommodated by conventional a-v-c circuits, provided that the minimum of the fade does not drop below the noise level and that the distortion resulting from selective fading is not excessive.

At high frequencies fading is in general more severe. Reception over a multiplicity of paths is common. The ground wave is rarely a factor in transmission, and differences between path lengths are often longer, giving values of Δf as low as a few hundred cycles.[2] Short-wave fading has many periods. As mentioned in connection with h-f propagation, seasonal, yearly, and diurnal variations take place, but in addition much shorter periods exist, some with periods less than $\frac{1}{10}$ sec., which are not readily accommodated with conventional a-v-c arrangements.

High-frequency fading is minimized through *diversity reception*. If two receiving antennas are spaced several wave lengths apart, it has been observed that the signals picked up do not fade in synchronism. Accordingly, if several antennas, normally three, spaced approximately 10 wave lengths apart, are employed, sufficient output is almost always available from at least one of the antennas to provide a useful signal. Distortion resulting from selective fading is usually worse on the poorer signals. Diversity radio-telephone systems, therefore, are commonly arranged to provide nearly all the low-frequency output voltage from the strongest signal automatically. The use of single-side-band signals also assists in avoiding distortion of this kind.

The effects of multipath transmission may be further avoided by the use of receiving antennas, directional in both the horizontal and vertical planes and aimed to pick up the strongest component of the signal. Antennas whose directivity is under the control of the operator have been developed for this purpose.[5]

[1] Bown, R., DeL. K. Martin, and R. K. Potter, Some Studies in Radio Broadcast Transmission, *Proc. I.R.E.*, **14**, 57–132, February, 1926; also Eckersley, T. L., Frequency Modulation and Distortion, *Exp. Wireless and Wireless Eng.*, **7**, 482–48[illegible], September, 1930.

[2] Potter, R. K., Transmission Characteristics of a Short-wave Telephone Circuit, *Proc. I.R.E.*, **18**, 581–648, April, 1930.

[3] Beverage, H. H., and H. O. Peterson, Diversity Receiving Systems of R.C.A. Communications, Inc., for Radio Telegraphy, *Proc. I.R.E.*, **19**, 531–561, April, 193[illegible]; Peterson, H. O., H. H. Beverage, and J. B. Moore, Diversity Telephone Receiving System of R.C.A. Communications, Inc., *Proc. I.R.E.*, **19**, 562–584, April, 1931.

[4] Friis, H. T., C. B. Feldman, and W. M. Sharpless, The Determination of the Direction of Arrival of Short Radio Waves, *Proc. I.R.E.*, **22**, 47–78, January, 1934.

[5] Friis, H. T., and C. B. Feldman, A Multiple Unit Steerable Antenna for Short-wave Reception, *Proc. I.R.E.*, **25**, 841–917, July, 1937.

As noted in connection with u-h-f propagation, the field intensity calculated from the diffraction of waves around the curved surface of the earth is increased by refraction in the lower atmosphere. Changes in the temperature gradient, therefore, are equivalent to a change in the propagation medium, and result in changes in the signal level.[1]

The approximate range of fading—maximum to minimum—which may be expected for frequencies of about 50 Mc during 1 day is shown in Fig. 18. The abscissas are given in terms of the ratio of the transmission distance to the line-of-sight distance. Since data on u-h-f fading are inadequate and, in any event, are too complex to represent in a single curve, the probable error in Fig. 18 is about ± 10 db. If the highest and lowest 5 per cent of the field strengths observed in 1 day are excepted, the range of variation is about half (measured in decibels) that shown in Fig. 18. The data have been assembled from several sources.

8. Noise. Among the factors limiting the usefulness of a received signal is noise, which may originate in any of the following places:

1. Within the receiver circuits.
2. Within the transmitter circuits.
3. Interfering signals.
4. Atmospherics (static) and man-made noise.

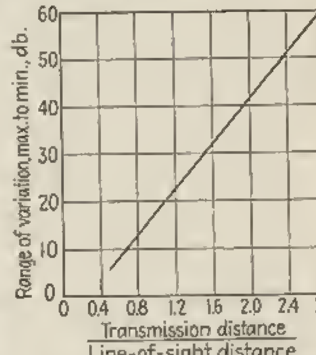


Fig. 18.—Ultra-high-frequency fading. Probable error in curve is ± 10 db.

The principal concern here will be with item 4.

Noise Wave Forms. Noise, in the broadest sense, is any type of interference. It may include continuous signals from undesired transmitters, for which the noise is contained within a known, relatively narrow, band of frequencies, and also discontinuous noises, for which the frequency band occupied is essentially infinite. Continuous disturbances are more easily studied by conventional methods; the term *noise*, therefore, is commonly restricted to discontinuous disturbances.

Discontinuous noises may be considered to be made up of sharp pulses, the frequency with which the pulses occur determining the character of the noise. If the pulses are relatively infrequent and clearly separated, the noise is said to be *impulsive*. If, on the other hand, the pulses follow each other so rapidly that they overlap and are not clearly distinguishable, then the noise is *random*. Between these two types any gradations may occur. Ignition noise is impulsive. Tube and thermal agitation noises are random.

Since the frequency spectrums of discontinuous noises are infinite in extent, their magnitudes will depend upon the band width of the device with which they are measured. Jansky[2] and others have shown that

[1] MacLean, K. G., and G. S. Wickizer, Notes on the Random Fading of 50-megacycle Signals over Non-optical Paths, *Proc. I.R.E.*, **27**, 501–506, August, 1939.

[2] Jansky, K. G., An Experimental Investigation of the Characteristics of Certain Types of Noise, *Proc. I.R.E.*, **27**, 763–769, December, 1939.

the peak, average, and effective voltages of discontinuous noises depend upon band width in the manner shown in the table:

	Impulsive	Random
Peak	Proportional to band width	Proportional to $\sqrt{\text{band width}}$
Average	Independent of band width	Proportional to $\sqrt{\text{band width}}$
Effective	Proportional to $\sqrt{\text{band width}}$	Proportional to $\sqrt{\text{band width}}$

For thermal agitation noise, Jansky found the peak to effective voltage ratio to be 4 and the average to effective voltage to be 0.85.

Atmospheric noise resembles random noise in that the individual pulses overlap. While the measured voltages increase, therefore directly as the square root of the band width, the ratios of peak to effective voltage are not constant as in the case of thermal noise.

The maximum tolerable noise level has not been measured under a wide enough variety of circumstances, up to the present time, to be able to specify its value for all conditions. The tolerable noise level depends upon a great many factors, including the following:

1. Type of service (sound, television, etc.).
2. Quality of service (excellent to poor).
3. Volume range of program material.
4. Width of frequency band.
5. Character of noise.
6. Type of modulation (amplitude or frequency; preemphasized or flat; etc.).
7. Method of measurement.

The effect of most of these factors has been only incompletely studied, particularly in so far as correlating noise levels with the physiological irritation they produce to the ear (or to the eye in television and facsimile).

Noise Measurements. The trend in noise measurements appears to be toward the use of a semipeak vacuum-tube voltmeter. The indicating instrument should have a natural period of 0.5 to 0.7 sec. and a damping factor between 10 and 100 (American Standards Association test methods) but the actual time constants of the noise meter should be determined more by the electrical circuit than by the indicating instrument. The charging time of the circuit should be approximately 10 millisec. and the discharge time approximately 600 millisec.[1] The addition of a frequency weighting network to simulate the ear's response is sometimes recommended.

Using an instrument of this type, the signal-to-noise ratio required at the output of a sound receiver for various qualities of service are, approximately, as follows:

Perfect signal	60–80 db
Excellent quality	40–50 db
For intelligibility	10–30 db

For television a peak signal to peak random noise ratio of 40 db gives a perfect picture, while a ratio of 30 db is intolerable. For single frequency noise which is a small multiple of the line frequency, the interference is barely perceptible for a signal/disturbance ratio of 50 db, while for a ratio of 35 db it is intolerable.[2]

[1] AGGERS, C. V., D. E. FOSTER, and C. S. YOUNG, Instruments and Methods of Measuring Radio Noise, *Trans. A.I.E.E.*, **59**, 178–192, March, 1940.

[2] JARVIS, R. F. J., and E. C. H. SEAMAN, The Effect of Noise and Interfering Signals on Television Transmission, *P.O.E.E. Jour.*, **32**, 193–199, October, 1939.

Atmospherics, or *static*, originates in lightning discharges. The impulses are frequent and overlap, so that the noise is more or less random, with sharp peaks exceeding the average level. Atmospheric static originates both in local storms, relatively infrequent in northern latitudes, and in more distant tropical storm centers. Static is propagated in the same manner as other radio waves; variations in distant atmospherics may often be predicted on this basis.

The signal strength from local storms varies approximately inversely as the frequency.[1] Thunderstorms and static are, of course, more intense in summer than in winter. The curves of Fig. 19 were measured near New York City, but probably are representative of most of the United States as well.

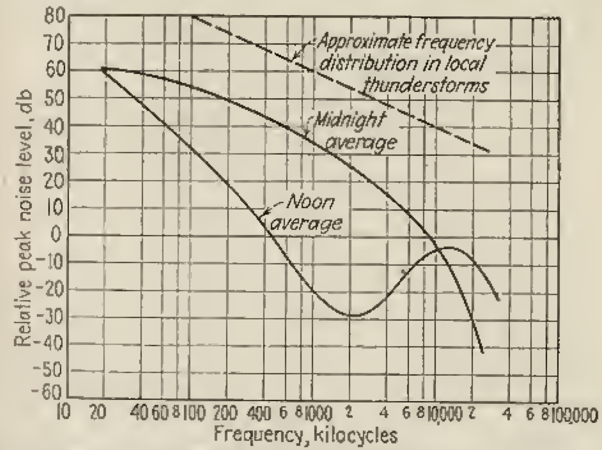


FIG. 19.—Peak energy distribution in static. Relative values are given. Absolute values depend on location and time. (*Potter, Proc. I.R.E.*, **20**, 1512, *September*, 1932.)

Since most static is of tropical origin, the lowest disturbance levels are found at distances remote from the equator, especially for low frequencies. However, at ultra-high frequencies and at broadcast frequencies during the day, for which only short-distance communication is feasible, most static is of local origin. Since local storms are infrequent in northern latitudes—about 30 per year near New York City—static causes little interference with communication above 30 Mc. At 150 Mc the peak voltage from thunderstorms has been found to vary approximately inversely as the distance, being 75 ± 10 db above 1 μv per meter when measured with a 1.5 Mc band width at a distance 1 mile from a storm.[2]

In the absence of either atmospheric or man-made static, Jansky has found that noise is still picked up by the receiver antenna, noise which he ascribes to stellar radiation. This noise, at frequencies between 9 and

[1] POTTER, R. K., An Estimate of the Frequency Distribution of Atmospheric Noise, *Proc. I.R.E.*, **20**, 1512–1518, September, 1932.

[2] SCHAFFER, J. P., and W. M. GOODALL, Peak Field Strength of Atmospherics Due to Local Thunderstorms at 150 Megacycles, *Proc. I.R.E.*, **27**, 202–207, March, 1939.

21 Mc, was found to be some 10 to 30 db above the level of thermal agitation noise in the receiver and, except for man-made noise, is the limiting noise at high frequencies a large portion of the time.[1]

Man-made noise may be generated by internal combustion-engine ignition systems, by power-line discharges, by diathermy machines, by motor-brush sparking, and by other electrical devices. Man-made noise usually reaches the receiver input in the following ways:[2]

1. Radiation or capacitance pickup by receiver antenna direct from noise source (relatively rare) or from power lines which convey noise voltages to vicinity of antenna.
2. By transmission over power lines direct to receiver.

Means for reducing such pickup are described in the literature.[3]

Typical noise voltages within the broadcast band measured at the terminals of noise-generating devices follow:[4]

TYPICAL NOISE VOLTAGES

Source	Line-to-line r-f voltage, millivolts	Line-to-ground r-f voltage, millivolts
Vacuum cleaner	3	3.5
Electric razor	40	5.6
Diathermy machine	250	37
Portable electric tool	20	26

The band width of the measuring device was presumably between 3 and 5 kc. These voltages may be reduced by the addition of simple noise-suppression filters.

During the past few years interference of radiation from physician's diathermy machines with radio communication has become objectionable at high frequencies. Jansky[5] has measured peak power levels of such interference ranging from 24 to 40 db below 1 $\mu\mu$w.

At ultra-high frequencies the most objectionable types of noise are ignition and diathermy. In some cases diathermy signal strengths in excess of 100 μv per meter are encountered in cities. The peak ignition noise produced by 90 per cent of the vehicles passing 100 ft. from an antenna 35 ft. high has been found to be less than between 9 and 20 μv per meter per kilocycle low-frequency band width. The higher value is for 40 Mc, the lower for 450 Mc. Vertical polarization appears to give ignition noise a little greater than horizontal.[6] In New York City, values varying between 1 and 40 are observed at typical antenna locations.

[1] JANSKY, K. G., Minimum Noise Levels Obtained on Short-Wave Receiving Systems, *Proc. I.R.E.*, **25**, 1517–1530, December, 1937.

[2] BLOCK, L., Radio Interference, *Philips Tech. Rev.*, **3**, 235–240, August, 1938; MERRIMAN, H. O., and F. G. NIXON, Radio Interference—Investigation, Suppression and Control, *Proc. I.R.E.*, **27**, 16–21, January, 1939; AGGERS, C. V., Methods of Controlling Radio Interference, *Trans. A.I.E.E.*, **59**, 193–201, April, 1940.

[3] For example, V. D. Landon and J. D. Reid, A New Antenna System for Noise Reduction, *Proc. I.R.E.*, **27**, 188–191, March, 1939. A survey of noise-reducing systems is contained in the "Radio Noise Reduction Handbook," 1938, Radio, Ltd., Santa Barbara, Calif.

[4] HASKINS, R. L., and C. W. METCALF, Station Coverage, *Communications*, **18**, 23–26, April, 1938; British Standard Specification for Limits of Radio Interference, *Brit. Standards Inst., Spec.* No. 800, 1937.

[5] JANSKY, *loc. cit.*

[6] GEORGE, R. W., Field Strength of Motor Car Ignition between 40 and 450 Megacycles, *Proc. I.R.E.*, **28**, 409–412, September, 1940.

TABLE II.—SUMMARY OF FREQUENCY ALLOCATIONS IN UNITED STATES
For details of frequency allocation in United States, see General Rules and Regulations, FCC, Part 2, U. S. Government Printing Office, and Order no. 67, FCC.

Frequency channels, kilocycles	Allocation	Frequency channels, kilocycles	Allocation
10–103	Fixed, government	11,010–11,685	Ship telegraph, maritime calling, government, coastal telegraph, fixed, aviation, miscellaneous
103–141	Coastal telegraph, government	11,710–11,890	International broadcast, government
143–193	Maritime calling, ship telegraph, fixed and coastal telegraph. (190 kc to state police and government)	11,910–13,990	Aviation, fixed, government, ship telegraph, coastal telegraph, miscellaneous
194–391	Government, fixed, airport, aircraft (375 kc to direction finding)	14,005–14,395	Amateur
392–548	Coastal telegraph, government, ship telegraph, aircraft, intership phone (500 kc to maritime calling and government)	14,410–15,085	Fixed
550–1,600	Broadcasting (1,592 to Alaska services)	15,110–15,330	International broadcast, government
1,600–1,712	Geophysical, relay, police, government, experimental, marine fire, aviation, motion picture	15,355–17,740	Fixed, government, aviation, ship and coastal telegraph, miscellaneous
1,716–2,004	Amateur	17,660–17,840	International broadcast
2,004–2,500	Experimental visual and relay broadcast, police, government, ship harbor, fixed, miscellaneous	17,860–21,440	Fixed, government, aviation
2,504–3,497.5	Coastal harbor, government, aviation, fixed, miscellaneous	21,460–21,650	International broadcast, government
3,500–4,000	Amateur	21,650–23,175	Coastal telegraph, government, ship telegraph, miscellaneous
4,005–6,000	Government, aviation, fixed	23,200–25,000	Aviation, government, miscellaneous
6,020–6,190	International broadcast, government	25,025–26,975	Broadcast, government
6,200–6,990	Coastal telegraph and phone, government, fixed, miscellaneous	27,000–27,975	Government, general communication
7,000–7,300	Amateur	28,000–30,000	Amateur
7,305–9,490	Government, fixed, aviation, ship telegraph, coastal telegraph, miscellaneous	30,000–42,000	Police, government, relay broadcast, coastal and ship harbor, miscellaneous
9,510–9,690	International broadcast	42,000–50,000	Broadcast and educational
9,710–11,000	Government, fixed aviation	50,000–56,000	Television, fixed
		56,000–60,000	Amateur
		60,000–112,000	Government, television
		112,000–116,000	Amateur
		116,110–139,960	Broadcast, government, aviation, police, miscellaneous
		140,100–143,880	Aviation
		144,000–400,000	Government, television, fixed
		400,000–401,000	Amateur
		401,000 and above	Experimental

HIGH-FREQUENCY USAGE

9. Frequency Allocation. The most recent international conferences at which frequency allocations have been agreed upon were held at Washington in 1927, Madrid in 1932, and Cairo in 1938. Frequency allocation between North American countries has been decided at the North American radio conferences, at Mexico City in 1933 and at Havana in 1937. The North American Regional Agreement, reached in Havana, went into effect in 1940, requiring a number of minor changes, particularly in broadcast-frequency allocations.

Technical information is supplied by the International Radio Consulting Committee (C.C.I.R.) which has had recent meetings at Lisbon (1934) and Bucharest (1937), and the International Scientific Radio Union (U.R.S.I.) with recent General Assemblies at London (1934) and Venice (1938).

In the United States licensing, regulation, and allocation are handled by the FCC, which succeeded the Federal Radio Commission under the Communications Act of 1934, as amended in 1937.

The general plan of frequency allocation in the United States is shown in Table II. The European system differs in several respects, one being the use of the 200- to 400-kc channel for broadcasting. In allocating channels to stations in the United States, care is taken not to permit operation at times and frequencies for which propagation is such that interference with foreign stations would be caused.

TABLE III.—BAND-WIDTH REQUIREMENTS

Service	Band Width
Continuous-wave telegraphy	Equals the telegraph speed in bauds (1 baud = 0.8 words per minute, for a telegraph code having 8 dots or blanks per letter) for the fundamental, 3 times this for 3d harmonic, etc.
Modulated continuous-wave telegraphy	Add twice the modulation frequency to the above
Commercial telephony	6 to 8 kc.
Broadcasting	10 to 30 kc.
Television	6,000 kc for both sound and sight, R.M.A. system (1940)
Wide-band f.m.	200 kc.
Facsimile	Approximately equals (number of picture components) × (time of transmission in seconds) + twice the subcarrier frequency, if used

10. Selection of Best Operating Frequency. One of the problems in the design of a radio system is the selection of the best frequency for carrying on communication. The first step is a decision as to the range within which the frequency to be selected will fall. Table V, which tabulates frequency ranges normally employed for communication over specified distances, will be found useful for this purpose. Other considerations besides purely technical ones will often dictate the choice of frequency range. The cost of transmitters, antennas, and receivers, as well as the availability of channels, must be studied. When a choice between the l-f and h-f ranges is indicated by the table, high frequencies should be employed, since the cost of such service will normally be lower. Exception must be made for periods during which the h-f band is useless for long-distance service because of magnetic storms, in which case low frequencies are sometimes used. Table V is intended to cover only the

TABLE IV.—FREQUENCY TOLERANCES

As specified by Cairo Convention (1938). See General Rules and Regulations, FCC, Part 2, U. S. Government Printing Office

Frequency Band	Tolerance
A. 10 to 550 kc:	
a. Fixed, land, mobile (other than under *b*, below) stations	0.1%
b. Mobile stations between 110–160 kc and 365–515 kc and aircraft	0.3%
B. 550–1,500 kc broadcasting	20 cycles
C. 1,500–6,000 kc:	
a. Fixed stations	0.01%
b. Land stations	0.02%
c. Mobile stations:	
1. 1,500–4,000 kc; 4,115–4,165 kc; 5,500–5,550 kc	0.05%
2. 4,000–6,000 kc	0.02%
d. Aircraft stations	0.025%
e. Broadcasting	0.005%
D. 6,000–30,000 kc:	
a. Fixed	0.01%
b. Land	0.02%
c. Mobile:	
6,200–6,250 kc; 8,230–8,330 kc; 11,000–11,000 kc; 12,340–12,500 kc; 16,460–16,660 kc; 22,000–22,200 kc	0.05%
Other frequencies	0.02%
d. Aircraft	0.025%
e. Broadcasting	0.005%

normal situations. In unusual circumstances other frequency ranges might be used in addition to those specified. For example, for com-

TABLE V.—SELECTION OF FREQUENCY RANGE

Distance	Frequency Ranges Normally Used (Nomenclature as in Table I)
Local (less than 15 miles)	Any
Short (15 to 150 miles)	1. Low (ground wave) 2. Broadcast (ground wave) 3. Ultra-high frequency
Medium (150 to 1,500 miles)	1. Low (ground wave) 2. Broadcast (sky wave) 3. High (sky wave)
Long (greater than 1,500 miles)	1. High (sky wave) 2. Low (sky wave)

munication between two airplanes, both flying at high altitudes, ultra-high frequencies may be used for medium distance communication.

After the frequency range has been decided, the problem of determining the most desirable frequency within that range remains. The conditions required are as follows: a specified signal/noise ratio; a minimum transmitter power; and a minimum of fading (especially of selective fading). The sky wave and ground wave should preferably have appreciably different amplitudes. Consideration must, of course, be given to the availability of channels.

The problem is divided into ground-wave and sky-wave transmission. Ultra-high frequencies will be considered separately.

When the transmission is to take place by means of the ground wave, the smallest attenuation in the earth, corresponding to maximum received field strength and a minimum of selective fading, will be obtained through the use of the lowest frequency in the range selected. On the other hand, the noise level increases at lower frequencies. An optimum frequency exists, therefore, for a given transmitter power which will result in the greatest transmission range. The use of a frequency somewhat lower than that for which the ground wave and the quasi-maximum sky wave are equal at the receiver is desirable.[1] The size and cost of the antenna structure should also be considered in the choice of frequency.

Sky-wave transmission depends upon the changeable characteristics of the ionosphere (discussed previously). The best signal strength and the minimum noise level, normally, are obtained for a frequency just below the "maximum usable frequency," as given in Figs. 6 and 7.

At ultra-high frequencies, for transmission within the line-of-sight distance, the highest frequency at which sufficient power for reliable communication can be generated should be employed. The signal/noise ratio for transmission within the line of sight varies approximately as f^n, where, if vertical antenna directivity is assumed at both receiver and transmitter and the transmitter output falls off with frequency, the exponent n probably lies somewhere between 2 and 5.

Beyond the line of sight the decrease in signal strength at ultra-high frequencies is more rapid the higher the frequency. An optimum frequency exists, therefore, which will provide the best signal/noise ratio at the receiver under a given set of conditions. If transmission far beyond the line of sight is required, then lower frequencies will be favored, between 35 and 45 Mc. If, however, transmission only slightly beyond optical distances is necessary, the use of higher frequencies is indicated. Usually, however, the optimum conditions are not critical, and considerable deviation from the optimum frequency is possible without too great an effect upon transmission.

TECHNICAL FEATURES OF H-F TRANSMITTERS

In this section the design of transmitters for use at frequencies above 2 Mc is considered, emphasizing those features in which h-f transmitters differ from low.

11. High-frequency Transmitter Requirements. No more can be given here than a list of the points which must be considered in the preparation of specifications and design of h-f transmitters:

1. Cost; size; weight.
2. Reliability; maintenance difficulties.
3. Power output; efficiency.
4. Fidelity; noise level.
5. Frequency range; variable or fixed frequency operation; frequency stability.
6. Antenna termination.
7. Power supply availability.
8. Type of modulation (a.m., f.m., telegraph, etc.) and class of service (telegraph, telephone, television, etc.; police, aircraft, army, amateur, etc.).

A typical h-f transmitter consists of a stable oscillator circuit (crystal, if for fixed frequency operation; master-oscillator, if variable frequency operation is necessary) followed by frequency multiplier and amplifier stages to raise the oscillator frequency and power to the desired level.

[1] A group of curves from which the approximate optimum frequency may be found (taking account of noise levels and fading) is published in the Report of Committee on Radio Propagation Data, *Proc. I.R.E.*, **21**, 1119–1438, October, 1933.

Amplitude modulation is most often accomplished by variation of one of the electrode voltages of the final amplifier stage. Frequency-modulation transmitters are usually modulated at a low level.

12. Considerations in the Design of Equipment. At high frequencies increased attention must be given to coupling between circuit elements, particularly through capacitances and mutual impedances in the ground circuits and leads. Amplifier units may frequently be better arranged by the use of link circuits. Capacitative coupling may be avoided, at the expense of increased capacitance to earth, by the use of static shields.

By-pass condensers pass through resonance, normally, at intermediate high frequencies. If effective by-passing is required, the impedance

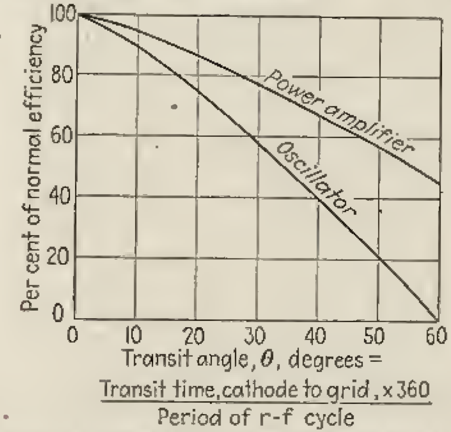


FIG. 20.—Variation in amplifier and oscillator efficiency against transit angle of electrons. (*Wagener, Proc. I.R.E.*, **26**, 401, *April*, 1938.)

of the by-pass capacitors at the operating frequency must be known. Frequently a lower impedance can be obtained by the use of a smaller capacitor, *i.e.*, choosing a capacitor size which more nearly resonates with the inductance of its leads.

Insulating materials are available with very low losses. Polystyrene (under its various names) has remarkable h-f properties but fails mechanically at temperatures above about 70°C. For the higher temperatures encountered in transmitters the use of one of the ceramic materials, or Mycalex, is preferable.

13. Power Amplifier Design. Up to a certain frequency (between 1.5 and 500 Mc, the exact frequency depending upon the tube used) the determination of vacuum-tube operation conditions is the same as that described in the section on Amplifiers. Beyond this frequency, transit time and dielectric losses increase, and the tube may become an appreciable portion of a wave length in dimensions, such that the efficiency drops and grid driving power requirements are increased.[1] Oscillator efficiency

[1] WAGENER, W. G., The Developmental Problems and Operating Characteristics of Two New Ultra-high-frequency Triodes. *Proc. I.R.E.*, **26**, 401–414, April, 1938; SAMUEL,

is lower than amplifier efficiency, as Fig. 20 indicates. Data on individual tubes should be obtained from the manufacturer.

Class C saturated amplifier stages are in widest use as h-f amplifiers, with a.m. accomplished at the highest power level. In some types, however, modulation at a low level, followed by class B linear amplifiers, is employed, the reduction in tube complement and high-power modulation equipment offsetting the increase in adjustment difficulties and the possible loss in efficiency. An example of the latter is a single-side-band transmitter, in which it is usually simpler to accomplish the relatively complicated modulation process at a low level and then to amplify the modulated signal in linear amplifiers. High efficiency linear amplifiers[1] are not widely used at higher frequencies, although amateur and experimental installations have been made.

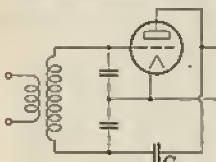


FIG. 21.—Imperfect grid neutralizing circuit.

Frequency multiplication is required when an oscillator of requisite stability cannot be built at the output frequency, where a wide range of output frequencies is to be covered and the oscillator frequency range is limited, or where neutralization must be avoided. Crystals are now (1940) available up to 20 Mc with temperature coefficients better than 2 ppm per °C. between 20°C. and 60°C., so that less frequency multiplication need be used, even in crystal-controlled transmitters. Higher frequency crystals are available but with poorer temperature coefficients. Frequency multipliers have the further advantage that neutralizing is not required, since output and input circuits operate at different frequencies and instability is less likely to occur. Multiplier efficiency is lower than that of amplifiers, however, and their use is avoided in high-power stages.

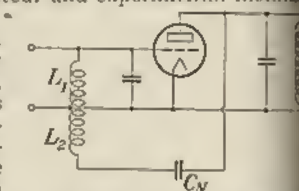


FIG. 22.—Circuit capable of satisfactory neutralizing if unity coupling exists between halves of grid coil.

If necessary, the frequency multiplying stage may be modulated, although higher distortion will be encountered. A frequency multiplier may also be used as a linear amplifier of a-m waves, although distortion is likely to be excessive and efficiency low.[2]

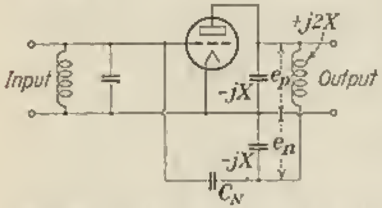


FIG. 23.—Plate neutralizing circuit. Capable of good neutralizing, if plate impedance of tube can be neglected.

A. L., A Negative Grid Triode Oscillator and Amplifier for Ultra-high Frequencies, *Proc. I.R.E.*, **25**, 1243–1252, October, 1937.

[1] Such as the circuit of W. H. Doherty, A New High Efficiency Power Amplifier for Modulated Waves, *Proc. I.R.E.*, **24**, 1163–1182, September, 1936.

[2] IVANOV, A. B., Amplitude Modulation of Frequency Multiplying Stage (in Russian), *Izvestia Elect. Slab. Toka.*, No. 7, pp. 34–38, 1938.

An f-m signal is not distorted in traversing a multiplier stage; in so doing the frequency deviation is multiplied by the same factor as the fundamental.

Neutralization. It is not always appreciated that the object of neutralization is actually twofold, the prevention, first, of amplifier instability, and, second, of reaction of the amplifier on preceding stages.[1] Many neutralizing circuits are incapable of accomplishing both of these aims. An example of a neutralizing circuit which does not provide perfect neutralization is shown in Fig. 21. This circuit, while it can be adjusted to prevent reaction on the driving stage, is degenerative. Figure 22 is better. If $C_n = C_{gp}$ and L_1 is approximately equal to L_2, and the coupling between L_1 and L_2 is $M = L_1$, good neutralization results. The same is true if the input is inductively coupled to the preceding stage. A satisfactory plate-neutralizing circuit is shown in Fig. 23. The output should not be coupled inductively to the plate tank coil but should be taken between plate and cathode, as shown. The plate tank circuit, comprising the two capacitors and the coil, may be considered equivalent to a half-wave transmission line, whose input and output voltages, e_p and e_n, are 180 deg. out of phase. There is, of course, an infinite number of circuits equivalent to a half-wave line, any of which may be employed in neutralizing circuits. Another equivalent, in a grid-neutralizing arrangement, is shown in Fig. 24. In push-pull circuits good neutralizing may be obtained by the use of cross-connected capacitors (Fig. 25).

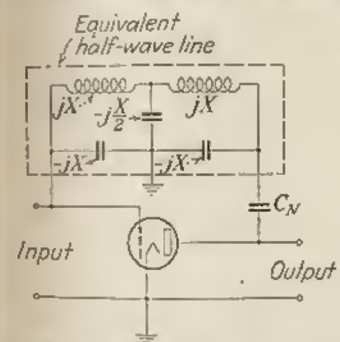


FIG. 24.—Neutralization with half-wave line.

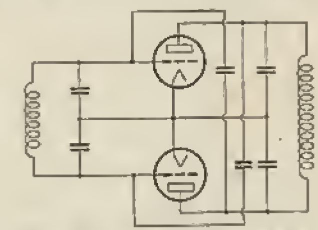


FIG. 25.—Push-pull neutralizing circuit.

Neutralizing is complicated further by the decrease in the grid-plate impedance and the increase in the lead impedances, particularly that of the cathode to ground, as the frequency is increased. The decrease in the grid-plate impedance can be mitigated by more careful adjustment of the neutralizing circuit. The increase in the lead impedances can be compensated by the use of more complicated neutralizing circuits.

The most troublesome of the lead impedances, that of the cathode to ground, is frequently eliminated by tuning the filament leads to series resonance and thereby effectively bringing the cathode to ground potential. This may be accomplished by the use of series condensers (Fig. 26*A*) or by making the heater-ground leads one-half wave length long (Fig. 26*B*).

[1] DOHERTY, W. H., Neutralization of R-F Power Amplifiers, *Pick-ups*, pp. 3–5, 21–23, December, 1939; MOMOTUKA, K., and K. SAZI, Considerations of Neutralizing Methods, *Nippon Elect. Comm. Eng.* No. 14, pp. 518–519, December, 1938.

One of the oldest neutralizing circuits[1] has recently been resurrected. It consists simply in tuning the grid-plate capacitance to parallel resonance by the addition of a shunt coil (Fig. 27) and thus raising the grid-plate impedance to a high value. The Q of the coil should be large enough so that the resonant impedance, XQ, where X is the reactance

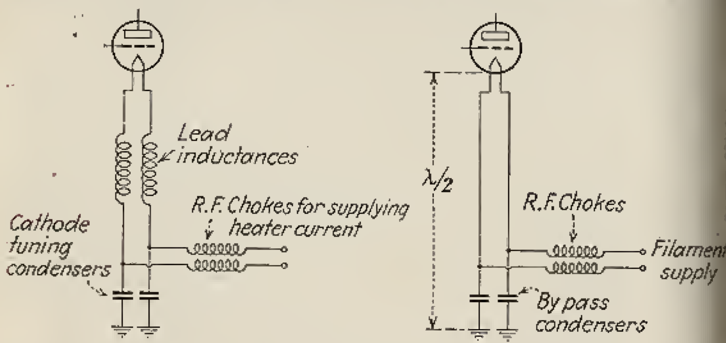


FIG. 26.—Methods of tuning filament leads to return filament to ground potential. At A, leads are tuned with condensers. At B, a half-wave line is used.

of the grid-plate capacitance at the resonant frequency, is much higher than the impedance of the grid-cathode circuit. At u.h.f. the shunt coil and blocking capacitor may be replaced by an open-circuited line slightly less than a half wave length long.

All the above neutralizing circuits suffer the disadvantage that the interelectrode capacitance is compensated exactly at only one frequency, usually the carrier frequency. The circuit must also be designed to be well balanced at the side-band frequencies in stages through which modulated energy is passing. At frequencies remote from the carrier, the stage will be under- or over-neutralized, either of which is usually undesirable. This lack of neutralization at frequencies remote from the carrier is a frequent cause of parasites, as is noted below. In addition, if the transmitter must be tuned rapidly over a wide range of frequencies, it is desirable to employ a wide-band neutralization circuit, requiring no readjustment of neutralizing as the transmitter frequency is varied. In wide-band, or "complete," neutralization, this is accomplished by duplicating each part of the tube structure, including the lead inductances, in a similar element in the neutralizing bridge. The bridge will then be balanced at all frequencies. A simple mechanical

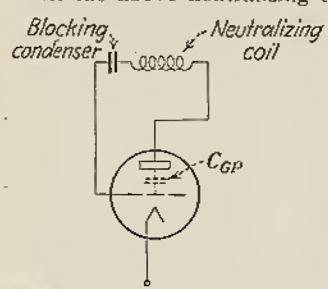


FIG. 27.—"Coil" or "shunt" neutralizing.

[1] NICHOLS, H. W., U. S. Patent 1325879.

arrangement[1] which accomplishes this for water-cooled tubes in a push-pull circuit is shown in Fig. 28A, while Fig. 28B shows the equivalent bridge circuit. The series cathode condensers, C_k, C_k', which tune the cathode leads, are used only when it is desired to keep the grid and plate tank voltages in phase with the corresponding electrode voltages. The

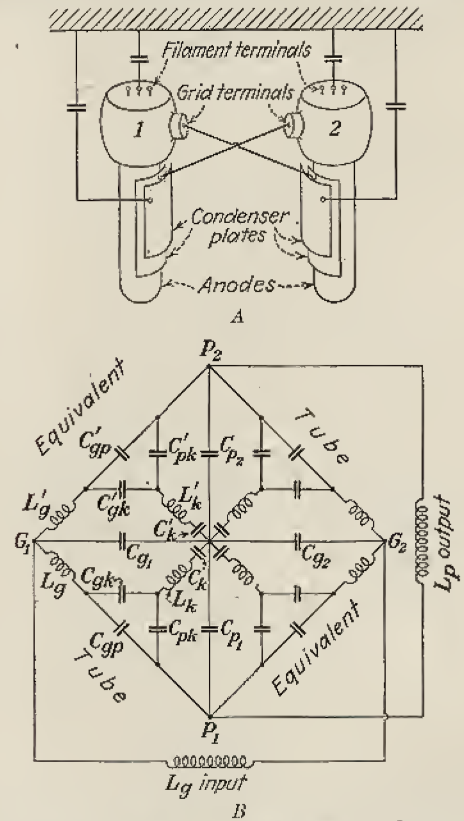


FIG. 28.—Wide-band, "complete," neutralizing. A shows the mechanical arrangement; B, the equivalent circuit. The condenser plates are adjustable.

anode connection in water-cooled tubes has no appreciable inductance, and it need not be compensated.

A recently developed circuit which requires little or no neutralizing[2] is shown in Fig. 29. The arrangement is inherently degenerative. If the

[1] BUSCHBECK, W., U. S. Patent 2002338.

[2] HAYES, L. W., and B. N. MACLARTY, The Empire Service Broadcasting Station at Daventry, *Jour. I.E.E.*, **85**, 321–369, September, 1939.

grids are grounded effectively, they screen the grid from the cathode circuit and no reaction on the exciter is possible. One disadvantage is that provision must be made for operating the filaments at high r-f potential. If the screening effect of the grid is incomplete, neutralizing condensers, shown dotted in the figure, should be used. In general, very small neutralizing condensers suffice, and the minimum plate tank reactance is cut nearly in half, as compared with conventional circuits.

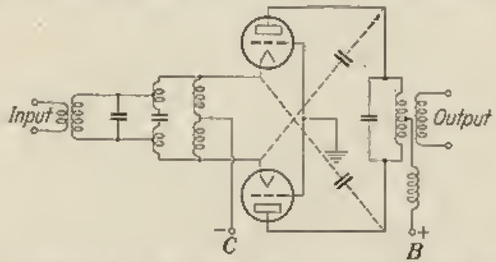


FIG. 29.—Grounded-grid amplifier. The neutralizing condensers, shown dotted, are required only when the shielding provided by the grid between cathode and plate is incomplete. The grid leads may be tuned to series resonance to improve the grounding. (*Hayes and MacLarty, Jour. I.E.E.*, **85**, 321, *September*, 1939.)

The commonest expedient for avoiding neutralizing difficulties is the use of screen-grid tubes. In the past the efficiency of such tubes has been low because of the high screen current required, but this is now avoided by the use of the beam principle. Many transmitter tubes of this type are available, including the 832 u-h-f push-pull tube and a 20-kw television tube.[1]

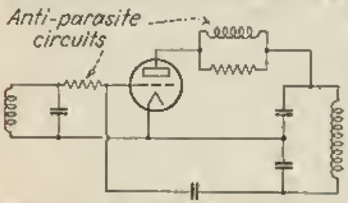


FIG. 30.—Parasitic suppression resistors and choke.

Parasites are undesired oscillations in transmitters,[2] to which h-f high-power transmitters are particularly prone. The presence of parasites is usually revealed by low efficiency, excessive distortion, high r-f voltages, overheating of components, or a combination of these effects. An absorption wavemeter may be used to locate the parasite and to determine its frequency.

A common cause of parasitic oscillation in single-ended r-f amplifiers is the existence of a very h f mode of oscillation whose frequency is determined by the lead inductances and the interelectrode capacitances. The cure is the insertion of resistance in series with the grid circuit and a small choke—a few microhenrys is common—in parallel with a resistor in the plate circuit, as in Fig. 30. The impedance of the parasitic suppression elements should not be so large that they interfere with the desired operation of the amplifier.

[1] HAEFF, A. V., L. S. NERGAARD, W. G. WAGENER, P. D. ZOTTU, R. B. AYER, and H. E. GIHRING, Development of a 20-kilowatt Ultra-high Frequency Tetrode for Television Service, Abstract in *Proc. I.R.E.*, **27**, 610–611, September, 1939; *Elect. Eng.*, **59**, 107, March, 1940.

[2] FYLER, G. W., Parasites and Instability in Radio Transmitters, *Proc. I.R.E.*, **23**, 985–1012, September, 1935.

Push-pull amplifiers frequently develop parasites of the type indicated above, the amplifier oscillating in push-pull at a higher frequency than normal. Wide-band neutralizing is very effective in suppressing such parasites. A push-pull amplifier may also oscillate in parallel, in which the two tubes are connected in parallel by the tuning condensers at h.f. or by the tuning inductances at l.f. The solution indicated in Fig. 30 applies here also. If oscillation occurs at l.f., the r-f feed chokes should be investigated as likely causes. It is usually preferable to center-tap the tank condenser, rather than the tank coil, to obtain the ground point.

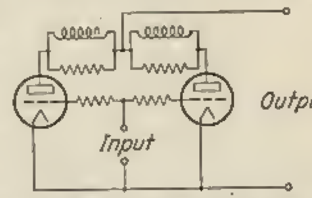


FIG. 31.—Parasitic suppression resistors and chokes for parallel-connected tubes.

Similarly, tubes operated in parallel often oscillate in push-pull at a frequency determined by the interelectrode capacitances and the leads connecting the tubes together. The cure is to insert isolating resistors and chokes between the tubes, as indicated in Fig. 31.

Many high-power tubes have an appreciable negative resistance over a portion of their grid characteristics which may cause a dynatron oscillation in some part of the grid circuit. The grid resistor of Fig. 30, plus parallel loading from grid to ground, is usually effective in suppressing such parasites. In some transmitters diode load tubes have been connected between grid and cathode for the same purpose.

In screen-grid tubes the screen by-pass connection should be as short as possible, to avoid the introduction of inductive reactance at a very h.f.

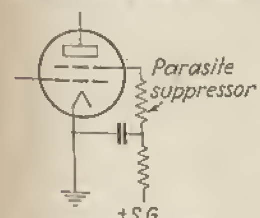


FIG. 32.—Parasitic suppression resistor in screen lead.

Beam power tubes, with their very high power gains, are prone to oscillations from this source. A small non-inductive resistor inserted directly in series with the screen (Fig. 32) is useful in suppressing such parasites.

14. Interstage coupling circuits frequently employed in h-f transmitters are illustrated in Fig. 33. The arrangements of A, B, and C claim attention because of their simplicity. Inductive coupling, as at D and E, is advantageous because no current flows in the ground impedance to add to feedback problems. Their use is desirable when the two amplifier stages are separated by considerable distances. The feeding line in such cases is sometimes made coaxial. If purely inductive coupling is desired, a static shield may be interposed between any two of the coils of D or E. F, G, and H are three arrangements for coupling transmission line tank circuits. I, J, and K are used to couple from a single-ended stage to a push-pull stage.

15. Push-pull versus Single-ended Circuits. The principal advantages of the push-pull connection are as follows: simpler neutralizing, cancellation (at least in part) of even harmonics, and simpler by-passing

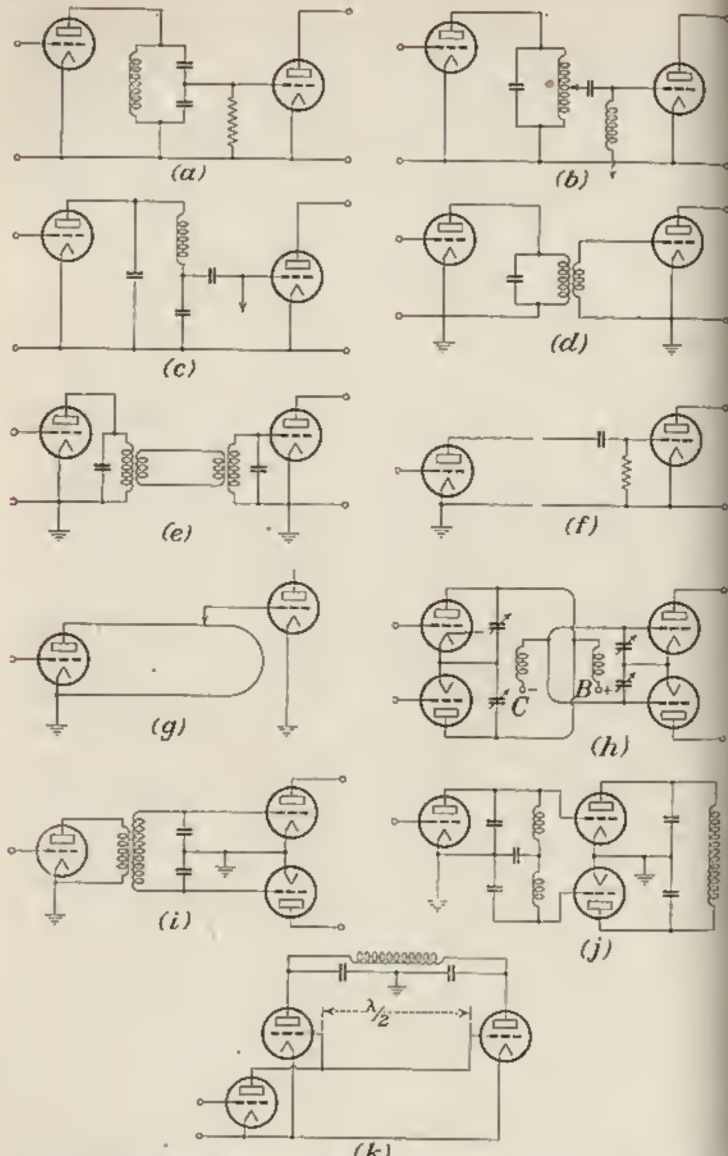


FIG. 33.—Interstage coupling circuits. *a*, *b*, and *c*, direct coupling; *d* and *e* inductive coupling; *f*, *g*, and *h*, transmission-line coupling; *i*, *j*, and *k*, single-ended to push-pull.

problems. The advantages of single-ended stages are as follows: lower tank voltages, generally lower tube cost per kilowatt, and simpler connection to grounded transmission lines and antennas.

Tank circuits may be either lumped or distributed. At higher frequencies transmission-line elements are employed as tank circuits, as in Fig. 34 in which a parallel-wire short-circuited line is shown for a

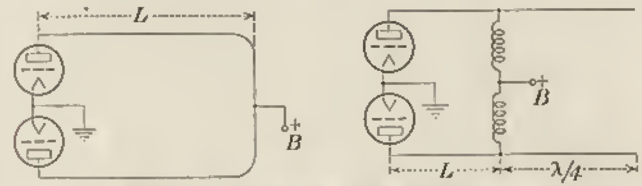


FIG. 34.—Transmission-line tank circuits.

push-pull stage. If the tube capacitance were zero, the line length would be an odd multiple of a quarter wave in length. Since the tube capacitance is never zero, the line length is shorter than a quarter wave. The length of line necessary to resonate with a given tube reactance is shown in Fig. 35. Occasionally an open-circuited line is used for a tank circuit, as in Fig. 34, in which case the line is a quarter wave length longer.

In a wide-band amplifier the power output of an amplifier is limited by the required band width and by the minimum tank capacitance. Thus, if the

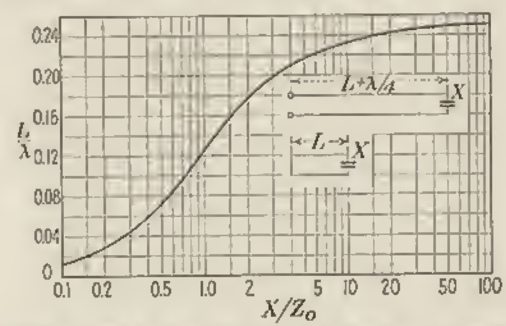


FIG. 35.—Length of line to resonate with condenser. X = reactance of condenser; Z_0 = characteristic impedance of transmission line.

tube has a maximum emission current I_m, a maximum modulation frequency, f_m, and a minimum output capacitance C, then the power output of the tube is proportional to I_m^2/Cf_m.

ULTRA-HIGH-FREQUENCY TRANSMITTERS

16. Ultra-high-frequency Circuits. Both the tube and the circuit problems become increasingly difficult as the frequency is raised, but during the past few years several developments have appeared to reduce the magnitude of some of these. The product of maximum realizable power by frequency is constantly being raised.

Up to between 250 and 500 Mc, more or less conventional tubes and circuits are employed in transmitters, with master oscillators at the output frequency, or at subharmonics of the output frequency, driving one or more multiplier or amplifier tubes. Amplifier-tube construction for ultra-high frequencies has vastly improved during the period between 1935 and 1940. For example, a screen-grid tube is available with an output of 20 kw at 120 Mc with a band width of 2 Mc.

Several circuit arrangements are in use. One scheme employs a crystal or other stable oscillator at a relatively l.f.—10 to 30 Mc—followed by multipliers and amplifier stages. By the use of multipliers in all amplifiers except, perhaps, the output stage, few neutralizing circuits are needed. Alternatively, an attempt may be made to secure higher efficiency in the amplifier stages by completing the multiplication at a low level and accomplishing most of the power amplification at the output frequency. Such designs are facilitated by the availability of u-h-f tetrodes, which need no neutralizing, such as the 832 and the 20-kw tube mentioned previously. Another circuit arrangement employs a stable oscillator circuit operating at the output frequency, sometimes of appreciable power output, followed by one or two stages of amplification. Such a system has the advantage of simplicity. A transmission-line stabilized oscillator is common in transmitters of this type.

Whenever possible, a.m. of u-h-f transmitters is effected in the final amplifier, either by grid or plate modulation, avoiding the use of linear amplifiers. Frequency modulation, on the other hand, is most easily accomplished at low level, since the amplifiers operate saturated and present no particular adjustment problems.

The most common type of tank and coupling circuit employed at u.h.f. is shown in Fig. 33*H*. The opportunities for the exercise of mechanical ingenuity in the arrangement of the circuits are plentiful. If the frequency range to be covered is wide, the effective Q of the tank circuit can be maintained substantially constant by simultaneous variation of the line length and the tuning capacitances. It should be noted that in h-f transmitters the effective Q is frequently higher than desired, and every attempt is made to reduce the tuning capacitance in order to reduce the tank losses and to obtain the required band width. Figure 35 may be used to calculate the length of line for any capacitance and frequency, although this may be altered somewhat by the effect of the coupled circuit. The coupling between the two lines is a combination of distributed mutual inductance and capacitance, unless a static shield is interposed between the two to eliminate the capacitative coupling. The use of such a shield may be desirable in some instances to improve the balance between the two sides of a push-pull stage.

17. Tubes for U.H.F. Triodes and tetrodes are now (1941) available which will furnish appreciable power output at centimeter wave lengths.[1] At 100 cm, for example, the 832 type will furnish 20 watts, while the 887 and 888 water-cooled tubes will furnish 350 watts. The 1628, with its double lead construction and dissipation of 40 watts, may be used as

[1] Wagener, W. G., The Developmental Problems and Operating Characteristics of Two New Ultra-high-frequency Triodes, *Proc. I.R.E.*, **26**, 401–414, April, 1938; Samuel, A. L., A Negative Grid Triode Oscillator and Amplifier for Ultra-high Frequencies, *Proc. I.R.E.*, **25**, 1243–1252, October, 1937.

low as 60 cm at full rating. Small Western Electric triodes are available with useful output down to 10 cm. The simpler centimeter-wave transmitters employ tubes such as these as self-excited oscillators.[1]

The most promising development in centimeter-wave transmitters, however, is the electron beam tube principle, exemplified by the Klystron of the Varians'. These are described in the section on Vacuum Tubes.

18. Variable ("Floating," "Controlled," or "Hapug") Carrier Transmitters. *Variable Carrier Operation.* The carrier level is made dependent in some way upon the amplitude of the modulating voltage.[2] There are many systems of this type. In one the carrier level is made to vary instantaneously according to the amplitude of the modulation voltage, and the percentage modulation, therefore, is kept constant at every instant.

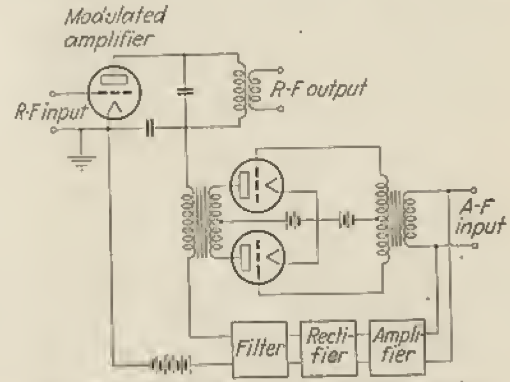


Fig. 36.—Variable carrier transmitter circuit.

In an alternative system the carrier level follows the syllabic variations, *i.e.*, the average peaks in the modulation voltage whose frequency is of the order of 1 to 10 c.p.s. A common practice is to make the minimum carrier level about 25 per cent of the maximum. The filter constants must be so chosen that the plate voltage of the r-f tube rises rapidly, or the transmitter will overmodulate on the leading edge of steep wave fronts. The current R.M.A. television signal is an asymmetric-sideband system with a variable carrier, whose amplitude is proportional to the average picture illumination.

The advantages of variable carrier are as follows:

1. Reduced noise level: in the absence of strong carrier the noise components beat only with each other and are reduced in amplitude.

[1] For examples of such self-excited centimeter-wave oscillators, see W. L. Barrow, Oscillator for Ultra-high Frequencies, *Rev. Sci. Inst.*, **9**, 170–174, June, 1938; O. Groos, "Einführung in Theorie und Technik der Decimeterwellen," S. Hirzel, Leipzig, 1937; "Radio Amateur's Handbook," 17th ed., American Radio Relay League, Hartford, Conn., 1939; "Radio Handbook," ed. by W. W. Smith, 6th ed., Radio, Ltd., Santa Barbara, Calif., 1939.

[2] Harbich, H., F. Gerth, and L. Pungs, Modulation with Variable Carrier Amplitude, *Hochf. u. Elek.*, **5**, 141–147, May, 1936; Fyler, G. W., Phone Transmission with Voice Controlled Carrier Power, *QST*, **19**, 9–12, January, 1935.

2. Reduced power consumption: since the average percentage modulation in typical audio program material is very low, the average power output is correspondingly low.
3. In television the d-c picture component is more easily transmitted.

The disadvantages are as follows:

1. Conventional receiver a-v-c circuits, which operate in proportion to the received carrier level, will not function properly.

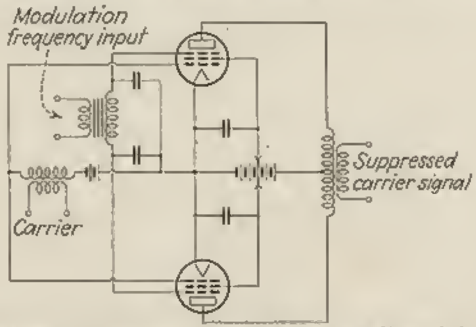


FIG. 37.—Balanced modulator. The balance is adjusted by varying the screen bias voltages.

2. Distortion on steep wave fronts may be excessive if the rectifier does not act rapidly.
3. Complexity of the modulation circuit. Because of the complexity and low efficiency of the modulation circuit, modulation is normally accomplished at low level and power output is increased by the use of linear amplifiers.

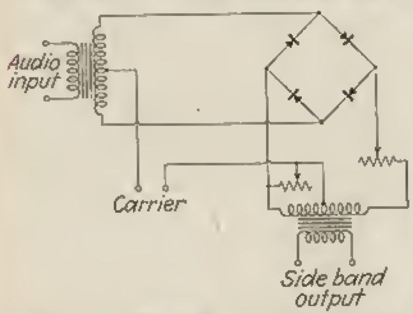


FIG. 38.—Balanced modulator using copper oxide rectifiers. The balance is adjusted with the two resistors. (*Koomans, Proc. I.R.E.*, **26**, 182, *February*, 1938.)

19. Suppressed Carrier Transmitters. In suppressed carrier transmission only the side bands are radiated. In the absence of modulation no voltage appears across the antenna. At the receiver the carrier is reintroduced in order to facilitate demodulation.

The carrier may be suppressed in any of a number of balanced modulator circuits. The tubes are operated on nonlinear portions of their characteristics. By careful adjustment of such balanced modulator circuits, the carrier can be suppressed some 50 to 60 db below its normal value in an a-m wave, but suppression beyond about 50 db is difficult to maintain over long periods of time without readjustment of the modulator tubes. More recently, copper oxide rectifiers have found

application in carrier suppression modulators.[1] By adjustment of the resistors a carrier suppression of 90 to 100 db is possible.

While it would seem possible to effect considerable power savings and noise reduction through the use of carrier suppression, in practice this type of modulation is rarely used because of the difficulties involved in replacing the carrier at the receiver. If the signal is to be demodulated without distortion, the replaced carrier must be not only of the correct frequency, but also of correct phase. This can be accomplished by transmitting a pilot frequency along with the side bands, from which the carrier is derived, but the difficulties are relatively great as compared with single-side-band transmission.

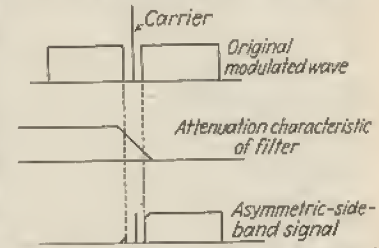


FIG. 39.—Asymmetric-side-band system, using filter to suppress undesired portion of spectrum.

20. Asymmetric (or "Vestigial") Side-band Transmitters. In asymmetric-side-band transmission all of one side band, except for low frequencies, is removed, the carrier is partially attenuated, and the other side band is completely transmitted, except for low modulation frequencies which are partially transmitted. Two systems are in use.

In the first system of asymmetric transmission[2] an ordinary double-side-band signal is passed through a filter whose characteristics are idealized in Fig. 39.

In the Koomans system[3] conventional double-side-band a.m. is employed from the lowest modulation frequency up to some intermediate modulation frequency (about 2,000 cycles for sound transmission), and single side band of double amplitude for higher frequencies. The

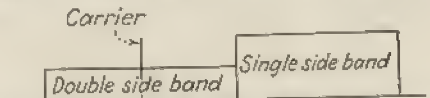


FIG. 40.—Energy distribution in Koomans asymmetric-side-band system. (*Koomans, Proc. I.R.E.*, **27**, 687, *November*, 1939.)

spectrum resulting from the application of a constant-amplitude, varying-frequency modulation voltage is shown in Fig. 40.

The advantages of asymmetric-side-band transmission are as follows:

1. Reduction in band width.
2. Reduction in deleterious effects of selective fading.

Furthermore, asymmetric side-band signals may be generated without complex modulation equipment and may be demodulated by conventional receiving circuits with little distortion.

[1] KOOMANS, N., Single-side-band Telephony Applied to the Radio Link between the Netherlands and the Netherlands East Indies, *Proc. I.R.E.*, **26**, 182–206, February, 1938.

[2] ECKERSLEY, P. P., Asymmetric-side-band Broadcasting, *Proc. I.R.E.*, **26**, 1041–1093, September, 1938.

[3] KOOMANS, N., Asymmetric-side-band Broadcasting, *Proc. I.R.E.*, **27**, 687–690, November, 1939.

21. Single-side-band Transmitters. Two methods are available for the generation of a single-side-band signal, one using filters, the other phase rotation of the modulation and r-f voltages.

In the filter system, the commoner of the two, the carrier is first suppressed by a balanced modulator. This is followed by a sharp cutoff filter which removes the undesired side band. These operations are carried out at a low carrier frequency. The resulting single-side-band signal is then converted to a higher r-f frequency by beating with a

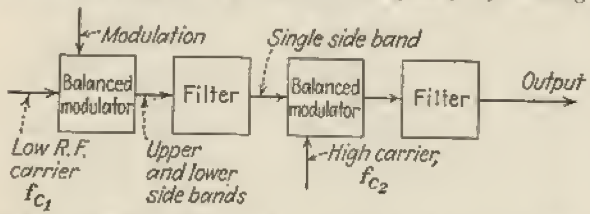


Fig. 41.—Filter system for generating single-side-band signal.

h-f carrier in a second balanced modulator and refiltering to remove undesired modulation products.[1]

The difficulty with such a system is that, if low modulation frequencies are to be transmitted, very sharp filters are necessary. If crystal filter technique is used, the lowest frequency which can be transmitted (in audio transmission) is limited to about 100 cycles. The initial carrier frequency should be chosen as low as possible, and in addition it may be necessary to reach the desired output frequency through several inter-

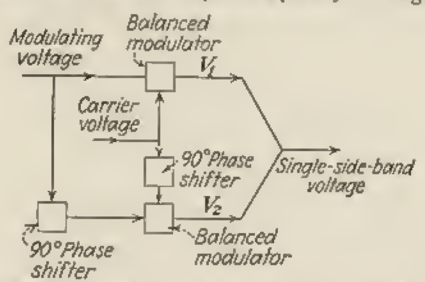


Fig. 42.—Phase-shift system for generation of single-side-band signal.

mediate modulations to simplify filtering the undesired modulation products in each of the subsequent balanced modulator stages.[2]

In the *phase-rotation system* for the production of single-side-band signals no sharp filters are needed. One embodiment is shown in Fig. 42. Two balanced modulators are employed, one of which is fed by modulat-

[1] Polkinghorn, F. A., and N. F. Schaack, A Single Side-band Short-wave System for Trans-Atlantic Telephony, *Proc. I.R.E.*, **23**, 701–718, July, 1935; Oswald, A. A., A Short-wave Single-side-band Radio Telephone System, *Proc. I.R.E.*, **26**, 1431–1454, December, 1938.

[2] Koomans, N., Single-side-band Telephony Applied to the Radio Link between the Netherlands and the Netherlands East Indies, *Proc. I.R.E.*, **26**, 182–206, February, 1938.

ing voltage, and also by carrier voltage, 90 deg. out of phase with the voltages fed to the other.

The outputs of the two modulators are, accordingly

$$V_1 = Vm \sin \mu t \sin \omega t$$

and

$$V_2 = Vm \cos \mu t \cos \omega t$$

where V and m represent the amplitudes of the carrier and modulation voltages, respectively, and ω and μ are the angular velocities of the same voltages. Adding gives

$$V_1 + V_2 = Vm (\sin \mu t \sin \omega t + \cos \mu t \cos \omega t) = Vm \cos (\omega - \mu)t$$

which shows that the output voltage of the system contains only the lower side band. By shifting the voltage to the second balanced modulator in the opposite direction the upper side band can be derived.

The modulation voltage phase shifter must be designed to give a constant 90-deg. phase shift, without change in amplitude, over the entire band of modulation frequencies. While this is difficult, it is not impossible, and various methods have been suggested for its accomplishment.[1]

A single-side-band signal has the following advantages, as compared with amplitude modulation:

1. Reduced channel width.
2. Secrecy; cannot be demodulated with conventional receiver.
3. Improved signal/disturbance ratio.
4. Reduced power consumption at transmitter.

To conveniently demodulate a single-side-band signal, it is usual to reinsert the carrier at the receiver. In contrast to the suppressed carrier system, the distortion is not excessive if the reinserted carrier deviates slightly—to 1 to 5 cycles—from the correct value. A "pilot" frequency is often transmitted along with the side band. The pilot frequency need not be the correct carrier frequency, but only a tone related to the carrier frequency, from which the carrier may be easily derived at the receiver.[2]

22. Single-side-band-plus-carrier Transmitters. In this case the side band may be generated independently, as in the preceding section, and then added to the carrier. Alternatively, if the phase-shift system of Fig. 42 is employed, conventional modulators may be substituted for the balanced modulators, in which case a single-side-band-plus-carrier signal is generated directly.

Single-side-band-plus-carrier has the following advantages:

1. May be demodulated by conventional detectors, without modification.
2. A gain in signal/disturbance ratio over amplitude modulation is attainable.
3. Selective fading is reduced.

23. Frequency-modulation Transmitters. A great many circuits have been proposed for f-m transmitters. Three of these have found application to the Armstrong wide-band system,[3] and will be described here.

[1] For example, see Byrne, J. F., Polyphase Broadcasting, *Trans. A.I.E.E.*, **58**, 347–350, July, 1939.

[2] Koomans, *loc. cit.*

[3] Armstrong, E. H., A Method of Reducing Disturbances in Radio Signalling by a System of Frequency Modulation, *Proc. I.R.E.*, **24**, 689–740, May, 1936.

Armstrong Circuit. A block diagram is shown in Fig. 43*A*. The unmodulated carrier is added to the side bands after the latter have been shifted by 90 deg. The resultant of the carrier and side-band voltages is thereby shifted in phase (Fig. 43*B*), and this change in phase is linearly

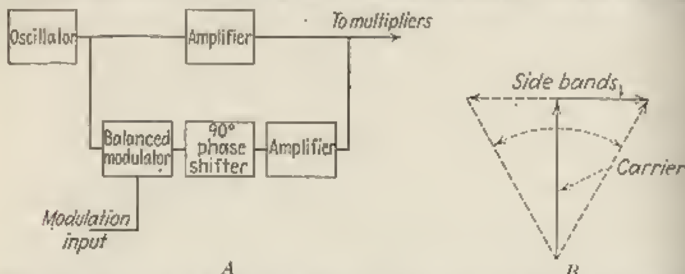


FIG. 43.—Armstrong modulation system. *A*, the circuit arrangement, (*Armstrong, Proc. I.R.E.*, **24**, 689, *May*, 1936); *B*, a vector representation of manner in which side-band voltage is added to carrier to generate p-m wave.

related to the side-band voltage so long as it is restricted to angles less than about 30 deg.[1] In addition to the phase shift the resultant also undergoes a small change in amplitude which is readily removed by the succeeding saturated amplifiers, the "limiter" stages.

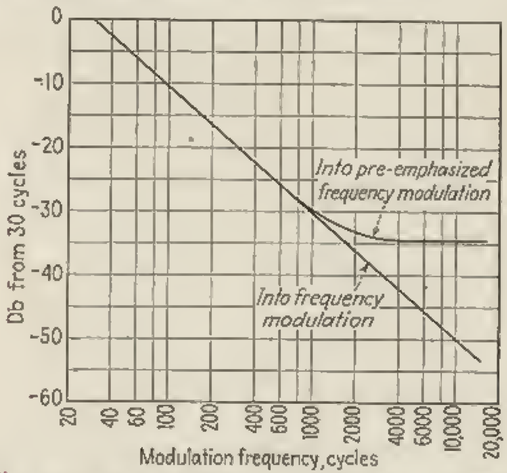


FIG. 44.—Frequency characteristics of audio system to convert phase modulation transmitter into f.m., and into pre-emphasized f.m.

The Armstrong circuit is fundamentally a phase modulator, since the phase deviation is independent of the modulation frequency. If it is

[1] JAFFE, D. L., Armstrong's Frequency Modulator, *Proc. I.R.E.*, **26**, 475–481, April, 1938.

desired that the output wave be frequency modulated, for which the phase deviation is inversely proportional to the modulation frequency, it is only necessary to introduce a correction circuit in the modulation-frequency amplifier whose response is inversely proportional to the modulation frequency. The wide-band system, in its present form (1940), however, makes use of a combination of f.m. and p.m., in which the radiated wave (for audio transmission) is frequency modulated between about 30 and 500 cycles and approximately phase modulated at higher modulation frequencies. This combination is called *pre-emphasized* f.m. The audio response curves employed by Armstrong for converting a p-m transmitter into pre-emphasized f.m. is shown in Fig. 44, along with the curve for converting a p-m transmitter into a f-m transmitter.

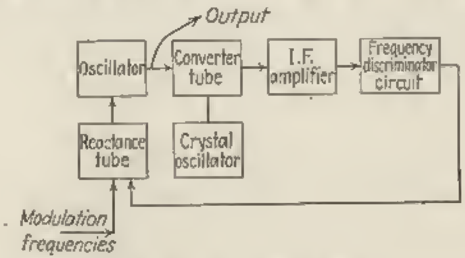


FIG. 45.—A-f-c f-m circuit.

For a sinusoidal modulation voltage the phase deviation, in any system in which the angular velocity (or frequency) is modulated, is related to the frequency deviation by

$$\Delta\phi = \frac{\Delta\omega}{\mu}$$

where $\Delta\phi$ = maximum phase deviation in radians
$\Delta\omega$ = maximum angular velocity deviation in radians per second = $2\pi\ \Delta f$
μ = angular velocity of modulation frequency in radians per second = $2\pi f_\mu$.

In the Armstrong system a frequency deviation of about 75 kc (corresponding to $\Delta\omega = 2\pi \times 75 \times 10^3 = 4.71 \times 10^5$ radians per second) is used. If the lowest a.f. to be transmitted is 30 cycles ($\mu = 2\pi \times 30 = 188$ radians per second), then the phase deviation needed is $\Delta\phi = \frac{\Delta\omega}{\mu} = \frac{4.71 \times 10^5}{188} =$ 2,500 radians. The amount of p.m. which may be produced in a circuit of the type illustrated in Fig. 43 is limited by non-linear distortion to a maximum of about ½ radian. The increase from ½ to 2,500 radians necessitates a frequency multiplication of about 5,000. In the Armstrong transmitters, this is accomplished in a series of low-power multiplier stages. A frequency multiplication of 5,000 may be carried out in 13 doublers, 8 triplers, 6 quadruplers, or combinations of these.

A-f-c Circuit. The second f-m transmitter circuit which has found practical application in the wide-band system is an adaptation[1] of the automatic frequency control system sometimes used in a-m broadcast receivers. If

[1] CROSBY, M. G., British patent 504766; CHIREIX, H., and P. BORIAS, U. S. patent 2076264. Detailed information on the operation of reactance tube and discriminator circuits may be found in the following papers: FOSTER, D. E., and S. W. SEELEY, Auto-

the grid voltage of a high plate resistance tube is fed with r-f voltage from the plate circuit through a 90-deg. phase shifting network, as in Fig. 46, then the impedance seen, looking into the plate circuit, is very nearly a pure reactance whose magnitude (in the absence of degeneration) is

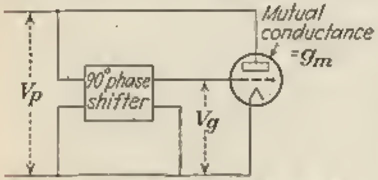


FIG. 46.—Reactance tube in a-f-c circuit.

$$X = \frac{V_p}{V_g g_m}$$

By varying the transconductance of a tube connected in this manner, called a *reactance tube*, the reactance may be varied from infinity to a minimum value indicated by the above equation. The transconductance may be varied by applying the modulation in series with one of the electrode voltages, such as that of the control grid.

In the a-f-c circuit a reactance tube is shunted across the tank circuit of a conventional self-excited oscillator. By varying the bias at an audio rate, the resonant frequency of the tank and the oscillation frequency are varied. To stabilize the mean frequency, a degenerative feedback circuit of the same type as that utilized in automatic frequency control in receivers is employed. The oscillator frequency is heterodyned to an i.f. by means of a converter and crystal oscillator and then passed to a frequency discriminator circuit, whose output voltage is proportional to frequency. This voltage is returned degeneratively to the reactance tube and serves to minimize frequency variations of the oscillator. If the circuit constants of the feedback circuit

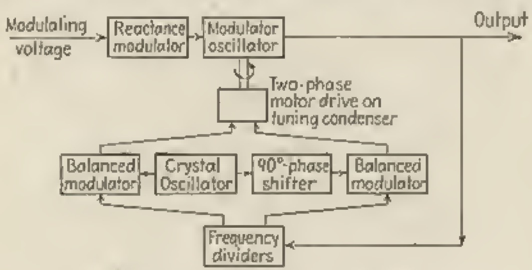


FIG. 47.—Morrison f-m circuit.

are suitably adjusted, the transmitter may be made degenerative for audio frequencies as well as for slower variations and the usual advantages taken of degenerative feedback. At 20 Mc linear frequency deviations of 100 kc or more are easily obtained with the a-f-c circuit, so that little multiplication is needed to adopt the circuit to the wide-band system. The frequency stability, however, depends upon the stability of the discriminator circuit in addition to that of the crystal oscillator, so that temperature or other control of the discriminator may be necessary. Hum problems may be minimized by the use of push-pull reactance tube arrangements.

matic Tuning, Simplified Circuits and Design Practice, *Proc. I.R.E.*, **25**, 289–313, March 1937; RODER, H., Theory of the Discriminator Circuit for Automatic Frequency Control, *Proc. I.R.E.*, **26**, 590–611, May, 1938.

Morrison Circuit. The third circuit to be applied to the Armstrong system was developed by Morrison.[1] The reactance modulator and oscillator are similar in principle to those used in the a-f-c circuit. The frequency correction, instead of being applied through the reactance tube, is furnished by a motor-driven tuning condenser. The motor is operated from vacuum-tube modulators, supplied by two voltages, one of fixed frequency from a crystal oscillator, the second proportional to the mean frequency of the output signal. The difference between the frequencies of the two voltages actuates the motor and corrects the output frequency accordingly.

The voltage whose frequency is proportional to the mean frequency of the output signal is derived without the use of frequency selective circuits. This is accomplished by passing the modulated wave through frequency dividers which reduce the depth of modulation to a small value. The advantage of the Morrison system is that the modulation circuits and the frequency control circuits are independent of each other. One function is not limited by the other, therefore.

TECHNICAL FEATURES OF H-F RECEIVERS

24. High-frequency Receiver Requirements. As in the case of h-f transmitters the details of the specifications for h-f receivers are fixed by the use to which the receiver is to be put. The items which are usually considered in the design of a receiver are the same as those listed for h-f transmitters and, in addition, the available signal strength, selectivity requirements, and image performance must be considered.

25. General Receiver-design Considerations. Most h-f receivers are of the superheterodyne variety, although at ultra-high frequencies other circuits, regenerative detectors, the superregenerative receiver, and even diodes or crystals followed by audio amplification are sometimes employed. To attain the best signal/noise ratios, it is necessary to amplify the signal before conversion to the i.f., or to employ a high-gain converter tube; improved tube designs are constantly extending the h-f limit at which such amplification or conversion is possible.

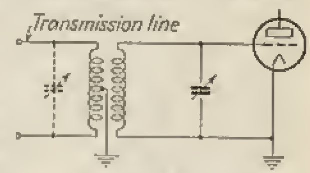


FIG. 48.—Typical input circuit of receiver.

The i.f. of the h-f bands of a home receiver is usually the same as that employed for the broadcast band, which is fixed by other considerations at about 455 kc. This is too low to give satisfactory image response and other characteristics for many specialized types of h-f receivers. Accordingly, in receivers primarily intended for h-f use, higher intermediate frequencies are found, values near 1.5, 3, 5, and 10 Mc being common.

From the point of view of signal/noise ratios the input circuits of the receiver are of primary importance. Figure 48 is a typical antenna coupling circuit, which may be idealized in the manner shown in Fig. 49, where the voltage e and the resistances have been reduced to terms of either primary or secondary quantities. The thermal agitation noise

[1] MORRISON, J. F., A New Broadcast Transmitter Circuit Design for Frequency Modulation, *Proc. I.R.E.*, **28**, 444–449, October, 1940.

power generated in the input circuits at room temperature is

$$W = 1.64 \times 10^{-20} F \text{ watts}$$

where F is the band width of the receiver in cycles. In a receiver of optimum design the only receiver noise affecting the signal/noise ratio is that resulting from thermal agitation in the input circuit. To accomplish this, the gain of the input circuit and of the first amplifier tube is made as large as possible, so that the thermal noise of the input circuit predominates over the tube noise in the first tube and in the superheterodyne converter and over the thermal noise in later circuits.[1] At low frequencies, where the tube input resistance is very high, this is accomplished by making the ratio R_2/R_1, Fig. 49, and the gain of the first tube large, or, in terms of Fig. 48, the transformer ratio and its Q are made high. At higher frequencies the input resistance of the tube fixes the maximum impedance of the first circuit. In such cases best operation is obtained if R_1 is made equal to R_2. The resulting signal/noise ratio is 3 db lower than the theoretical maximum.

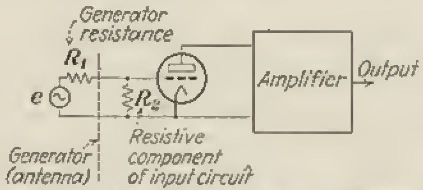


FIG. 49.—Equivalent input circuit of receiver. (*Johnson, Llewellyn, Elec. Eng.*, **53**, 1449, *November*, 1934.)

If the receiver is connected to the antenna through a transmission line, it is desirable (and in television, necessary) that the transmission line be terminated in its characteristic impedance to minimize reflections and the attendant distortion in the frequency characteristic. This also corresponds to making $R_2 = R_1$ in Fig. 49.[1] In the design of wide-band receivers it is necessary also that the frequency response of the input circuit be considered.

Converter circuits are similar to those employed at low frequencies, except that more attention must be paid to interlocking ("pulling") between oscillator and converter circuits since the ratio of signal to intermediate frequencies is usually high. If separate oscillator and mixer tubes are employed, the 1851 and 1852 tubes will be found to have high conversion transconductances, of the order of 3,000 micromhos, and low noise.[2] The 6K8 tube is the best combination mixer-oscillator tube available at present (1940) for use in h-f superheterodynes.

26. Ultra-high-frequency Receivers. Up to about 100 Mc u-h-f receivers follow the same patterns as other h-f superheterodyne receivers. Differences are as follows: r-f amplifier and conversion gains are lower, loading of tank circuits by tubes is more troublesome, higher intermediate frequencies are employed, and interlocking of oscillator and amplifier tuning is more difficult to avoid.

[1] JOHNSON, J. B., and F. B. LLEWELLYN, Limits to Amplification, *Elec. Eng.*, **53**, 1449–1454, November, 1934.
[2] KAUZMANN, A. P., New Television Amplifier Receiving Tubes, *RCA Rev.*, **3**, 271–289, January, 1939.

To avoid excessive loading of the tank circuits (Fig. 50) by the tubes, the plate and grid connections are tapped down on the tank circuits. Lead lengths should, of course, be short, and the use of by-pass capacitors with low-impedance leads is essential.

The 6K8 tube as an oscillator-converter or the 1852 as a converter with a separate oscillator were most widely used in 1940.

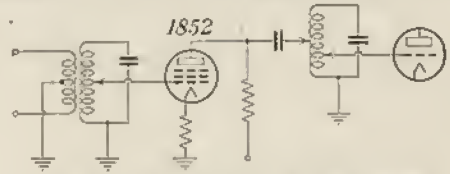


FIG. 50.—Typical u-h-f amplifier circuit.

Above about 500 Mc it is difficult to amplify the received signal at the carrier frequency using conventional u-h-f tubes. A diode converter circuit successfully used in a 700-Mc receiver[1] is shown in Fig. 51. The third harmonic of the oscillator heterodynes with the incoming signal in the special diode to produce a 10-Mc i-f beat.

Because of the difficulty of amplification and frequency conversion of centimeter waves simpler receiver types than superheterodynes are often employed. A crystal detector followed by an audio amplifier, Fig. 52, is the simplest of these and, while insensitive, is frequently used in laboratory receivers. Regenerative detectors, as in Fig. 53, have also been employed. To increase the sensitivity, the supergenerative principle is often used.[2]

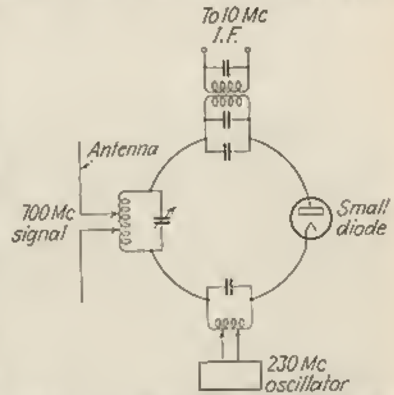


FIG. 51.—Centimeter-wave mixer circuit. (*Bowles, Barrow, Hall, Lewis, Kerr, The CAA-MIT Instrument Landing System, presented at A.I.E.E. Convention, Jan.* 22, 1940.)

Perhaps the most promising development in centimeter-wave receiver technique is the application of the electron-beam principle to converter and amplifier tubes.[3] It is probable that this principle will be widely used in the near future and that appreciable amplification at frequencies above 500 Mc will be attained thereby.

[1] BOWLES, E. L., W. L. BARROW, W. M. HALL, F. D. LEWIS, and D. E. KERR, The CAA-MIT Instrument Landing System, presented at A.I.E.E. Convention Jan. 22, 1940.
[2] Many such receivers are described in the following: GROOS, O., "Einfuhrung in Theorie und Technik der Decimeterwellen," S. Hirzel, Leipzig, 1937; "Radio Amateur's Handbook," 17th ed., American Radio Relay League, Hartford, Conn., 1939; "Radio Handbook," ed. by W. W. Smith, 6th ed., Radio, Ltd., Santa Barbara, Calif., 1939.
[3] HAHN, W. C., and G. F. METCALF, Velocity-modulated Tubes, *Proc. I.R.E.*, **27**, 106–116, February, 1939.

27. Reception of Single-side-band-plus-carrier and Asymmetric-side-band Signals. Signals of these types may be amplified and demodulated with conventional receivers. In the asymmetric-side-band case the carrier in the side-band filter should normally be located at the midpoint of the filter attenuation curve, *i.e.*, at the point where the filter is 6 db down. Such operation normally gives minimum distortion.

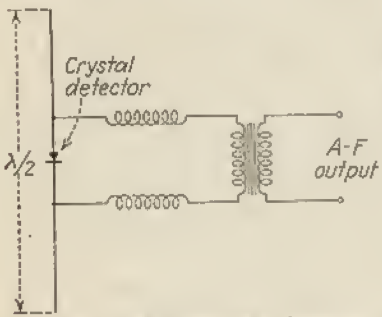


FIG. 52.—Centimeter-wave detector.

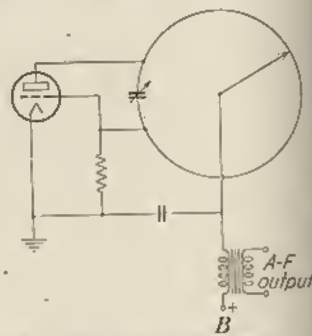


FIG. 53.—Centimeter-wave regenerative detector.

A single-side-band-plus-carrier signal suffers a certain amount of nonlinear distortion when rectified by a linear rectifier. Figure 54 shows the r-m-s total of the harmonics produced by the demodulation of a sine-modulated single-side-band signal by means of a linear rectifier.[1] This distortion is largely second harmonic, and may be partially avoided by the use of a full-wave or of a square-law demodulator. Some evidence indicates also that the distortion produced by the demodulation of a single-side-band signal by a linear rectifier is not so objectionable to the ear as the values indicated by the curve would indicate. For sinusoidal modulation no distortion is produced if the single-side-band signal is demodulated by a square-low rectifier. The modulation depth *m* of Fig. 54 equals 2 when the side band and carrier are of the same amplitude.

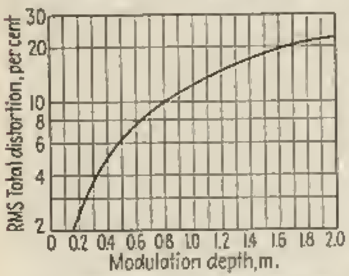


FIG. 54.—R-m-s total harmonic distortion introduced in demodulation (by linear detector) of sinusoidally modulated single-side-band-plus-carrier signal.

In asymmetric-side-band transmission the demodulation distortion is avoided by the use of both side bands at modulation frequencies for which the percentage modulation is likely to be high. At high modulation frequencies the energy content in typical program material is low, so

[1] WHITE, P. J., Modulation Distortion (in Dutch), *Tijdschr. Nederland. Radiogenoot.* **7**, 99–114, April, 1936.

that the distortion resulting from rectification is correspondingly low.[1] An alternative expedient that may be employed is the accentuation of the carrier in the receiver. No simple means of accomplishing this is available, however.

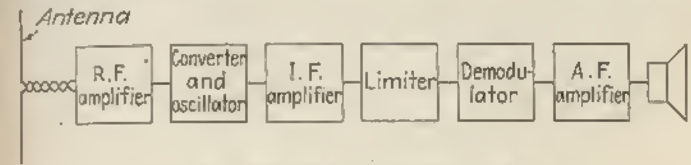


FIG. 55.—Frequency-modulation receiver.

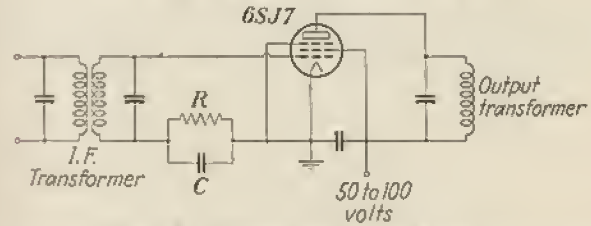


FIG. 56.—Conventional limiter circuit.

While the demodulation distortion is higher in single-side-band-plus-carrier and in asymmetric-side-band systems than in a conventional a-m transmission, it should be noted that distortion resulting from selective fading, which may be very objectionable, is reduced.

28. Single-side-band Receivers. In order to demodulate a single-side-band signal with a conventional rectifier, a carrier of approximately the correct frequency must be added to the received signal. For highest quality reception the replaced carrier must be within 1 to 3 cycles of the correct position, since all the frequencies in the received signal will be shifted by the amount by which the replaced carrier deviates.

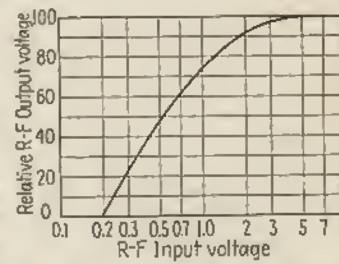


FIG. 57.—Typical limiter characteristic.

The simplest means for replacing the carrier is to add the output of a stable oscillator to the signal in the i-f or r-f channel of the receiver. A crystal oscillator is convenient for this purpose. If the carrier is replaced in the i-f circuit, the heterodyne oscillator of the superheterodyne must also have good stability.

In many single-side-band signals a pilot frequency is transmitted along with the signal, from which the carrier is derived at both the receiver

[1] ECKERSLEY, P. P., Asymmetric-side-band Broadcasting, *Proc. I.R.E.*, **26**, 1041–1093, September, 1938.

and the transmitter. If this arrangement is employed, the receiver carrier frequency cannot depart from the correct value. In some arrangements the pilot frequency is filtered from the signal and used to operate an automatic frequency-control circuit connected to the beating oscillators. A number of schemes have been devised for this purpose.[1]

29. Frequency-modulation Receivers. In Fig. 55, from the antenna through the i-f amplifier, the receiver is quite conventional and, for the wide-band system in present use (1940), should have a band width of 150 to 200 kc. Following the i-f amplifier is a limiter stage, designed to

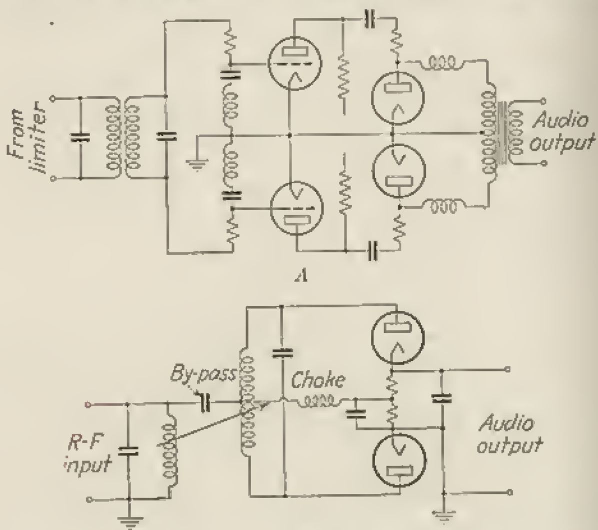


FIG. 58.—Frequency-modulation demodulator circuits. *A* shows Armstrong circuit (*Armstrong, Proc. I.R.E.*, **24**, 689, *May*, 1936); *B* shows frequency discriminator. By-pass condenser is for r.f.

remove amplitude variations from the signal as completely as possible. The time constant of the *RC* combination is preferably less than 10 microsec. in an audio receiver. It is essential that the horizontal portion of the input-out characteristic of a limiter stage be flat, and it is desirable that it extend to low values of input voltage. New developments in limiter circuits will undoubtedly improve these two limiter properties. Full advantage cannot be taken of the benefits possible with the wide-band system unless sufficient gain is provided preceding the limiter so that the input signal is always beyond the flat part of the curve. If this

[1] KOOMANS, N., Single-side-band Telephony Applied to the Radio Link between the Netherlands and the Netherlands East Indies, *Proc. I.R.E.*, **26**, 182–206, February, 1938; POLKINGHORN, F. A., and N. F. SCHAACK, A Single Side-band Short-wave System for Trans-Atlantic Telephony, *Proc. I.R.E.*, **23**, 701–718, July, 1935; OSWALD, A. A., a Short-wave Single-side-band Radio Telephone System, *Proc. I.R.E.*, **26**, 1431–1451, December, 1938.

can be accomplished, an input signal only a few decibels above the noise level can be made to produce noise-free output voltage.

The demodulator circuit, which follows the limiter, is designed to convert frequency variations into I-f output and also to assist the limiter in suppressing amplitude variations. The two circuits of Fig. 58 are in current use. Figure 58*A*, Armstrong's circuit,[1] employs two series *LC* circuits resonant above and below the mid-band frequency. The voltages of the two series circuits are rectified and combined in the polarity which partially cancels amplitude changes. Figure 58*B* is an adaptation of the frequency discriminator, employed in automatic frequency-control circuits.[2] It affords the same advantage of partial cancellation of amplitude variations as the circuit above.

[1] ARMSTRONG, E. H., A Method of Reducing Disturbances in Radio Signalling by a System of Frequency Modulation, *Proc. I.R.E.*, **24**, 689–740, May, 1936.

[2] FOSTER, D. E., and S. W. SEELEY, Automatic Tuning, Simplified Circuits and Design Practice, *Proc. I.R.E.*, **25**, 289–313, March, 1937; RODER, H., Theory of the Discriminator Circuit for Automatic Frequency Control, *Proc. I.R.E.*, **26**, 590–611, May, 1938.

SECTION 16

CODE TRANSMISSION AND RECEPTION

By John B. Moore, B.S.[1]

1. Radio communication, as distinguished from radio broadcasting of educational and entertainment programs, is carried on chiefly by means of some one of the recognized telegraph codes. Radiotelegraph signals are, therefore, made up of short and long periods of constant signal strength separated by idle periods of proper duration to correspond to the combinations of dots, dashes, and spaces comprising the characters of the code being used. The design of the entire system must be such that the lengths of the dots, dashes, and spaces in the signal supplied to the receiving operator are substantially the same as they were made by the transmitting operator. In a simple system operated at slow speeds no special difficulties are encountered in meeting this requirement. Present-day commercial systems, however, which utilize remote control from a central traffic office and which are operated at high keying speeds, impose severe requirements on all the equipment used.

2. Standard Codes. In international communication the International Morse Code is used. Specially marked and accented letters such as are used in German, French, and the Scandinavian languages have special characters which are used when working a station in the same country or its possessions. When communicating with a foreign station, these letters are either replaced by a combination of unaccented letters or in some cases the unaccented letter is transmitted alone. Some countries such as Japan and Egypt having alphabets differing radically from the Latin alphabet use special codes for working within the country or to ships. Nationals of such countries desiring to transmit a message in their own language to a foreign country must spell out the sounds of their words in one of the languages using the Latin alphabet.

3. Business Codes. Business concerns that have a large volume of telegraph communication use so-called five-letter or ten-letter codes. Standard codes for such use are available and consist of groups of letters arranged alphabetically; each group standing for a complete sentence or part of a sentence. Special and private codes are also used, and large concerns often have a department for the coding and decoding of coded telegraphic messages.

4. Printing Telegraph Equipment. Various types of printing systems, in which the received signal is automatically printed in standard letters on a paper tape, are being employed on the higher grade radio circuits of the world. The trend is toward such automatic reception, as a substitute for manual transcription.

Codes employed are the Standard International Morse Code, the "five unit" Baudot Code, and the recently developed[2] "seven unit" code.

[1] Research receiving engineer, R.C.A. Communications, Inc.
[2] U.S. Patent 2,183,147

A	·—
B	—···
C	—·—·
D	—··
E	·
F	··—·
G	——·
H	····
I	··
J	·———
K	—·—
L	·—··
M	——
N	—·
O	———
P	·——·
Q	——·—
R	·—·
S	···
T	—
U	··—
V	···—
W	·——
X	—··—
Y	—·——
Z	——··
Period .	·—·—·—
Comma ,	——··——
Colon :	———···
Question mark, or request for repetiton of a transmission not understood ?	··——··
Apostrophe '	·————·
Dash or hyphen –	—····—
Fraction bar /	—··—·
Parenthesis (before and after words) ( )	—·——·—
Underscore (before and after words or part of sentence)	··——·—
Equal sign =	—···—
Understood	···—·
Error	········
Cross or end-of-telgram or end-of-transmission signal	·—·—·
Invitation to transmit	—·—
Wait	·—···
End of work	···—·—
Starting signal (beginning every transmission)	—·—·—
Separation signal for transmission of fractional numbers (between the ordinary fraction and the whole number to be transmitted) and for groups consisting of figures and letters (between the figure-groups and the letter-groups)	·—··—

Ä (German)	·—·—
Á or Å (Spanish-Scandinavian)	·——·—
CH German-Spanish	————
É (French)	··—··
Ñ (Spanish)	——·——
Ö (German)	———·
Ü (German)	··——

1	·————
2	··———
3	···——
4	····—
5	·····
6	—····
7	——···
8	———··
9	————·
0	—————

Fig. 1.—The Continental code.

The International Morse Code consists of dots and dashes, as depicted on page 576. The Baudot Code is built up of all possible combinations of five consecutive and equal time intervals, numbered 1 to 5, into which the length of time allotted to the transmission of any code group is divided. The "seven unit" printer code divides the time allotted for transmission of any one code group into seven consecutive and equal time intervals. Only three of these possible seven pulses are used for any code group or character. The receiving equipment is so designed that a received group containing fewer or more than three pulses or marking intervals will cause a special "error sign" to be printed instead of an incorrect letter or figure.

5. Multiplex operation over a radio circuit has certain very definite advantages from the viewpoint of the traffic man. No single operator can keep traffic moving at 100 words per minute. The radio circuit, however, is often capable of handling twice this speed or better. Economical operation then requires that two or more operators be assigned to the circuit. Multiplex equipment permits doing this in the most expeditious and straightforward manner since each operator then has a channel under his complete control. This makes it possible to efficiently use suitable printing telegraph equipment—each such channel being handled by a single operator at a speed of approximately 50 words per minute. Three such channels give a circuit capacity of 150 words per minute, with no complications such as are experienced when such high-speed operation is attempted over a single-channel circuit employing tape transmission and reception.

Two basic types of multiplex system have been employed. One utilizes two or more modulating frequencies, which are applied to the radio transmitter. The other employs the time-division principle. This latter is a more recent development, as applied to radio communication systems. Its chief advantage is that it can be applied to any radio-telegraph circuit which will properly handle the keying speeds required by the particular system and equipment. Time-division multiplex systems now in use on long-distance radio-telegraph circuits provide a total of two, three or four separate channels over the one radio circuit.

6. Character Formation. The unit used in code characters, and in figuring speeds of transmission, is the dot. Present practice, based on automatic transmitting equipment, is to speak of dots per second. On this basis the time required to transmit one dot includes the duration of the space separating the dot from the next element of the character. As the duration of the dot itself and of the following space are equal, they constitute a cycle. Keying speeds are, therefore, commonly stated in dots, or (square) cycles, per second. The equivalent time required for the transmission of the other elements of the code are as follows: a dash, two dots; space between letters, one dot; space between words, three dots. For traffic purposes speeds are generally stated in words per minute. The ratio of words per minute to dots or cycles per second is generally accepted as being 2.5:1 for usual commercial traffic, 100 words per minute being equivalent to 40 cycles keying frequency.

In the Baudot code used for printing telegraph equipment, the duration of the character is divided into five equal periods. For any one of these periods either a marking, or a spacing (no current or reverse current) impulse may be transmitted. One impulse is required between letters, and in the non-synchronous type of equipment an additional impulse is required at the start of each character to set the receiving mechanism in motion. The total number of elements per character is, then, either six or seven depending on the type of equipment used. The space between words is a full-length character. The code consists of a different combination of marking and spacing impulses for each character, there being a total of 32 possible combinations for the five periods utilized. For calculation of keying frequency the single period or element, which is the shortest impulse required to be transmitted, corresponds to the marking portion of a dot in the Morse Code. This is one half cycle. For the non-synchronous printer equipment each letter requires, for its transmission, seven half cycles or three and one-half full cycles. On the basis of five letters per word and a space between words, the ratio of words per minute to keying cycles per second is 2.86 to 1. This is the figure realizable with automatic tape transmission. Where the impulses go directly from the keyboard-operated machine to the line, the dot speed will remain unchanged, but the number of words per minute that can be transmitted will be reduced on account of the unavoidable irregularities in the speed of the typist.

7. Required Frequency Range. A square-wave shape such as a succession of dots, where the value of the current or voltage rises instantly to a steady value at which it remains for one half cycle and then instantly drops to zero, can be analyzed into the fundamental and all of its odd harmonics. The equation of the voltage wave is

$$e = \frac{4E}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots\right) \tag{1}$$

which holds for values of x between $-\pi$ and $+\pi$. For most practical telegraphic purposes it is only necessary for the system to pass the fundamental, third, and fifth in their proper intensity and phase, as terms of higher order do not add sufficiently to the fidelity to warrant building the equipment to handle them. The frequency range required by a sufficient number of higher order harmonics to give appreciable improvement can often be used to better advantage for additional channels.

For any service where the received signal strength rises to the same maximum value on every dot and dash, it is not necessary to pass even the third harmonic of the keying frequency. A system which will pass the second harmonic of the fundamental keying frequency is satisfactory. The receiving equipment can be adjusted to operate at a fairly definite level on the building up and decaying of the current or voltage wave so as to give characters which are neither too heavy (long) nor too light (short) as compared to the spaces. However, in a system where the received signal may vary by 2:1 or more in intensity at fairly short and frequent intervals, it is necessary to have quite a steep rise and fall of the received signal at make and break in order to obtain a constant "weight" of keying. This applies particularly to automatic reception, where the signal operates a recording device either directly from amplifiers or through a relay of either the mechanical or vacuum-tube types. For aural reception it is desirable to retain the harmonics of the keying frequency, as the signal then sounds cleaner cut and more definite, making it easier to read.

Cases of interference, in both the radio and the land-line portions of a system, are sometimes encountered where it is necessary slightly to round off the sharp, square envelopes of the dots, in order to reduce or eliminate the interference or cross talk caused by the too sudden rise and fall of current.

Where the exact effect of a given circuit on the shape of a square input wave is desired, the range of frequencies passed by the system must be considered as a continuous band rather than dealing with only odd harmonics of the keying frequency.

The usual modulation and side-band theory of radio telephony is applied to code transmission by considering the fundamental keying frequency, and such of its harmonics as are passed, to modulate the carrier 100 per cent. The total band width required to be passed by the entire system is equal to twice the frequency of the highest harmonic of the keying speed that it is desired to retain. (See Arts. 30 to 33 for actual values.)

8. Speeds Attainable. Speeds of transmission range from about 15 up to 300 words per minute; the corresponding keying frequencies being 6 to 120 square cycles per second. Work with ships and with aircraft is carried on mainly at speeds up to about 35 words per minute. Transmission is by means of a manually operated telegraph key. Reception is by ear. In point-to-point service, such as transoceanic, traffic speeds normally range from 30 up to 250 words per minute depending upon the type of equipment used, transmission conditions, and the amount of traffic to be handled. Keying is done by machine almost entirely, hand-operated keys being used only for minor service communications. Reception is generally by means of an ink recorder, the telegraphic characters on the tape being transcribed on a typewriter by the operator. Aural reception is resorted to only under adverse conditions. In radio systems where multiplex equipment is employed on the circuits, each channel of the two or three going over a single circuit will operate at approximately 50 words per minute. This gives the circuit a total capacity of 100 or 150 words per minute.

9. Fidelity of the mark-to-space ratio, while important at all speeds, requires special attention when automatic operation at speeds in excess of 100 words per minute is to be maintained. Where the duration of the mark portion of a dot is only $\frac{1}{60}$ sec. or less, factors that are disregarded at slow speeds become of primary importance. Automatic transmitters, relays, and electrical circuits should be fast enough so that the signal supplied to the recording equipment will not be heavier than 60/40 or lighter than 40/60 in mark-to-space ratio at the highest speed used. At 200 words per minute, which is not exceptional in present-day short-wave work, this means a variation of not more than 1.25 millisec. in the duration of a dot. While it is sometimes possible to compensate for heavy or light keying characteristics by means of relay adjustments in another portion of the system, this should not be depended upon for obtaining the desired over-all fidelity. Each unit of the system should be capable of giving the required fidelity at a speed in excess of the maximum operating speed, the margin required depending on the number of elements in the over-all system and the fidelity of each.

10. Checking the keying characteristics of portions of, and of the entire, system is done by means of keying wheels which send out either a single word over and over, or a succession of dots of 50/50 mark-to-space ratio. For speeds up to about 100 words per minute the usual high-speed ink recorder can be used for checking character formation

quite satisfactorily. For accurate information, especially at higher speeds, some form of oscilloscope or oscillograph must be used. The low-voltage type of cathode-ray oscilloscope is admirably suited to this work where photographic records are often not required. Associated amplifiers must be better than the equipment being tested.

11. Requirements for Facsimile. Facsimile service requires equipment capable of handling keying frequencies up to about 500 square dots per second. This speed is possible only on short-wave equipment and requires a band width of about 5,000 cycles. In the transmission of facsimile half tones higher keying speeds may require a total band width of 10,000 cycles. This system of facsimile is now practically obsolete.

RADIOTELEGRAPHIC SERVICES

Services. Code-communication channels and equipment can be classified, according to the type of service rendered by them, under the general headings of transoceanic, shorter distance point to point, ship to shore, aircraft, special mobile services, and military.

12. Transoceanic (long-wave), long-distance communications were, prior to 1928, handled almost exclusively on frequencies ranging from about 14 to about 30 kc. Great-circle distances covered on such commercial circuits range from 2,000 to 5,000 miles, roughly. To cover distances greater than this with commercial reliability requires so much power to be radiated from the transmitter that it becomes uneconomical.

Approximate values of signal strength to be expected are calculated from the Austin-Cohen transmission formula

$$E = 120\pi\frac{HI}{\lambda D}\sqrt{\frac{\theta}{\sin\theta}} \times \epsilon^{-u} \tag{2}$$

$$u = \frac{0.0014D}{\lambda^{0.6}}$$

where HI = effective height times current for transmitting antenna in meter amperes
λ = wave length in kilometers
D = great-circle distance in kilometers
θ = arc of great circle between transmitter and receiver
E = received field strength in microvolts per meter

or the slightly different expression

$$E \text{ in } \frac{\mu V}{m} = \frac{377HI}{\lambda D}\epsilon^{-u} \tag{3}$$

where

$$u = \frac{0.005D}{\lambda^{1.25}}$$

which is derived from data taken on the New York to London circuits at frequencies ranging from 17 to 60 kc.[1]

13. Field Strength Required (Long Wave). For successful operation the received field strength must be sufficiently above the level of atmospheric disturbances and other local sources of noise to give fully readable signals. Automatic recording requires a signal-noise ratio of at least 2:1. This is based on the general, or average, noise level. Moderately

[1] Espenschied, Anderson, and Bailey, *Proc. I.R.E.*, February, 1926.

severe atmospheric disturbances such as "crashes" and "clicks" will be from several to perhaps ten times as strong as a normally satisfactory signal. Field strengths obtained on transoceanic circuits range from 10 or less up to 250 μv per meter. A value of 20 is about the minimum for satisfactory communication under average conditions. Modern high-powered transmitting stations have an antenna input power of from 40 to 500 kw with output ratings up to some 130,000 meter-amp.

14. Short Wave. During the last few years "short waves" have assumed increasing importance in long-distance radio communication of all types. Frequencies used range from about 4,000 to 23,000 kc, depending upon distance, season of year, time of day, and path traversed. Proper choice of frequency allows of reliable communication between any two points on the earth with transmitters of modern design. Power output of the equipment ranges from 1 to 40 kw. Owing to the extreme variations in transmission conditions encountered at these frequencies, it is necessary to have available at least 10 kw output from the transmitters for high-speed automatic operation over the longer distances. Even with the maximum output of present transmitters and with directive antennas for both transmission and reception, communication is slowed down or even stopped, at times, by severe disturbances in transmission conditions. Normal field strengths obtained at the receiving antennas range from 0.1 up to 100 μv per meter or more, depending on transmitter radiation, path, and transmission conditions. The minimum signal required for reliable commercial operation depends partly on the noise level at the receiving point. Atmospheric disturbances (static), while troublesome at times, are not so serious as in the case of long waves. Fading requires the use of a greater signal-noise ratio on short waves. Utilization of space, frequency, polarization, or time diversity of fading will overcome, to a great extent, the bad effects of static and permit successful operation on much weaker signals. A very rough estimate of the minimum field strength ordinarily required for code communication, with automatic recording, is 5 μv per meter. Slow-speed aural reception can be carried on with field strengths of as low as 0.1 μv per meter.

Minimum field strength required is determined by (1) directional distribution of *noise* at the receiving point; (2) directivity and pickup of the antenna system, which are both effective in determining the *gain* of the antenna in signal-noise ratio as compared with a standard vertical doublet; (3) the *noise equivalent* of the receiver itself.

15. Short Waves *versus* Long Waves. Advantages of short waves for transoceanic code communication are (1) lower first cost of equipment and antennas, (2) smaller power consumption, (3) higher keying speeds of which the equipment is capable, (4) less trouble from static, (5) directive transmission, (6) greater distances covered with a reasonable and practicable transmitter power. Disadvantages are (1) interruption of service due to severe magnetic disturbances, (2) effects of fading, (3) necessity of having several frequencies, a separate antenna being required for each, for 24-hr. service the year round.

Advantages of long-wave operation are (1) freedom from interruption of service by magnetic disturbances, (2) comparative reliability and steadiness of signal strengths. Long-wave arcs, alternators, and tube sets are used. Tube transmitters, only, are used for short-wave operation.

16. Point-to-point communication for distances up to some 2,000 miles is carried on at frequencies ranging from approximately 30 kc

up to 100 kc. These stations are used for domestic service and also for the shorter international circuits. Certain bands in the 6,000- to 23,000-kc portion of the spectrum are also used for these shorter circuits.

Types of equipment used for 30- to 100-kc work include spark (obsolete), arc, frequency multipliers, and tube transmitters. For short-wave operation, tube transmitters are used exclusively.

17. Ship-to-shore and ship-to-ship communication is an entirely different class of service, in all respects, from point to point. Except at the larger coastal stations and on a very few ships, transmission is entirely by hand and copying is by ear. This is because of the nature of the service; a coast station usually has not more than 10 to 20 messages for one ship at a time, and vice versa. Automatic transmission and reception are used only when traffic on hand amounts to some 40 messages or more. The same operator generally handles both transmission and reception, which is not the case in point-to-point work. Owing to the great number of ships, and to the intermittent nature of their traffic, the marine frequency bands must be shared by all ships. This creates interference and traffic-handling problems that are not encountered in point-to-point work. A marine operator must be located at the receiving equipment. Remote control is used only on the transmitters of coastal stations, the transmitting and receiving stations being separated by distances of up to 50 miles to permit of simultaneous transmission and reception.

Frequencies utilized lie within the 100- to 550-kc band; those around 150 kc being used for long-distance work to the larger ships, while those from 400 to 550 kc are for shorter distance work, mainly to the smaller ships, and for distress calls (500 kc). Coastal stations using efficient 5- to 10-kw transmitters and directive reception can normally work ships about 1,500 miles and up to 3,000 miles under favorable conditions, at the lower frequencies. Operation in the 400- to 550-kc band is more variable, a 5-kw transmitter having a normal daytime range of around 500 miles and a night range of several thousand under favorable conditions.

Spark (obsolete), arc, and tube transmitters are used at the lower frequencies. On the higher frequencies tube sets are replacing the old spark equipment. These operate either cw or icw as desired.

Short waves have been coming into more and more use for the handling of ship-to-shore telegraph traffic and special services. The chief advantage is the great distances that can be covered with a low-powered transmitter, as compared with conditions existing on the 500-kc and lower frequency marine bands.

TRANSMITTING SYSTEMS AND EQUIPMENT

18. The high-frequency alternator is one of the most used types of transmitter for long-wave transoceanic code communication. The *Alexanderson alternator* used in this country is a high-speed inductor-type machine having a large number of poles so that frequencies up to 30 kc and higher may be obtained directly. These machines have an output of 200 kw and are driven by a 600-hp. two-phase induction motor through a set of gears to give the desired alternator speed. The stator is built in sections to facilitate dismantling for repairs and maintenance and has 64 separate windings which are connected to separate windings on the antenna-input transformer. One winding is used to supply a tuned circuit, the output of which is rectified and used for automatic speed control. Forced lubrication and water cooling are used on account of the high speed and relatively high losses as compared with commercial power-frequency machinery. Such an alternator intended for operation at 27,200 cycles is driven at a speed of 2,675 r.p.m., has 1,220 poles, and requires a field current of 2 amp. at about 120 volts.

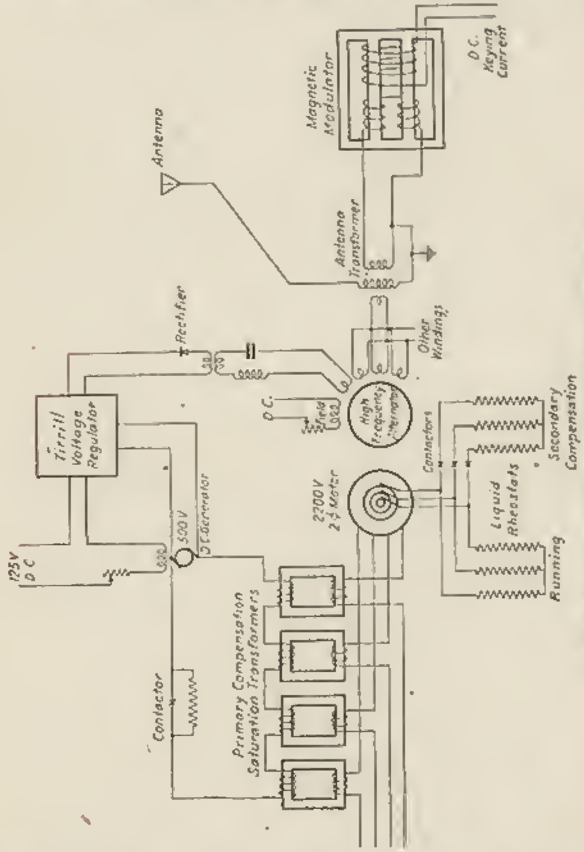


Fig. 2.—Alexanderson alternator equipment.

To maintain the frequency constant to approximately 0.1 per cent and to have it the same under conditions of full load and practically no load, elaborate compensating means are provided as shown on the schematic diagram. Primary compensation saturation transformers each have an a-c and a d-c winding so connected that the voltage at the motor depends upon the impedance of these transformers which, in turn, depends upon the value of current in the d-c winding. Connected to the

slip rings of the wound rotor are two banks of liquid rheostats, the "running" bank being connected at all times and the compensation bank being thrown on or off by the contactors. These contactors, and the contactor in the primary compensation d-c control circuit, are operated from a master relay which is controlled from the central traffic office. Compensation adjustments are made to maintain the machine at the same speed with the control key open or closed.

19. Method of Keying. Keying the output is accomplished by means of a magnetic modulator which is a special transformer having an a-c winding and a differentially connected d-c saturation winding. When the control key is open, a relay closes this d-c circuit, and the resulting drop in impedance of the a-c winding detunes the antenna and reduces the alternator output voltage so that practically no current circulates in the antenna circuit. For key closed, the d-c winding is deenergized and the antenna circuit now becomes resonant to the alternator frequency, so that normal antenna current is obtained. Owing to the low frequency of the system and the low resistance of the antenna circuit, also on account of the large contactors required in the compensation circuits, keying speeds are limited to about 120 words per minute on long-wave transmitters.

20. Goldschmidt Alternator. Another type of h-f machine that has been used to some extent is the Goldschmidt alternator. The fundamental frequency generated is usually one-fourth of that desired. This is then changed successively to the second, third, and fourth multiples by utilizing the e.m.f. generated in one winding by the rotating field due to current of the next lower order frequency which is flowing in the other winding. The heavy circulating currents are obtained by tuning the respective windings, the output circuit being arranged to deliver energy to the antenna at the desired multiple frequency. The object of this method of obtaining radio frequencies is to use a comparatively low-speed machine rather than to attempt direct generation at the desired frequency, which requires the use of a high-speed machine having a large number of poles.

21. Static Frequency Multipliers. Present practice favors the use of static frequency multipliers where it is desired to use an alternator of comparatively low frequency. Two general methods, both of which depend upon the use of special transformers having d-c saturation windings, are employed. The first utilizes either two or three transformers connected in such a manner that the second or the third harmonic of the fundamental is in phase in the several output windings. The second may utilize but a single transformer, with a d-c saturation winding. The output winding is tuned to the desired harmonic frequency and receives its energy by "shock excitation." This is accomplished by so adjusting the d-c and a-c supply currents that voltage is induced in the secondary winding for only a small portion of a cycle of the supply frequency. In this manner harmonics of the fifth, and higher, orders may be obtained.

22. Arc transmitters are used, to some extent, for long-wave transoceanic work. There have been two main objections, however, to the use of such equipment. Most arc transmitters emit two frequencies, one for mark and the other for space. As there must be a sufficient frequency difference between these to allow of their being separated in the receiving equipment, one such transmitter really requires two communication channels for its operation. The other objection has been that most arc sets emitted strong harmonics. These can, however, be prevented from radiating strongly by proper shielding and the use of properly arranged circuits for feeding the antenna. Elimination of the space wave or "back wave" is rather difficult in transmitters of this type, especially when the output may be as high as 1,000 kw in large installations. The actual power output of the arc cannot be keyed, as the arc, to be stable, must draw a fairly constant current while in operation. Keying is generally accomplished by changing the inductance of the resonant circuit associated with the arc, thereby changing the frequency of the emitted wave. This is done by short-circuiting a few turns that are coupled to the main tuning inductance.

Methods have been proposed for shifting the output of the arc to a dummy antenna, or absorbing circuit, for keying the actual power radiated on but one frequency. Such methods have not come into general use.

The arc is operated from a d-c source, usually motor generators, at a voltage of from 300 to 3,500 volts depending upon the power rating of the unit. It burns in an atmosphere rich in hydrogen, which is supplied

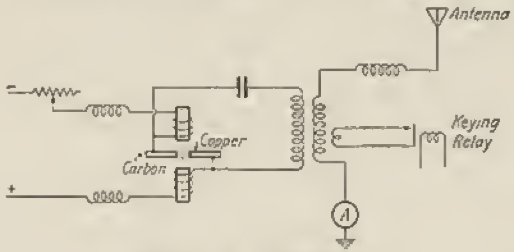


FIG. 3.—Arc transmitter.

by gas or by the vaporization of some such liquid as alcohol which is fed into the arc chamber. For the efficient production of undamped oscillations the arc must burn in a transverse magnetic field. This is supplied by a large electromagnet, the poles of which are respectively above and below the arc chamber and the coils of which are energized by passing the arc current through them. The intensity of magnetic field required for optimum results is inversely proportional to wave length and also depends upon the material used to furnish the hydrogenous atmosphere in the arc chamber. Values normally range from about 2 to 20 kilogausses. A water-cooled copper anode is used with a carbon cathode which is slowly rotated by means of a motor while the arc is in operation. A current-limiting resistor, normally used while striking the arc, is shorted out when the arc is running.

23. Tube transmitters have been used but little at frequencies between 14 and 30 kc for long-distance communication. Tubes to handle the power required have not been available until quite recently. This meant that a number of tubes had to be operated in parallel in the power-amplifier stage. Such transmitters have rated outputs of from 40 to 500 kw and are of the usual master-oscillator power-amplifier type.

24. Long-wave antennas of the various familiar types such as the T, inverted L, and umbrella have been used. Masts for these structures have, in some cases, been as high as 1,000 ft. Ordinarily they range from

400 to 800 ft. high. The technical problem is to get as many amperes in an antenna of as great an effective height as possible with a given power input. Voltages from antennas to ground may easily be 100 kv or more so that corona and insulation considerations place a limitation on the design. Of the total power supplied to the antenna, the useful portion is that radiated. The remainder is accounted for by conductor losses, coil losses, leakage, and corona (if present), and by loss in the resistance of the ground-return path. In a structure where most of the capacity is from the flat top to earth, and where the dimensions are considerably less than a wave length, the radiation resistance is given approximately by the relation $R = 1{,}600(H^2/\lambda^2)$, where H is the effective height of the antenna and λ the length of the radiated wave. Approximate calculation of H is possible in simple cases by summing up the products HI for all sections of the structure and dividing by the total current. This is done by calculating the capacities to earth of the various sections, and by measurement of the total value. Experimental methods of determining the capacity from small-size models are described by Lindenblad and Brown.[1]

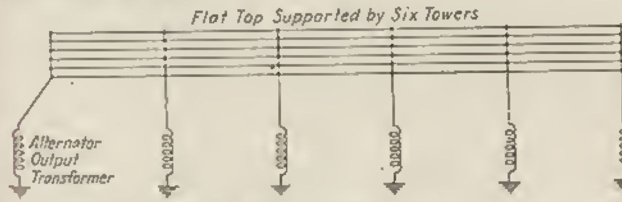


FIG. 4.—Multiple-tuned antenna.

25. The multiple-tuned antenna, consists of a long, flat top supported by towers and having down-leads at a number of points which pass through tuning inductances to earth. The total antenna current is the sum of all the currents measured at the base of the tuning coils. A system of buried wires and overhead conductors connected to them through current-equalizing coils is laid out to give a uniform distribution of current in the earth under the antenna. This is approximately the condition for minimum earth resistance. This uniform distribution is sometimes altered, by experiment, to still further reduce the losses. Such antenna and ground systems often have a total resistance of less than $\frac{1}{2}$ ohm. Total antenna currents of 700 amp. and more are obtained, by this means, from a transmitter output of 200 kw. For N tuning points the inductance of each down-lead and coil is approximately N times that which would resonate with the total antenna capacity at the desired frequency. The physical length of such an antenna for operation at 17 kc, or thereabouts, may be 1 or $1\frac{1}{2}$ miles, with as many as six tuning points.

26. Removal of Ice. In climates where sleet is experienced the antenna wires should be counterweighted, rather than solidly anchored, in order to lessen the chances of breakage. A heavy coating of sleet on the wires, with the attendant increase in sag, throws the antenna out of tune as well as endangering it mechanically. When this becomes

[1] LINDENBLAD, N., and W. W. BROWN, Main Consideration in Antenna Design, *Proc. I.R.E.*, June, 1926.

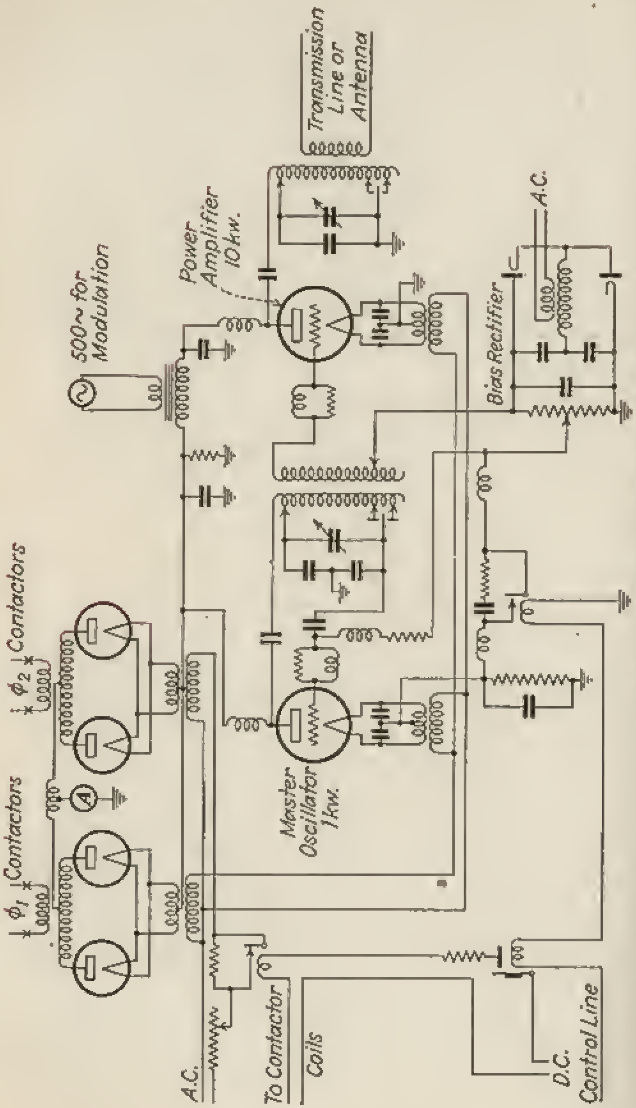


FIG. 5.—Marine coastal transmitter.

serious, it is necessary to melt the sleet from the wires in order to get normal antenna current. For this purpose break insulators and by-pass condensers are so arranged in the antenna wires that a series circuit of all (or part) of the wires is obtained at the low power-supply frequency. Special transformers supply power at about 2,000 volts for the purpose. This is sent through the antenna conductors just long enough to heat them sufficiently to melt off the sleet or ice.

27. Marine Transmitters. For marine work, tube transmitters are replacing the older spark and arc equipment. The radiated energy is confined more to a single frequency, which is essential for reducing interference; and systems for simultaneous transmission and reception, for break-in operation, and for remote control are much more easily built up by the use of tube transmitters. With a well-filtered plate supply the beat note obtained by use of a heterodyne or autodyne receiver is fairly pure, and its pitch can be changed at will by the receiving operator to suit conditions. For attracting the attention of ships standing by on a calling wave, or for working ships not equipped for heterodyne reception, the radiated energy can be modulated at an a-f rate.

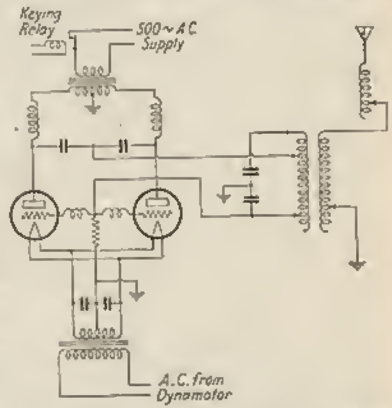


FIG. 6.—Essential circuit of i-c-w marine transmitter with a-c plate supply.

Transmitters for coastal stations usually have an output of from 5 to 10 kw. An air-cooled 1-kw tube functions as master oscillator and drives the 10-kw power-amplifier tube, which is of the water-cooled type. Plate supply is obtained from a full-wave kenotron rectifier, the output of which is filtered to some extent. Bias voltages are normally obtained from a small rectifier, to eliminate as much rotating machinery as possible. Filament supply is a.c. from step-down transformers. Because of the nature of the service, interruptions due to equipment trouble must be reduced to a minimum. For this reason two power-amplifier tubes are mounted so that either one can be used. Cooling water systems are provided in duplicate and equipped with pressure- or flow-operated relays which will shut down the transmitter in case of water failure. In some cases it is advisable to locate the antenna at a distance from the transmitter proper. A two-wire transmission line is used for this purpose, being matched to the power-amplifier and antenna-circuit impedances at its ends by means of air-core transformers.

To make the transmitter instantly available, the tube filaments are operated at reduced voltage, with plate supply off, when not in actual use. The "starting" relay operates contactors which apply full voltage to the filaments and close the low-voltage circuit to the plate-supply transformers. For remote control, the starting and keying relays can be operated from a single line by using double-current keying with a polar "keying" relay and a neutral line relay with weighted armature for "starting." The 500 ~ source, for production of icw, may also be relay operated. Wave change can be arranged by relay-operated contactors which change taps on the tuning inductances, these contactors being operated by a polar relay controlled from the operator's table.

28. Transmitters for shipboard use are generally of smaller power output than are those for coastal stations. Cost and space requirement are also important factors which must be kept down. The usual equipment is, therefore, more simple and compact than that treated above. The master-oscillator power-amplifier arrangement with d-c plate supply, or a.c. at a frequency of 350 cycles, meets the requirements very well in the intermediate frequency bands. The master oscillator holds the frequency steady regardless of changes in antenna capacity due to rolling of the ship, and the elimination of a separate rectifier saves space. Where space permits, a high-voltage d-c generator is used for plate supply. Medium power tubes require about a 2,000-volt supply. Change of wave is accomplished by changing taps on the tuning inductances. Choice of several frequencies in the band is provided by means of a multipoint switch operated from the front of the panel. The normal power-supply mains being d.c., a motor generator is required to furnish the plate-supply voltage. Another machine may furnish a.c. for the filaments. On small transmitters satisfactory keying can be effected in the low-voltage a-c plate supply by means of a relay controlled from the operator's key.

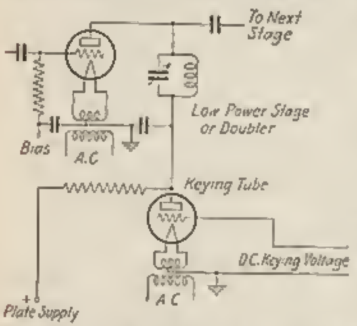


Fig. 7.—Tube keyer for transmitter.

29. Short-wave Technique. Channel spacings resulting from the ever greater demand for frequency and channel assignments, in the range from approximately 3,000 to 23,000 kc, require ever greater stability of the frequency of emitted carrier waves. Government regulations, based on international agreements, are yearly becoming more severe. To maintain a tolerance of plus or minus 0.01 per cent—which is what can be expected of a good short-wave transmitter—requires the use of either a very carefully stabilized and compensated tube oscillator or of some control device such as a quartz crystal. Crystal control has found most favor in this country to date.

Commercial short-wave code transmitters used for long-distance communication have an output of from 20 to 40 kw. The crystal is kept at a constant temperature and operates at one-eighth or one-fourth of the final frequency desired. The oscillator stage is followed by a screen-grid "buffer" stage, to isolate it from feedback and detuning effects, then by two or three frequency-doubling stages before the first amplifier stage operating at the signal frequency. Screen-grid tubes used in these stages, with proper shielding of tubes and circuits and filtering of supply leads, eliminate troublesome feedback effects without the use of neutralization. Water-cooled triodes used in the final power amplifier must be employed in a balanced stage with proper neutralization of feedback through the tube capacities. The tank circuit of the power

amplifier is coupled either directly, or through a transmission line, to the antenna.

For high-speed telegraphic operation the voltage regulation of all plate and bias supplies must be good. If poor regulation exists, the envelope shape of the characters will be triangular or irregular, instead of rectangular. (A small amount of lag may be introduced intentionally, in some cases, to round off the corners in order to eliminate trouble from keying clicks in near-by receivers.) For this reason hot-cathode mercury-vapor rectifiers are used for supplying the high d-c potentials required. These tubes, together with the high-voltage transformers, have very good voltage regulation at high values of output voltage.

For continued operation at keying speeds up to 250 words per minute (100 cycles per second) it is inadvisable to use a system of keying which employs electromechanical relays. A vacuum-tube keying stage is therefore used to key one of the low-power stages of the transmitter.

Where a plate supply having good regulation is not available, the load on it can be held constant by using two power amplifiers one of which supplies the antenna and the other a resistance load. Keying is accomplished by shifting the load from the main amplifier to the absorbing tube by biasing the amplifier grids below cutoff and bringing the absorbing tube grid bias up to such a value that the load drawn from the plate supply is the same as when the amplifier is supplying energy to the antenna. For receiving systems which rely partly upon frequency diversity of fading, it is desirable to modulate the wave radiated from the transmitter at an a.f. of something under 1,000 cycles per second. To prevent interference with signals on adjacent channels, this modulation should be reasonably free of harmonics. Otherwise, the higher order side bands will extend over into the adjacent channels and cause interference.

RECEIVING SYSTEMS AND EQUIPMENT

30. Long-wave Receivers. Long-wave receiving equipment must be designed to reduce trouble from static to a minimum and to separate transmitters differing in frequency by only about 200 cycles, which is the approximate spacing of assigned channels. The use of four efficient tuned circuits provides the required selectivity together with moderate ease of handling. For commercial work it has been the practice to obtain the h-f selectivity ahead of an aperiodic amplifier, then to go to a heterodyne detector of either the single-tube or balanced-modulator type which is followed by as much a-f amplification as is required. The final selectivity may, if necessary, be obtained by the use of narrow a-f band-pass filters. For complete separation of signals on adjacent channels this is often necessary. Owing to the difficulty of obtaining complete shielding at these comparatively low radio frequencies, it is generally advisable to use astatic pairs of coils in all tuned circuits, couplers, oscillators, etc., in addition to the use of a reasonable amount of shielding. Transformers and couplers are built with electrostatic shields to prevent capacity coupling, where this is undesirable.

In a multiplex receiving station, where it may be necessary to receive from 10 to 20 signals from approximately the same direction, a single aperiodic antenna system is the most economical and practical. The individual receivers are fed by means of "coupling tubes" operated from a common, or from individual, antenna-output transformers. All tuning is done beyond these coupling tubes so that operation of the individual receivers is entirely independent of all others.

31. Directional Antennas. Reduction of static is accomplished by the use of directive-antenna systems. Arrays of large loops, or of loop and vertical combinations, are one means of obtaining directivity. Where the nature of the soil is such as to produce a considerable tilt of the wave front, the Beverage wave antenna is used to advantage. This antenna consists of one or two wires strung on poles at a height of about 20 ft. and extending in the direction of the desired signal for a distance of approximately one wave length. The antenna is highly directional, and small signal voltages obtained from stations to the rear can be compensated for by feeding into the signal circuit a small voltage of proper amplitude and phase obtained from the damping resistance connected between antenna and ground, or by setting up reflections in the antenna itself.

As keying speeds on long-wave transoceanic circuits seldom exceed 100 words per minute (40 cycles per second) and signal strengths are

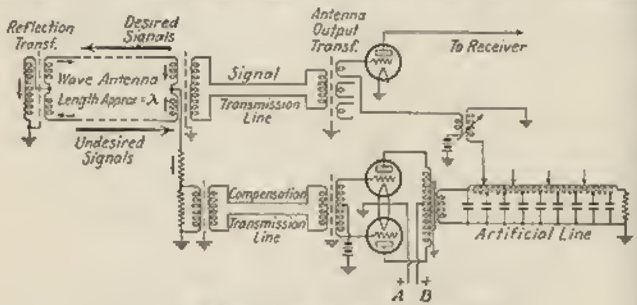


FIG. 8.—Wave antenna and output circuits.

steady, such a channel requires only a total band width of about 160 cycles. Frequency variations of the transmitters can be kept within about 0.1 per cent or 20 cycles in 20,000, and heterodyne oscillators used for reception should have as good stability.

32. Ship-to-shore Receivers. Receiving equipment for ship-to-shore service must cover the frequency range of 500 down to 14 kc in order to operate in the regular marine bands and also to receive broadcasts and time signals from high-powered long-wave stations. Receivers for shipboard use are of the autodyne type embodying a tuned antenna circuit coupled to the oscillating detector, which latter has a "tickler coil" for regeneration control and generally two stages of a-f amplification. By means of tapped inductances the receiver may tune from about 1,000 down to 60 kc. For the lower frequencies a set of loading inductances is used. The chief requirements are ease of operation and rapidity of tuning. Regeneration control allows the receiver to be operated oscillating for cw reception or non-oscillating for reception of spark, icw, or modulated signals. Provision is made for disconnecting the receiver from the antenna when transmitting.

Important coastal stations have separate receivers to cover the lower and higher frequency marine bands of approximately 115 to 171 kc and

375 to 500 kc, respectively. Such receivers should have but a single tuning control and, to obtain the required selectivity, should be of the superheterodyne type. An i-f oscillator, which can be used at will by the operator, must be provided for cw reception. The over-all selectivity should be such that a total band width of not more than 1 kc is passed at 80 per cent peak response.

As in long-wave reception, reduction of static and interference is accomplished by the use of directive antennas. For the lower frequency band the Beverage wave antenna has the advantage of relatively large pickup; good directivity with compensation, and the ability to supply a number of receivers operating at the same or different frequencies. Where reception from all directions is required and for the higher frequency bands where the wave antenna is unsuitable for night reception, antennas of the flat top, inverted L, T, vertical, or loop types are employed. The loop and vertical combination, giving a cardioid directive diagram, can be arranged with crossed loops and a goniometer so that the operator can rotate his antenna reception diagram at will.

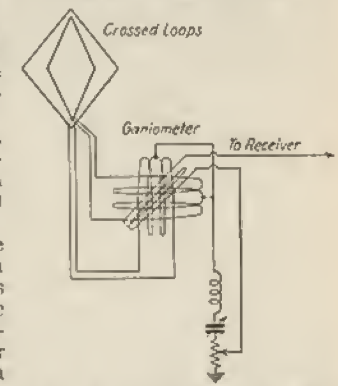


FIG. 9.—Loop-vertical antenna for directive reception.

33. Short-wave receiving equipment, for the reception of commercial radiotelegraph signals, comprises two general classes, *viz.*, (a) point to point and (b) mobile.

For commercial point-to-point service the receiving equipment must deliver a signal which is as nearly perfect as is possible. This requires a high degree of frequency stability, the best practicable over-all selectivity, and means for reducing the effects of fading to a minimum. The receiver should have a total band width such that it will provide an attenuation of at least 60 db at the frequencies of the channels adjacent to that on which reception is being carried on. In calculating selectivity requirements, the assigned channel spacing must be reduced by twice the frequency tolerance permitted on each channel. This gives the frequency spacing between two signals on adjacent channels, when the frequencies of the two transmitters have drifted toward each other. Protection against all other types of interference, such as those encountered in superheterodyne receivers, should be not less than 70 db. At the same time, the useful band width must be sufficiently great so that no undue amount of attention will be required to keep signals fairly well centered in the pass band of the receiver. With present-day stability of transmitter frequencies, and of receivers, this means a useful band width of from 1 to 4 kc depending upon the carrier frequency.

Present-day receivers, to provide the required performance, are generally of the multiple-detection, or superheterodyne, type in which one or two i-f systems are employed. It is only by the use of a relatively low final i.f. that the necessary selectivity and useful band widths can be obtained. The required i-f characteristics are obtained by use of either a band-pass filter or a number of stages of amplification employing one or more tuned transformers per stage. Choice of more than one band width in the i-f system is highly desirable and often necessary.

In equipment used for high-speed automatic operation, the signal is amplified, beat down to a lower frequency, and then rectified. The rectified output, consisting of short and long pulses of d.c., is used to operate a relay of either the electromechanical or vacuum-tube type. The former operates into a simplex, duplexed, or quadruplexed d-c telegraph line to the central traffic office. The tube relay, or "keyer," controls the signal fed to the tone line from a local a-f source. The receiving operator is thus supplied with an audio signal of constant frequency and intensity regardless of any changes in the actual radio signal which are not great enough to make it drop out of the receiver. By means of a-f filters six or more keyed tones of this sort may be handled over a single, two-wire tone line.

To minimize the effects of fading, receiving equipment is arranged to take advantage of the diversity of fading existing, at a given instant, either on slightly different frequencies at the same location or on the same

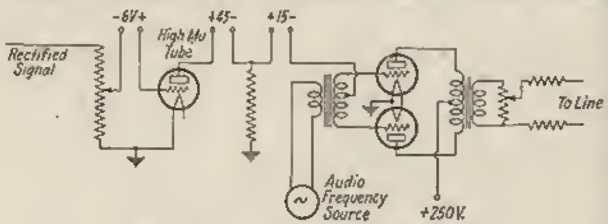


FIG. 10.—Tone keyer for receivers.

frequency at points separated 10 wave lengths or more apart. *Frequency diversity*, in practice, is most economically obtained by modulating the carrier with an a.f. of not higher than 1,000 cps, and preferably of not higher than 500 cps, in order to minimize interference to signals on adjacent channels. This results in radiation on the carrier and on an upper and a lower frequency. If the band width of the receiver is sufficient to pass these three frequencies and if the normal signal strength on any one of these frequencies is sufficient to operate the keying device, considerable diverse fading on the several frequencies received can be tolerated. In spite of the fact that a lesser peak voltage can be obtained from a modulated signal than from a pure cw signal, considerable improvement is obtained, under practical conditions of fading, by its use. Where *space diversity* is utilized, a pure, unmodulated signal is to be preferred. In this case two or three separate receivers are fed from separate directive antennas spaced 10 wave lengths or more apart. The rectified outputs from these receivers are combined and made to operate the keying device. Confining the radiated energy to a single frequency means greater signal strength for a given transmitter power, and combination after rectification eliminates the consideration of instantaneous phase relations which might be such as to cancel rather than add.

34. Use of Limiting Circuits. Under conditions of high signal-noise ratio and violent fading, the use of considerable limiting in the receiving equipment is desirable. This should be done following the final selec-

tivity and must be in a system having small enough time constants so that the decaying transients occurring after each overload do not occupy an appreciable portion of the interval between characters. In order to use such limiting successfully, it is essential, as stated before, to pass up to about the fifth harmonic of the keying frequency. If this is not done, wide variations in mark/space ratio of the final signal will occur as the degree of limiting varies with the signal strength.

Character formation can be maintained, in some cases of overloaded systems, by the use of a so-called "sliding bias" on the rectifier. The signal may be amplified up to some 30 or 40 volts maximum value and applied to the grid circuit of a rectifier tube which begins to take grid current at a relatively low applied signal voltage. By proper choice of grid- and plate-circuit resistors, and the use of a condenser across the grid-circuit resistor to give a relatively large time constant, only the tops of the character envelopes will be effective. In using such a system, however, reliance must be placed upon some form of diversity reception to prevent drop-outs, and splitting of characters, due to rapid fading.

Recent practice has been to use some system of automatically controlling the gain (A.G.C.) of the r-f amplifier stages. The circuits are similar to those used in broadcast receivers and are superior to those which operate on the final detectors, because they minimize overloading r-f and i-f amplifiers and first detectors.

35. Commercial Receiving-center Problems. In a large receiving station for long-distance communication there may be from 10 to 100 individual receivers installed and intended for simultaneous operation. To do this requires that each unit be effectively shielded and that all battery-supply leads be well filtered for the frequencies at which the respective units operate. High-frequency equipment must also be protected from l-f voltages which might be present on the battery supply busses, as such voltages may cause undesirable modulation of signals if allowed to get to the tube circuits. Transmission lines, where used, must be of a type which has negligible stray pickup and radiation. Satisfactory types of line, depending upon the equipment with which it is to be used, are (*a*) the balanced four-wire line, (*b*) the two-wire transposed line, and (*c*) the concentric-pipe line. The first consists of four wires arranged at the corners of an imaginary square, diagonally opposite wires being connected together at both ends of the line. The four-wire and two-wire types are used where the system is to be kept balanced with respect to earth. Antenna systems which operate against earth generally use the concentric-pipe line in which the outer pipe is grounded. The two types of systems are sometimes connected together by means of suitable tuned transformers.

To obtain the full benefits of good shielding, stray feedback through the battery-supply leads must be eliminated by means of properly proportioned, and located, filter circuits. This is of especial importance in short-wave equipment and in medium-wave equipment for marine coastal station use.

36. Power supply for commercial receiving equipment must be absolutely reliable and not subject to interruption. Storage batteries operated on either a floating or a charge and discharge basis are used for this service.

Charging equipment consists of motor-generator sets for filament batteries, where relatively heavy currents are required, and either motor generators or rectifiers for batteries of smaller rating such as used for plate and bias supply. Where receiving antennas may be located fairly close to the building that houses the charging equipment, this must be located in a specially shielded room to prevent direct radiation into the antennas. Equipment used for floating batteries that are in service must be provided with effective filtering between it and the battery and load bus.

Where the nature of the radio service does not warrant the expense of installing and maintaining storage batteries, reliance may have to be placed on the continuity and reliability of a-c power service provided by the local power company. In such cases the most economical and flexible arrangement for a small station is to provide each receiver with its own filament transformer and its own plate and bias supply rectifiers.

An emergency power supply should be provided in all cases of a-c operated equipment. Where storage batteries are installed for supplying the receivers during power failures, additional emergency power supply may or may not be necessary. In some cases the cheapest arrangement may be a battery installation that will take care of normal short-period outages and an emergency power plant to care for longer periods of failure of the public power service.

CONTROL METHODS AND EQUIPMENT

37. Central Office. In commercial radiotelegraphic systems the transmitters are controlled from a central traffic office, and received signals are conveyed to this central office from the receiving station by land lines. Transmitting and receiving stations are, in some cases, as much as 500 miles distant from the central office. The tendency, however, is to keep this distance below 100 miles to reduce initial and maintenance costs, or rentals, of land lines. Long control and tone lines are justified only if a distant location of the transmitter will effect a considerable saving in the power required to obtain satisfactory service, or if the distant receiving site is considerably superior to near-by ones in signal-noise ratio. In long-wave transoceanic and medium-wave marine work the use of long land lines is often well worth while. In short-wave work the over-all results are not so dependent upon geographical location. Suitable sites are generally available within 100 miles of the city to be served.

38. Automatic Transmitters. In "automatic" operation of code circuits a tough paper tape is perforated by means of a machine which has a keyboard similar to that of standard typewriters. This tape is then fed through the "automatic transmitter" in which two cam-operated steel rods come up against the tape at every point where a perforation might exist. Where one is, the rod goes on through, and a contact operated by a lever on the lower end of the rod is closed. These two rods controlling the "make" and "break" contacts alternate in coming against the tape and are sufficiently offset in the direction of travel of the tape so that perforations in the upper (make) and lower (break) rows, when opposite the same center hole, give a dot and when opposite adjacent center holes give a dash. (Sample tape appears below.)

T H I S I S A S A M P L E

The two contacts supply current, in opposite directions, to a polar relay which, in turn, keys the control circuit going to the transmitting station. For speeds much above 100 words per minute it is desirable to have as few mechanical relays as possible between this main polar relay and the keying circuit of the radio transmitter. The time required for a relay armature to travel from one contact to the other, while short, becomes important when the duration of a dot is less than 0.010 sec.

Printing telegraph equipment employs a special model of automatic tape transmitter, which is adapted to the different code used for such systems.

In installations of multiplex equipment employing the principle of time division, automatic tape transmitters supplying the several channels are synchronized and phased to give the required over-all performance of the multiplex system.

39. Tone-control Circuits. Where only a few transmitters are to be controlled from one point, d-c double-current keying is the most economical and satisfactory. A complete metallic circuit is to be preferred

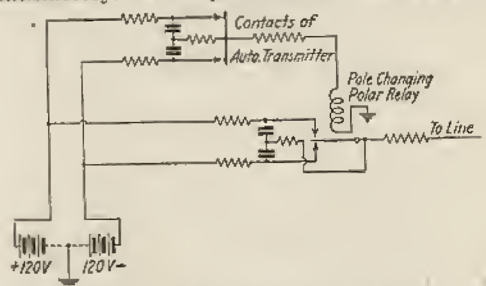


FIG. 11.—Double-current control circuits.

to a single wire with ground return, although the latter is entirely satisfactory in many cases.

In a large central-office system the number of control lines required can be greatly reduced by the use of multiplex tone, or "voice-frequency carrier," control. By the use of a number of different frequencies and band-pass filters at both ends of the circuit as many as 10 channels can be obtained on a two-wire line which will pass frequencies from about 400 cycles up to 2,500 cycles with approximately equal attenuation. In one such type of equipment the a-f supply is a multifrequency inductor-type alternator having a separate winding and rotor for each frequency. Energy from this machine is keyed by means of either electromechanical or vacuum-tube relays which are controlled by the automatic tape transmitter and supply current to the control line. Band-pass filters in the individual control channels reduce the harmonic content of the signal supplied to the line to a low value and also round off the corners of the square keying envelopes.

The band width required in filters for tone-control work depends (1) upon the maximum keying speed which must be handled and (2) upon the fidelity of envelope shape required for the particular application. Where great fidelity is not required or where the over-all transmission gain of line and associated equipment does not vary more than about

20 per cent, it is sufficient to pass the second harmonic of the keying frequency. This means a total band width of four times the keying frequency. To obtain fairly square envelope shape, with a mark-to-space ratio of about 60:40, it is necessary to pass up to the third harmonic or a total band of six times the keying frequency, at least.

For the lengths of line normally used between central offices and outlying stations, and for present-day code keying speeds, the matter of phase distortion due to the line is of relatively small importance.

40. Control equipment used at transmitting stations may be of either the d-c or tone-operated type, depending upon the system used at the central office. In a double-current d-c system the conventional polarized telegraph relay is used as a main-line relay for speeds up to some hundred words per minute. Where normal operating speeds run much above 100 words per minute, special high-speed relays of the polarized type must be used. Large keying and compensation relays and contactors used in long-wave transmitters are controlled by the line relay or a heavier intermediate relay. In tube sets—especially short-wave equipment—higher keying speeds are possible and require the use of a minimum number of mechanical relays. For d-c control the main-line relay may operate directly into a tube keyer incorporated in the transmitter.

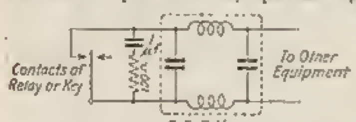


FIG. 12.—Spark absorber and click filter.

In tone-control systems the equipment at the transmitting station comprises band-pass filters and amplifier-rectifier units. The rectified output may be used to operate either electromechanical relays or tube keyers. Where such equipment is used at large high-powered transmitting stations, it may have to be protected from stray fields of the transmitters, transmission lines, and antennas. The amount and disposition of shielding and filtering required by control equipment and associated wiring depend on numerous factors such as the following: (1) lowest frequency radio transmitter installed or contemplated, (2) highest control frequency planned, (3) r-f field intensities, and (4) level of control signals and voltages. It will be obvious that a high-power long-wave transmitter operating on a frequency of about 20 kc will create serious problems where it is desired to employ control channels ranging in frequency from, say, 400 to 20,000 cps or higher.

Tube keyers, while more elaborate than the usual mechanical relays, are capable of operating at practically any speed desired. They also eliminate relay maintenance and adjustment. In the simpler arrangements the control tone is amplified, rectified by either a two-element or a three-element tube rectifier, then passed through a smoothing circuit or low-pass filter. The d-c pulses thus obtained are applied to the control elements of the keying-stage tube or tubes.

41. Received Signal Transfer. Systems for transferring signals from the receiving station to the central office are similar to the transmitter-control systems. In short-wave work the actual radio signal, after heterodyne detection, is amplified and rectified and applied to a tube keyer. This may be arranged to supply d.c., or tone, for transfer to the traffic office. Audio-frequency filters, of the same type used for tone control, allow a number of channels to be handled over one line.

Where tone lines are long enough to require the use of one or more repeaters, care must be taken that the sum of the voltages of all channels is not high enough to cause any overloading of the repeaters. If this takes place, intermodulation between channels will be caused, which results in mutilated signals at the central office. With repeatered lines and the usual band-pass filters, it is essential that all channels be kept at approximately the same signal level. A maximum difference of 2:1 between any two channels should not be exceeded. Large differences in channel levels are apt to cause interference on the weaker ones.

In medium-wave and short-wave receiving stations the contacts of all telegraph keys and relays must be prevented from sparking, and the wires to and from the contacts must be properly filtered. If these precautions are not taken, serious click interference will be experienced in the receiving equipment. The same applies to commutator-type electric motors. Circuit breakers should preferably be located in a shielded room.

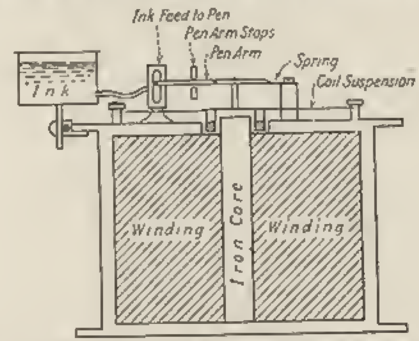


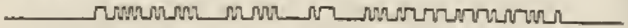
FIG. 13.—Ink recorder. Paper tape and tape guide not shown.

TRANSCRIBING METHODS AND EQUIPMENT

42. High-speed Reception. As the average operator copies at a rate of only about 40 words per minute, aural reception must be replaced by some method in which a record is made of the signal, on the high-speed circuits, the recorded signal then being copied off at a slower speed by one or more operators. The older dictaphone and photographic methods of recording were not entirely satisfactory. Most systems now use some form of "ink recorder" in which the movement of a pen is controlled by the incoming signal and makes short and long characters on a moving paper tape.

Reception by tape has the double advantage of speed and of there being a record to which the operator may refer or which may be looked up later in case any question arises.

43. Ink Recorder. One commonly used type of ink recorder consists of a small coil suspended in a strong unidirectional magnetic field supplied by an electromagnet. The signal is amplified and rectified and the d-c pulses sent through the recorder coil which, in turn, moves the pen arm up against an upper stop. With no signal current flowing, the pen is held against the lower stop by the spring of the pen arm and coil suspension. To improve the action of the device at high speeds, the coil is suspended midway between the stops, and current reversals are used, in place of pulsating d.c., to operate the coil. This is obtained from a pole-changing relay operated by the rectified signal, or from a special amplifier-rectifier unit which gives an output d.c. in opposite directions for "mark" and "space."



T H I S I S A S A M P L E

44. Printers. Where printing telegraph equipment is employed, manual transcription of the incoming signal is eliminated. The printed tape coming from the receiving machine is simply pasted on message blanks. Errors may be corrected by obtaining the required correction from the distant radio terminal, and pasting it over the original which contained the error.

References

AUSTIN, L. W.: *Proc. I.R.E.*, June, 1926.

BEVERAGE, RICE, and KELLOGG: The Wave Antenna, a New Type of Highly Directive Antenna, *Proc. A.I.E.E.*, March, 1923.

BYRNES and COLEMAN: 20- to 40 kw High-frequency Transmitters, *Proc. I.R.E.*, March, 1930.

CARSON, JOHN R.: "Electric Circuit Theory and the Operational Calculus," McGraw-Hill Book Company, Inc., New York.

ESPENSCHIED, ANDERSON, and BAILEY: Transatlantic Radio Telephone Transmission, *Proc. I.R.E.*, February, 1926.

FULLER, L. F.: The Design of Poulsen Arc Converters for Radio Telegraphy, *Proc. I.R.E.*, October, 1919.

SHEA, T. E.: "Transmission Networks and Wave Filters," D. Van Nostrand Company, Inc., New York.

STEINER and MASER: Hot-cathode Mercury-vapor Rectifier Tubes, *Proc. I.R.E.* January, 1930.

SECTION 17

AIRCRAFT RADIO

By HARRY DIAMOND[1]

1. Importance of Radio Communication to Aircraft. The success of any transportation system depends in a large measure upon the rigorous maintenance of safe, scheduled operation. Probably nothing has contributed more to the safety and reliability of transportation systems than the associated communication systems. Radiotelegraph, radiotelephone, the radio beacon, and the radio direction finder have been important elements to such safety in both sea and air transportation.

Radio serves as a communication means between airplanes and between airplane and ground. It furnishes the pilot with weather information, tells him when he is on or off his course, helps him to land under conditions of poor visibility, and is beginning to be of value in preventing collision with other planes or with fixed objects. It provides the operations office continuous contact with each aircraft in flight and thereby affords full control of all flight operations to conform with existing meteorological conditions and traffic requirements. For the airport traffic manager it furnishes a rapid and certain means for communicating with arriving or departing airplanes and directing their landings or take-offs in a safe and orderly sequence. For the weather man it serves as a useful tool in the accumulation of upper-air weather data needed in making his forecasts.

2. Organization of Civil Radio Facilities. Aviation radio facilities may be broadly classified according to who furnishes the service: (1) government systems; (2) transport company systems; (3) airport operator facilities. As the government system is operated for the benefit of all fliers, the transport companies plan their systems so as to incorporate the service rendered by the government. The airport operators, in turn, design their radio facilities to tie in efficiently with both the government and the transport company systems.

1. *The government*, through the agency of the Civil Aeronautics Authority, has constructed a network of radiotelephone broadcast stations for the dissemination of weather information to aircraft in flight and a system of radio range beacons supplemented by radio marker beacons for the guidance of aircraft over the civil airways. The provision of radio-landing aids at terminal airports to facilitate the landing of airplanes under adverse visibility conditions has been begun. CAA also operates an extensive system of teletype lines for the collection of weather information to be used in the radiotelephone broadcasts and for airways traffic control. In collecting weather information this agency has the coopera-

[1] Principal Radio Physicist, National Bureau of Standards.

tion of the U. S. Weather Bureau which maintains a large number of weather stations at the airports and at points off the airways. Practically all these stations are on the weather teletype network. The teletype lines interconnecting the airways traffic-control offices form a separate network to facilitate the control of some 20,000 military, commercial, and private airplanes flying the airways.

On July 1, 1940, there were approximately 28,000 miles of lighted airways in the United States, practically all of which were radio equipped. Some 27,000 miles of teletype were in use in the weather network and 10,000 miles in the traffic control network. The radio facilities included nearly 250 radio range beacons with voice broadcast facilities at each of these stations, 50 l-f radio marker beacons at strategic points on the airways, 180 u-h-f cone markers for giving positive indication of the location of the beacon stations, and 115 fan-type markers for defining control points along the airways at which arriving airplanes are kept (at various altitudes) while awaiting permission to land during adverse visibility conditions. Ten radio-landing installations were in the process of completion to afford service tests under actual airway conditions.

2. *The air transport companies* have adopted and installed two-way communication equipment at approximately 200-mile intervals along the airways of the nation and in all their airplanes. This system permits continuous contact between the offices of each company and their aircraft in flight, thereby allowing flight operations to be controlled according to existing weather conditions and traffic requirements. The transport companies also operate teletype circuits and point-to-point radio stations which provide the rapid communication between operating offices which is essential to the successful operation of high-speed passenger, mail, and express service. The facilities of the different air transport companies are coordinated through Aeronautical Radio Inc., an association organized for this purpose and having its headquarters in Washington, D. C. The radio facilities already enumerated are sufficient for air transport companies operating over the civil airways of the United States. In the case of international routes, such as the route to South America, the transpacific route, and the transatlantic route, the operating companies must provide the additional facilities which are necessary for the guidance of their aircraft.

3. *Airport operators* provide as standard airport equipment short-range two-way communication equipment used in directing from a central point all take-off and landing maneuvers of transport, military, and private airplanes. Voice communication with arriving and departing airplanes within a 25-mile radius is essential to the safe and efficient operation of a busy airport.

3. Military Radio Facilities. The communication and navigational requirements of military aircraft are naturally considerably different than for civil aircraft. Here, the emphasis is on mobility and flexibility of both the ground station and aircraft equipment. Operation is required over geographical areas rather than along fixed routes. Simplicity of radio equipment is paramount whereas the service conditions are generally more difficult. The research work carried on by the military agencies to secure suitable equipment and methods exerts great influence on the state of the art. The military radio developments are of particular applicability to civil air transport operation on routes outside the United States, as to South America or in the transoceanic service.

TABLE I.—RADIO FREQUENCIES IN CIVIL AVIATION

Service	Present setup	Proposed u-h-f setup
Radiotelephone weather broadcast and radio range beacon	200–400 kc (49 shared, 9 exclusive frequencies)	123.000–126.000 Mc (31 exclusive frequencies); 126.000–127.000 Mc (10 shared frequencies)
Airport traffic control	278 kc	129.300, 129.780, 130.300, 130.860, 131.420, 131.480 Mc
Two-way communication between airplane and ground	2,900–3,500 kc (night), 4,100–6,600 kc (day) (80 frequencies)	140.240–143.880 Mc (28 frequencies)
National calling and working and itinerant service	3,105, 3,120 kc (night), 6,210 kc (day)	140.100 Mc
Instrument landing group:		
Runway localizer beacon*	....	109.500, 109.900, 110.300 Mc
Landing beam*	....	93.500, 93.900, 94.300 Mc
Radio marker beacon*	75.000 Mc	75.000 Mc
Radio teletype	....	60.180–65.860 Mc (45 frequencies)
Aviation instruction group	....	33.420–39.060 Mc (4 frequencies)
Transport company point to point	2,700–18,000 kc (23 frequencies)	
Miscellaneous aviation services group:		
Public message traffic, collision prevention, radio altimeter, and others	....	Frequency requirements at present unknown

* 400-kc guard bands.

4. Radio Frequencies in Civil Aviation. The radio frequencies used for the various radio aids are indicated in Table I. Except for the instrument landing and radio marker beacon groups, service at the time of writing has been largely in the lower frequency ranges. The radiotelephone weather broadcast stations and the radio range beacons operate in the 200- to 400-kc band. Airport traffic-control transmitters operate at 278 kc. The air transport company communication systems use frequencies from 2,900 to 6,600 kc. However, it is now planned to move the different facilities into the u-h-f region, as shown in the table. By 1945 use of the lower frequencies will probably be limited only to such service as cannot be afforded at ultra-high frequencies. The reasons dictating the move are as follows:

Atmospheric disturbances arising from electrical storms have constituted a service limitation to reception in the 200- to 400-kc band and to somewhat lesser degree in the 2,900- to 6,600-kc band. Another form of disturbance, called *precipitation static* and of importance only in aircraft reception, constitutes a second, and often even more serious, limitation to reception at these frequencies. This form of disturbance has been found to accompany rain, snow, and even sand storms and appears to be caused by oscillating corona discharge from points on the airplane fuselage to the surrounding atmosphere. Its intensity is often sufficient to paralyze the ground-to-airplane services. Accumulated experience indicates that reception on ultra-high frequencies is practically free from atmospheric disturbances and is to an appreciable extent, less influenced by precipitation static.

A second advantage of u-h-f propagation is the freedom from dependence upon ionospheric conditions. In the present communication band such dependence results in severely fluctuating received signal intensities and renders these frequencies generally unsuitable for direction determination by either transmission or reception. Even in the beacon band, ionosphere propagation (at night) tends to prove troublesome. It is to be noted, however, that u-h-f propagation is not entirely free from variable effects owing to variable tropospheric bending.

A third advantage of the ultra highs is the greater directivity of transmission or reception possible. This is important, *e.g.*, in the radio range-beacon service. At low frequencies the only directive patterns available are the figure of eight and the cardiod. In forming a course with such patterns, considerable radiation exists in directions at large angles to the course. The return of such radiation to the course, *e.g.*, by reradiation or by reflection from mountain sides, produces an interference pattern which results in bent and multiple courses. The possibility of using more directive patterns at ultra-high frequencies offers means for reducing the side radiations and hence the troublesome effects described.

Finally, the rapidly expanding aviation radio facilities require an increasing number of r-f channels which are not available in the portion of the spectrum hitherto utilized.

5. Propagation Characteristics of Aviation Radio Frequencies. Because of the widely different radio frequencies used in aviation, a complete discussion of their propagation characteristics would require a volume. A few words on their more general characteristics will, however, be given here.

The l-f services, 200 to 400 kc, rely upon ground-wave propagation. Sky-wave propagation, at night, sets a limit to the distance separation between stations operating on the same or adjacent frequencies. Graph 2 of Fig. 1 shows the ground wave corresponding to a ground conductivity typical of the plains regions, while graph 3 shows the same data corresponding to a ground conductivity characteristic of the mountainous regions. Graph 4 shows the estimated sky-wave intensity. The various intensities are computed along the direction of maximum radiation (45 deg. off course). Figure 2 differs from Fig. 1 only in the frequency of operation (371 instead of 200 kc). Based on a minimum service field intensity of 50 μv per meter and a maximum tolerable interfering field intensity of 12.5 μv per meter, two stations of this type operating on the same r.f. may be spaced within 400 to 500 miles of each other at 200 kc. At 400 kc, the minimum allowable spacing increases to 600 to 800 miles. The allowable spacing corresponding to operation on adjacent frequencies (3-kc separation) depends on the selectivity of the average receiver used and varies from 200 to 400 miles depending on the ground conductivity and the operating frequency.

Radio wave propagation in the h-f communication band depends chiefly on sky-wave radiation returned to earth from the ionized layers. The ground wave is generally of negligible importance beyond distances of about 30 miles. The transmission characteristics are therefore dependent upon highly variable phenomena and cannot be definitely specified. An approximate idea of the

daytime propagation at these frequencies is given in Figs. 3 to 5, based on data obtained by Bell Telephone Laboratories in 1929.

From these graphs it is seen that the higher frequency appears to be best suited to daytime operation. This has been borne out in practical operation, so that the daytime working frequencies throughout the country are of this order.

Similar graphs for transmission during night, showing field strength as a function of distance, are given in Fig. 6. It is even more difficult to generalize from these graphs than for the case of daytime transmission, the movement of the ionized layer involved being more erratic. The graphs do show, however, that the lower frequencies are more reliable for nighttime transmission, the transmission on 5,690 kc being unsatisfactory due to excessive fading.

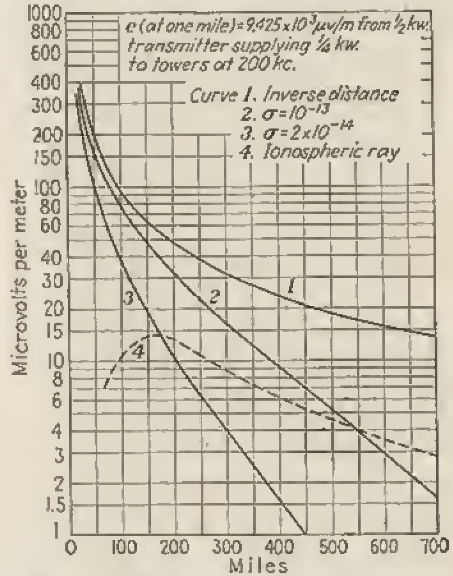


FIG. 1.—Field-intensity attenuation of 500-watt radio range-beacon station at 200 kc, 125-ft. towers.

Experience has shown that the choice of day and night communication frequencies, shown in Table 1, was a wise one considering the non-availability of ultra-high frequencies at that time.

The u-h-f services depend upon ground-wave propagation. The ground wave may be considered to consist of three components as follows: (1) the direct wave which travels directly between the transmitting and receiving antennas, (2) the ground-reflected wave which reaches the receiving antenna only after reflection from the ground surface, (3) the surface wave which is the component of the ground wave remaining when both the transmitting and receiving antennas are at zero height—at grazing incidence, the reflection coefficient of the ground is −1 so that the direct and ground-reflected waves cancel.

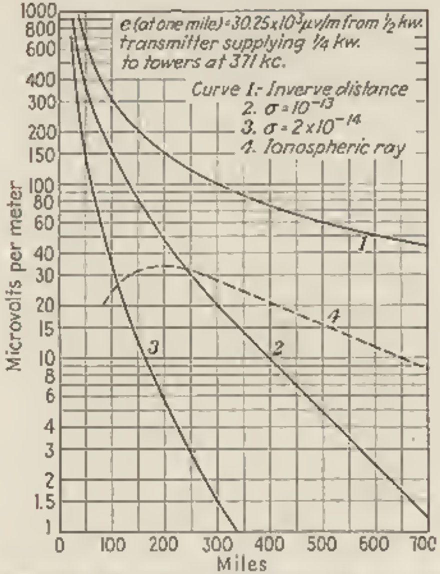


FIG. 2.—Field-intensity attenuation of 500-watt radio range-beacon station at 371 kc, 125-ft. towers.

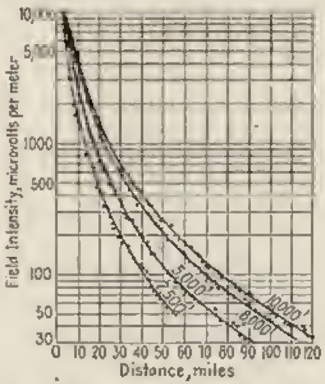


FIG. 3.—Average strength of daytime signals received in an airplane from 500-watt station on 1,510 kc. (Airplane at altitudes designated on graphs.)

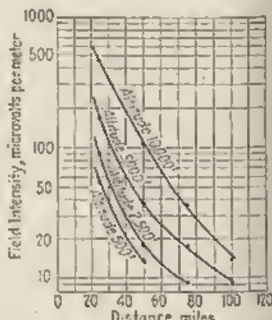


FIG. 4.—Reception from airplane using 50-watt transmitter on 1,625 kc. (Airplane at altitudes designated on graphs.)

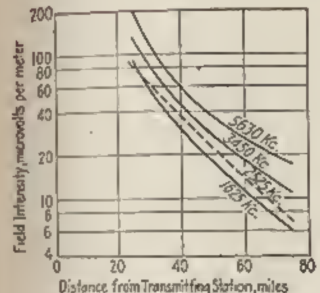


FIG. 5.—Effect of frequency on attenuation of 500-watt ground station.

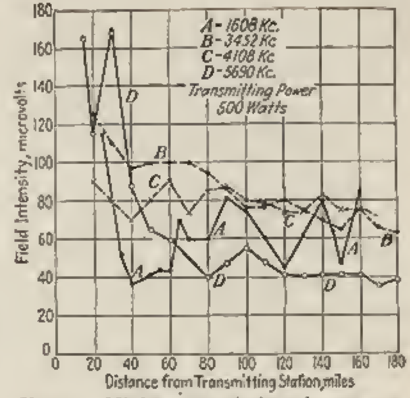


FIG. 6.—Night transmission phenomena.

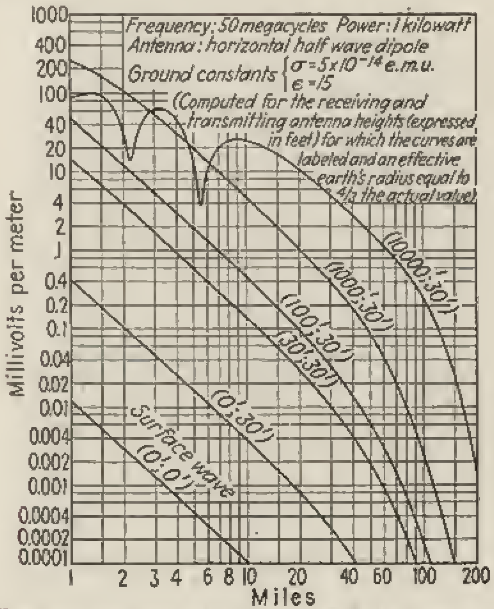


FIG. 7.—Field-intensity attenuation of 1-kw 50-Mc transmitting station (horizontal transmitting and receiving antennas at altitudes designated on graphs).

In Fig. 7 horizontal polarization is considered, and average electrical properties of the ground are taken. The graphs take into consideration the spherical shape of the earth and the average decrease in dielectric constant of the lower atmosphere with altitude (producing bending). The graphs show the striking dependence of the received field intensity upon the heights of the transmitting and receiving antennas. It is this feature which renders ultra-high frequencies peculiarly adaptable to aviation use.

GROUND-STATION EQUIPMENT

6. CAA Radio Range-beacon and Weather-broadcast Stations (200 to 400 kc). The most modern CAA installation, used at nearly 100 locations, employs a transmitter having two independent r-f channels controlled by two matched A-cut quartz plates 1,020 cps apart. Complete stand-by equipment with an automatic transfer relay for placing

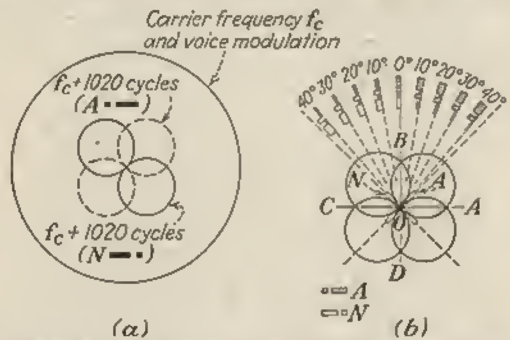


FIG. 8.—(a) Space radiation pattern of simultaneous radio range-beacon and weather-broadcast station; (b) range-beacon signals produced.

it in service in event of failure of the regular unit is provided. The antenna system comprises five self-supporting base-insulated steel towers 125 ft. high, four of which are placed on the corners of a square 300 to 500 ft. on a side and the fifth at the geometrical center of the square. One of the r-f channels of the transmitter delivers 400 watts of carrier power (which may be modulated 70 per cent by speech) to the central antenna. The other delivers 275 watts of unmodulated carrier power to a coupling system which feeds the four corner antennas. In the absence of speech modulation, the setup forms a single side-band system having 1,020-cycle modulation; the carrier is radiated non-directionally by the central radiator, whereas the side band has the characteristic radiation of the radio range beacon (see Fig. 8a). When special modulation is applied, the central tower radiates, in addition, the speech side bands which are also non-directional.

The system affords means for the simultaneous radiation of weather broadcasts and directional guidance signals. To avoid interference between the 1,020-cycle beacon signals and the speech frequencies, a band rejection filter is inserted in the input circuit to the speech modulation for eliminating the speech frequencies in the neighborhood of 1,020

cycles. A combination band-pass band-rejection filter is used in the output circuit of the aircraft receiver so that one circuit carries only the range-beacon (1,020-cycle) signals and the other circuit carries the speech signals. By means of a switch the pilot may select one or the other signal, or he may obtain reasonably satisfactory reception of both.

The four corner towers replace the two-loop antenna systems, crossed at right angles, which are used at older-type radio range-beacon stations. They constitute two directional antenna systems, each formed by two opposite towers on diagonal corners of the square. These are fed in opposite phase so that they correspond to the vertical conductors of the older loop antennas and give the same figure-of-eight radiation characteristics in the horizontal plane (see Fig. 8b). In this way radiation is confined to the vertical antennas, and the transmission of horizontally polarized electric-field components in the sky wave, such as from the horizontal wires of the loop antennas, are avoided. With the loop antennas these transverse horizontal components upon reflection from the ionized layers produce serious and erratic errors in the indicated beacon courses, often called *night errors* because they occur only at night in the frequency range used.

The principles of operation of the radio range beacon whereby radio-marked courses are set up are evident from Fig. 8b. The intensities of the side-band emissions formed by the two directional antenna systems, and hence of the detected signals produced by beating the side-band emissions with the non-directional carrier, are equal along the lines OA, OB, OC, and OD which bisect the angles between the two antennas. An airplane may therefore follow a course along the bisectors referred to if means are provided for distinguishing the radiations from the two directional antennas. For this purpose an automatic keying relay, connected in the coupling circuit from the side-band channel of the transmitter to the directional antenna system, is used for keying the radio power to one of the directional antennas in accordance with the Morse characteristic N(—.) and to the second directional antenna in accordance with the Morse characteristic A(.—). The coded signals are sent out in groups and are interlocked so that along any one of the four courses they form a long dash, or continuous monotone signal, interrupted every 24 sec. by the station identification signal. The course signals are obtained along zones, 2 to 3 deg. wide. Off the course the monotone signals break up into the component N and A signals, one or the other being of greater intensity depending upon the side "off course." The pilot is thus enabled to return to the course if the airplane should drift to one side or the other for any reason.

The coupling system between the transmitter and the directional antennas incorporates the link-circuit relay, a goniometer, a course-shifting pad, artificial line sections, concentric transmission lines to the tower antennas, and antenna coupling and tuning equipment (see Fig. 9). The relay, of the polar type, is energized by an automatic motor-driven keying device (not shown) so as to key the r-f power to the primary windings P_1 and P_2 of the goniometer in accordance with the N-A sequence indicated in the foregoing.

7. Goniometer. The goniometer is used for convenience in orienting the beacon space pattern and consists of two primary and two secondary tuned windings. The primary windings are crossed at 90 deg., as are also the secondary windings, the two sets of windings being made concentric. One set of windings is fixed and the other set rotatable about the common axis. The angle between the primary and secondary windings may therefore be varied at will. Each primary winding, acting in conjunction with the two crossed secondary windings and the two crossed directional antennas, sets up a system which is electrically equivalent to a single directional antenna. The plane of this phantom antenna is dependent upon the relative coupling of the secondary coils to the primary coil under consideration. Since there are two primary windings, two such phantom antennas exist, the angle between their planes being equal to the angle between the primary windings. The two phantom antennas may therefore be rotated in space (thus changing the position of the equisignal zones or courses formed by their space patterns) by changing the relative position between the primary and secondary windings. Without the use of the goniometer it would be necessary mechanically to rotate the directional antenna system to secure the same result. In practice, the rotation of the beacon space pattern is convenient in the first adjustment of the beacon, the goniometer being locked in position after this adjustment. Actually, other conditions (to be discussed

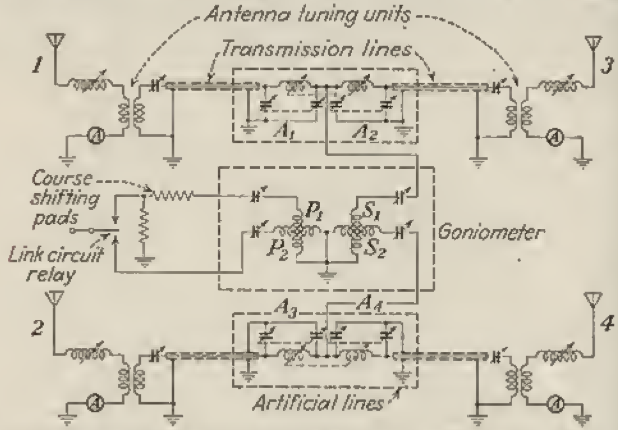


FIG. 9.—Schematic diagram of radio range-beacon antenna-coupling system.

dictate that the antenna orientation be chosen so that the goniometer may be left, preferably, at its 45-deg. setting.

8. Course Orientation to Coincide with Airways at Arbitrary Angles. The course-shifting pad and the artificial π line sections A_1, A_2, A_3, and A_4 are used for shifting the range-beacon courses from their 90-deg. relationship in order that they may be aligned with the airways. The course-shifting pad reduces the r-f power fed to goniometer primary winding P_1, thereby reducing the relative amplitude of the corresponding figure-of-eight radiation pattern. The resultant effect on the course orientation is shown in Fig. 10*a*. The artificial line sections allow modification of the 180-deg. phase relationship between the currents in the two towers forming each directional antenna (1 and 3 or 2 and 4) so that the space pattern corresponding to the N or A radiation may be made to depart from a figure of eight. This provides for a non-reciprocal relationship of the normally 180-deg. courses, as shown in Fig. 10*b*.

9. Course Stabilization. Special precautions are taken to ensure maintenance of the space patterns so that shifting of the courses will not exceed 1.5 deg., owing to the changes in phase or magnitude of the current

in one or more of the towers which may be produced by changes in tower capacity under varying weather conditions, etc.

It can be shown that the two towers of each directional antenna may be made to maintain nearly exact equality of current amplitude and 180-deg. phase if the condensers in series with the primary windings of the antenna coupling transformers are adjusted so that

$$X_1 = Z_0 \tan \theta$$

where X_1 = primary reactance (with the secondary antenna circuit open)
Z_0 = characteristic impedance of the transmission line
θ = electrical length of the line to each antenna (including the artificial line section).

This is equivalent to tuning the transmission line to resonance. The stabilization is affected somewhat if other than a 180-deg. phase relationship

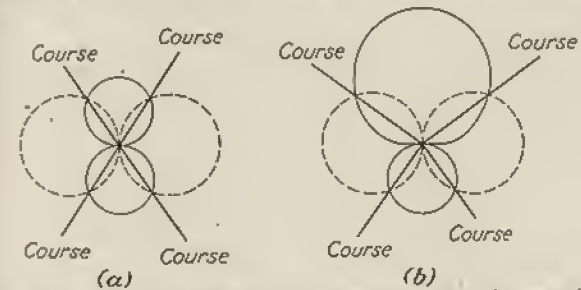


FIG. 10.—Alignment of range-beacon courses with the airways: (*a*) course squeezing; (*b*) course bending.

between the two towers is desired. Attenuation in the transmission lines and resistance in the primary transformer windings also affect the degree of stabilization.

With the goniometer set at zero degree (S_1 coupled only to P_1 and S_2 to P_2, in Fig. 9) this stabilizing arrangement would still allow the relative magnitude of one radiation pattern corresponding to one pair of towers to vary with respect to the second pattern corresponding to the second pair of towers. This is because of the possibility of an effective change in the load impedance offered to the transmitter and is overcome by setting the goniometer at 45 deg. so that both secondary windings of the goniometer are coupled equally to each primary winding.

10. Ultra-high-frequency Two-course Beacon with Visual Indication and Sector Identification. The space patterns of a radio range beacon which is undergoing service tests on the airways for u-h-f operation is shown in Fig. 11. This arrangement incorporates several features which present-day knowledge and advanced flying technique have indicated as being desirable.

Only two useful courses are provided (by the intersection of the full-line patterns), as compared to four courses in the case of the present l-f range beacons. This simplification materially reduces the orientation problems which the pilot is frequently called upon to solve under special conditions. For example, when near the station during strong winds, with the four-course beacon, the pilot may drift into an N or A quadrant and experience difficulty in determining which N or A quadrant he is in and in which direction to fly to get on the desired one of the four courses. The orientation problems with the two-course beacon are further simplified by the sector identification signals produced by the two dash-line patterns shown in Fig. 11. Thus, if the main courses are oriented east and west, the criterion for determining whether the east or west leg of the beacon is being followed consists in ascertaining whether the *PTE* or *PTW* identification signal is the stronger. (Here *PT* is the station identification signal and *E* or *W* the sector identification signal.)

The main patterns forming the two-beacon courses are distinguished from each other by modulating signals, 90 and 150 cps, instead of by the *N* and *A* coded signals in the present l-f range beacons. This allows the use of a visual course indicator in the output of the beacon receiver on the airplane and, since the two antenna systems are excited simultaneously, renders a-v-c reception feasible. The course indicator consists of an electrical filter and balanced-rectifier circuit for separating the two modulating frequencies, rectifying them, and applying the resultant rectified signals in phase opposition through a zero-center pointer-type microammeter. When the airplane is on-course, the two rectified signals

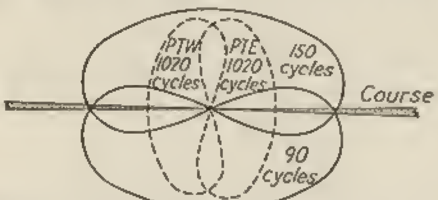


FIG. 11.—Visual two-course range beacon with aural sector identification.

are equal and the pointer remains at zero, the "on-course" position. When the airplane deviates, say, to the 150-cycle side of the course, the corresponding rectified signal becomes greater and the pointer deflects to the sector of the instrument marked 150 cycles. The pilot is thus afforded a continuous visual indication of the position of the airplane with respect to the desired course.

At predetermined intervals, approximately 30 sec., the identification signals are transmitted on a modulating frequency of 1,020 cps, which, by means of suitable filters in the aircraft receiver output, are directed to the pilot's headphones. Besides furnishing sector identification, these signals serve an additional purpose, in that they indicate to the pilot that the range beacon is operating. Such information is essential when using a zero-center–type course indicator, since an on-course indication may be obtained when the received signal intensity is zero.

The patterns shown in Fig. 11 are produced by five-element antenna arrays; the directivity thereby obtained being helpful in reducing bent and multiple courses (because of the reduced side radiation). Experiments have indicated that the use of horizontally polarized waves rather than vertically polarized waves also serves to reduce such effects, probably because reradiation occurs more frequently from vertical obstacles. Hence present plans call for use in the range-beacon antenna arrays of an antenna element which sets up only horizontally polarized waves. Such

an element is shown in Fig. 12 and is equivalent to a horizontal loop antenna.

The use of u.h.f. does not reduce the number of frequency channels required for a national network of range beacons, as might at first appear in comparison with the l-f network. The reason for this is that the ultra-modern air liner flying in the substratosphere may use every second, third, or even fourth range beacon and thus will be interfered with by intermediate stations unless they operate on different frequencies.

11. Low-frequency Marker Beacons. These stations, located at intervals along the airways, have served two different purposes. One is to mark the meeting points of adjacent radio range-beacon courses or to denote a particular locality along the airway, such as an intermediate landing field or an abrupt change in the elevation of the topography. For the former, transmitters capable of transmitting alternately on the two frequencies of the adjacent radio range beacons are employed, while for the latter only single-frequency transmitters are used. Each marker beacon station has a characteristic identifying signal. Its range is limited to 5 to 10 miles so that it may effectively localize the point desired.

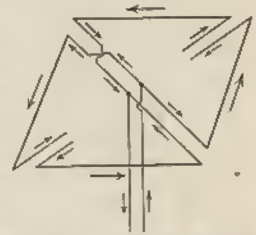


FIG. 12.—Ultra-high-frequency loop-antenna element for use in antenna arrays producing horizontally polarized waves.

The second purpose filled by marker beacons is one of directional guidance as well as marking of locality. Radio marker beacons of this type are miniature radio range beacons. These are located either at points along the airways so as to fill in gaps between the more powerful radio range beacons or at intermediate landing fields to enable pilots to locate the landing areas during adverse weather conditions. In these applications they are hardly to be distinguished from radio range beacons or runway localizing beacons, respectively, and their number varies as the facilities along the airways are modernized.

12. Ultra-high-frequency Cone Markers. There exists approximately directly over the l-f type radio range beacon a small zone of zero signal called the *cone of silence*. This zone arises from a combination of the directive properties of the beacon-transmitting antenna system and the receiving antenna on the airplane and has been used extensively by pilots for obtaining a definite "fix" over the station. Because of the essentially negative nature of the indication and because of the possibility of obtaining false indications caused by momentary equipment failures, radio transmission vagaries, etc., a cone-type marker beacon is now used at range-beacon stations to provide positive identification of the station location.

The cone marker consists of a 5-watt 75-Mc crystal-controlled transmitter, modulated 100 per cent at 3,000 cps and feeding a directive antenna array which produces a conical lobe of energy radiated upward. Its radiation pattern in the vertical plane is shown in Fig. 13. The horizontal-plane radiation pattern is non-directional, *i.e.*, circular.

The antenna system for obtaining the desired patterns is shown in Fig. 14*a*. It is installed one-fourth wave length above a coarse (3- by 3-in.) mesh screen which in turn is erected approximately one-half wave length above ground. The screen counterpoise provides an effective level reflector for the antenna, yet allows snow to fall through and vegetation to grow beneath.

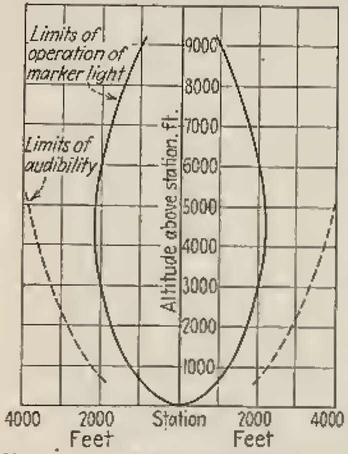


FIG. 13.—Vertical-plane radiation pattern of cone-type marker.

The antenna and counterpoise are located at sufficient distance from the central beacon tower to avoid pattern distortion, r-f power being fed from the transmitter (located in the station house) by means of a parallel-conductor transmission line. The latter consists of two ½-in. seamless copper pipes spaced 1 in. apart and supported centrally in a 3- by 3-in. copper shield. A short matching stub, connected at an appropriate point at the antenna end of the transmission line (not shown), eliminates standing waves along the line.

The vertical radiation pattern is obtained by the quarter-wave separation between the antenna and its counterpoise, coupled with the directive properties of a half-wave horizontal receiving antenna running fore and aft along the belly of the airplane. The non-directional horizontal radiation pattern is obtained by feeding the east-west half-wave radiating elements of the an-

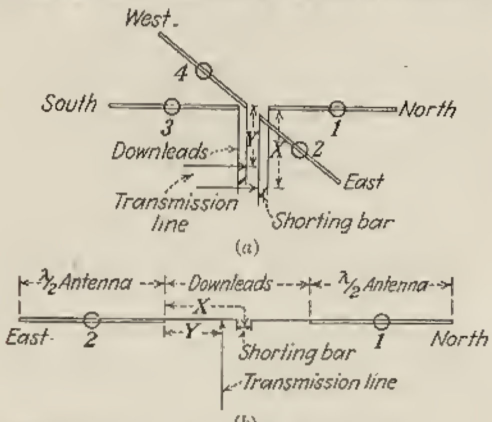


FIG. 14.—Details of antenna system for cone-type marker: (a) method of feeding; (b) detail of 10-deg. phasing.

tenna system 90 deg. out of phase with the north-south radiating elements (see Fig. 14a). This is accomplished by making the dimen-

sion X of the transmission lines connecting antennas north and east together and south and west together exactly one-fourth wave length and adjusting dimension Y so that the currents in the north and east antennas are in phase quadrature and likewise the currents in the south and west antennas. Figure 14b, which shows the north and east antennas in a straight line with their respective down-leads, will clarify the principles involved.

The transmitter is constructed in duplicate and includes, in addition, a monitor unit. The latter functions to disconnect the regular channel in event of its failure to deliver a predetermined power output and starts up the stand-by channel, connecting it to the transmission line in place of the regular channel.

13. Ultra-high-frequency Fan Marker. The procedure for aviation traffic control, evolved as the result of many years' experience, calls for the division of authority over traffic along the airways and in the vicinity

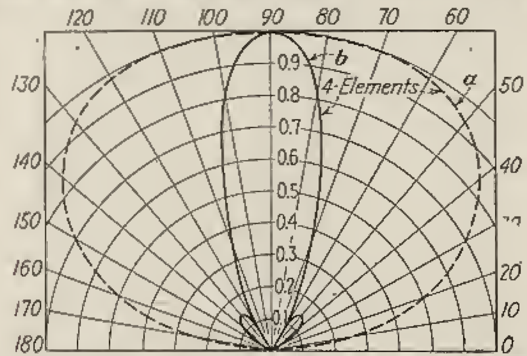


FIG. 15.—Vertical-plane radiation patterns of fan-type marker: (a) in plane transverse to range-beacon course; (b) in plane parallel to range-beacon course.

of and at airports. The pilot follows a flight plan approved by the airways traffic control officer until he reaches a control point marking the beginning of the airport control zone. Up to this point, the airways control officer has kept track of the pilot's position throughout the flight (via the periodic contacts between the airplane and the ground stations of the air transport companies, CAA facilities, etc.) and may alter the flight plan to conform with traffic requirements or meteorological conditions.

During good visibility conditions, the pilot reaching this control point communicates with the airport control officer for landing instructions and automatically passes out of the control jurisdiction of the airways control officer.

During adverse visibility conditions, the latter retains control authority until the airplane comes into visual contact with the airport control tower. Depending on traffic conditions, he may order the pilot to circle at a specified altitude above the control point until other airplanes effect their landings.

In either case there is seen to be a need for a radio aid which may define the control points, generally about 25 miles from the airport on each of the airways. The fan-type marker beacon serves this purpose.

The fan marker consists of a 100-watt 75-Mc crystal-controlled transmitter feeding a directive antenna array. The setup produces a sheet of energy radiated upward and transverse to the range-beacon course. The transmitter is in duplicate, as in the case of the cone marker, with provision for putting the stand-by unit in operation in event of a predetermined change in performance of the regular unit. A modulation frequency of 3,000 cps is used with distinctive keying to serve for identification.

Useful radiation extends up to about 20,000 ft. At 7,000 ft. the radiation is about 18 miles wide and 4 miles thick. Although it appears that sufficient precision of position indication is afforded by considering the entering point of the radiated field, some consideration has been given to increasing the definition by setting up crossed patterns.

The antenna system used for obtaining the pattern shown in Fig. 15 consists of four horizontal half-wave antennas located in line along the range-beacon course and fed in phase by a transmission line from the transmitter house. The antennas are placed one-fourth wave length above a coarse-mesh screen counterpoise which, in turn, is approximately one-half wave length above ground. The details of the transmission line and of the counterpoise are substantially the same as for the cone-type marker.

14. Ground-station Equipment for Two-way Communication Systems. To date, two-way communication between ground and airplane as carried out by the domestic air transport companies has been chiefly in the h-f band, 2,000 to 6,600 kc. (Experimental use of u.h.f. is in progress.) Communication is by voice because of the greater speed of operation; the use of radiotelegraphy is largely confined to companies operating outside the United States. The radio equipment used at the fixed terminal of a typical two-way radiotelephone system has reached a remarkable degree of refinement to meet the particularly exacting requirements encountered in this service.

The transmitter must be capable of operation on any one of a group of frequencies, with facilities for rapid change-over to any other frequency in the group. This is necessary since each transport route has a day and night frequency for communication with aircraft and also separate frequencies for point-to-point communication. Moreover, when the ground station is located at the junction of several routes operated by the same company, provision must be made for communication on either the day or the night frequency corresponding to each route. Each frequency channel is crystal controlled, the frequency being held constant to within 0.025 per cent. Approximately 400 watts of r-f power on each frequency is required in the antenna to effect reliable communication over the desired distance range; some ground-station transmitters use up to 3 kw.

The receiving equipment must be highly selective because of the many channels that must be accommodated within the comparatively narrow band of frequencies allocated to this service, stations at the same airport operated by different transport lines being frequently less than 1 per cent apart. Extremely high sensitivity coupled with excellent a.v.c. is required to provide substantially constant output under the varying transmission characteristics which usually obtain in this frequency range and because of the varying distance between the aircraft and

ground. Provisions for remote operation must be made since, in order to secure freedom from man-made interference, the receiving equipment is frequently located as much as 30 miles distant from the operating staff.

15. Ground-station Transmitter. Typical of the advanced type of transmitting equipment required for this service is the Western Electric Type 14 transmitter. This transmitter provides crystal-controlled telephone, continuous wave, or tone telegraph transmission on any one of 10 frequency channels within the range of 2 to 18.1 Mc. It employs a crystal oscillator; two intermediate buffer-amplifier stages which function either as amplifiers or doublers, depending upon the frequency used; a modulating amplifier preceded by two audio stages; and a power amplifier. Each frequency has its own quartz plate and a set of interstage and output coils. The set is so arranged that any one of the frequencies desired may be selected, by single-digit operation of a telephone dial, within about 1 sec. The transmitter can be operated with push-button, telegraph-key, or voice-operated carrier control. The carrier is suppressed automatically during unwanted periods. Provision is made for remote frequency selection and starting and stopping of the carrier. A set of three simple vertical antennas approximately 15, 30, and 60 ft. high may be used to cover the entire range, or any combination of directional and non-directional antennas. The transmitter is a-c operated, requiring approximately a 4-kva 220-volt three-phase supply. In a setup of this type of equipment by the Eastern Air Lines, Inc., on the New York to Atlanta route, these transmitters provide telephone ground-to-airplane communication, telephone point to point, continuous-wave telegraph point to point, and RCA facsimile point to point for the transmission of long weather sequences and long routine company business. On the telephone point to point, provisions are made for the use of Western Electric speech inverters so that the conversations may be of a private nature.

One problem in the design of ground-station transmitters lies in the suppression of overmodulation products which produce adjacent-channel interference. An automatic constant-level speech input amplifier which operates to prevent momentary high a-f peaks from reaching the transmitter has been developed for this purpose. Its use in conjunction with crystal-resonator preselectors in the receiver input allows successful reception only 1,000 ft. from a 400-watt transmitter when tuned to a carrier only 20 kc off the transmitter frequency.

16. Remote-control Receiver. Use of remotely located receiving installations controlled from the radio offices at the airport has become widespread except for emergency operation. The receiver is generally a multifrequency superheterodyne with quartz-plate control of the heterodyne oscillator. Control of the remote receiver is accomplished by dial impulse-generating equipment and a telephone wire line. Relay circuits at the receiver, which respond to the transmitted impulses, provide for frequency change, volume-control settings, etc. The receiver output is fed back over the line and is generally amplified at the operator's end for loud-speaker service.

One type of remote-controlled receiver incorporates a Codan "carrier-operated device, antinoise" which keeps the receiver silent while in stand-by position and feeds the loud-speaker only when a modulated carrier is being received on its preselected frequency. This device operates reliably under high noise conditions and does not require adjustments to compensate for variations in the noise level. This receiver is arranged for mounting in weatherproof cabinets which may be fastened to the same telephone pole which supports the receiving antenna and is provided with an emergency-battery power-supply system which is automatically connected to the receiver in the event of failure of the normal a-c power supply. The set has a sensitivity of 1 μv to give 50-mw output. The a.v.c. will hold the output level constant within 4 db for a variation in input voltage of 100 db. Cutoff of audio frequencies below 200 and above 3,000 cps is provided to reduce noise. Its selectivity is such that an interfering modulated carrier 10 kc away (30 per cent modulation at 400 cps) must be 20 db above the desired carrier level to produce an interfering voltage 20 db below the desired signal voltage.

17. Radio Facilities at a Modern Airport. A brief tabulation of all the various radio aids and facilities available at a modern airport will serve to emphasize the important part radio has in aviation. The facilities at La Guardia Field, New York City, are taken for example.

1. *CAA Radio Room.* Teletype facilities for collecting weather information. Remote control equipment for operating radio-range beacon and telephone transmitters and for making weather broadcasts. Receivers for maintaining watch on national calling frequencies.
2. *Air Transport Company Communications Systems.* Transmitters and remote-controlled receivers operated by each of five air-line companies for two-way communication with aircraft and for point-to-point communication. Direct communication facilities with CAA radio room and with airways and airport control offices. Private teletype system.
3. *Marker Beacons.* Cone-type marker beacon at radio range-beacon station. Fan-type marker beacons at control points on each of the four incoming airways.
4. *Radio Landing Aids.* Four-way instrument landing facilities providing for landing in either direction along the airport's two longest runways. Each direction utilizes an u-h-f runway localizing beacon, a landing beam, and two fan-type low-approach marker beacons. (A complete description of radio landing aids is given in Art. 30.)
5. *Airways Control Office.* Teletype facilities for collecting information on airways traffic conditions. Direct communication with CAA radio room, air transport company dispatchers, and airport control tower for collecting information on airway's traffic conditions and for directing traffic (through CAA, company's and airport radio-facilities).
6. *Airport Control Tower.* Remote-controlled airport radiotelephone transmitter (with complete emergency stand-by unit) for directing airport traffic. Remote and direct-controlled receivers for standing watch on national calling frequencies, air transport companies' two-way communication frequencies, U. S. Army frequencies, and others. At the time of writing, 15 remote-controlled receivers are used, 13 for maintaining a watch on the following frequencies and 2 in reserve. Three multifrequency direct-controlled receivers provide stand-by emergency facilities.

Frequency, Kilocycles	Service
3,105	Itinerant aircraft
4,495	U. S. Army
3,232.5 3,257.5 5,612.5 5,632.5 5,652.5	American Airlines
4,422.5	Eastern Air Lines
2,870	Pan American Airways
3,088 4,937.5	Transcontinental & Western Air
5,572.5 3,162.5	United Air Lines

AIRPLANE RADIO INSTALLATION

18. Special Requirements and Installation Practice. Because of the special nature of the installation of radio equipment aboard aircraft, special mechanical, electrical, and aerodynamic requirements are imposed. Reliability and simplicity of operation are essential. The equipment must be constructed to withstand continued vibration and landing shock without change in performance and must operate under all conditions of weather encountered in flight. Space and weight must be kept down to a minimum. The equipment must be capable of quick removal from the airplane for servicing or replacement. Simple but complete remote control of the equipment, including frequency change-over, etc., is essential. An adequate, efficient power supply is required. The antennas must be of sound aerodynamic design. Special precautions are needed in eliminating various electrical disturbances arising on the airplane.

19. Airplane Antennas. An aircraft antenna must have a good effective height, must be of sound aerodynamic design, and must be convenient to use under varying air-transport operation conditions. The trailing wire fulfills the first requirement but fails to meet the second and third requirements. It is still used in modified form in some modern installations for transmission, because of its efficiency and comparatively greater freedom from ice formation. A typical fixed transmitting antenna consists of a mast approximately 6-ft. high, mounted above the fuselage and with flat-top wires extending toward the wing tips and the vertical rudder post. In larger ships this form may be modified to produce a front to rear, wing tip to wing tip, or V antenna. The mast may constitute the lead-in, in which case it is insulated from the fuselage. Lengths of antennas range from 30 to 75 ft. depending upon the antenna form and the size of the airplane.

A whip antenna extending 4 to 6 ft. vertically above the fuselage has an effective height of about 1 meter, sufficient for use with sensitive receivers. Errors in course indication on the radio range beacon are introduced unless the receiving antenna on the airplane is entirely nondirectional. This restriction limits the antenna configuration to either the vertical-pole antennas or to a vertical antenna with flat-top loading, the flat-top elements of which are so arranged that their horizontal effects neutralize each other. The symmetrical, longitudinal, or transverse T antennas with vertical lead-in are examples of the latter type. The V antenna, mounted well forward below the fuselage with its apex leading and the lead-in connected to this point, is another antenna of this type; control of the angle of the V provides for a symmetrical antenna and counterpoise system. Considerable use is made in practice of a single wire inclined backward and upward toward the vertical rudder post. With this arrangement the directional errors are utilized to compensate for the tendency of a pilot to weave about the beacon course.

Antennas for transmitting and receiving at u.h.f. may consist of half-wave dipoles, generally horizontal, with conventional transmission-line coupling. The beacon receiving antenna may be of the form shown in Fig. 12 mounted well above the fuselage and forward so that it will retain its free-space characteristics.

20. Aircraft Power Equipment. Five determining factors enter into the choice of the power system to be adopted: (1) reliability, (2) weight,

(3) availability when the main power plant of the airplane is crippled, (4) electrical performance, and (5) maintenance required during service. Several distinct types of power-supply systems are available. The receiving-set power requirements are satisfactorily provided by the combination of the 12-volt battery and dynamotor plate supply. The transmitting-set plate-supply requirements, being considerably larger, have led to the development of a number of different arrangements. These include dynamotors driven from the aircraft storage battery, airplane-engine-driven generators, wind-driven generators (now practically obsolete), and auxiliary gasoline-engine-driven generators. In comparing these systems, consideration must be given to the ever-increasing electrical load requirements on a modern transport airplane other than radio power supply. These include lighting, motor starters, motors for operating adjustable pitch propellers, retractable landing gear, flaps, fuel and oil pumps, remote-controlled switches and solenoids, etc.

The most widely applied system utilizes a 14-volt d-c charging generator, driven from an airplane engine and provided with a voltage regulator so as to maintain substantially constant generator voltage for all possible airplane-engine speeds. The generator charges the 12-volt airplane storage battery, which in turn drives the necessary dynamotors for obtaining receiver and transmitter plate power supply. Some of the larger transport airplanes use a dual battery and generator system, each battery being of nominal 65 amp.-hr. capacity and each generator having a 50-amp. rating. In one arrangement, provision is made whereby the two systems normally operate independently each carrying half the load; when desired, the full load may be applied to either battery and both generators may operate in parallel for charging it. In a second arrangement, one system is kept as a stand-by so that a fully charged battery will be available for emergency operation in event of failure of the airplane engines. A 24-volt battery and charging-generator system has also been adopted for large transport airplanes.

An alternate system suitable for airplanes having very high electrical load requirements, employs an auxiliary gasoline engine driving an a-c generator. The generator may be 115 volts, 800 cycles, single phase, or it may be 115 volts, 400 cycles, three phase; the three-phase system is somewhat more suitable when the electrical load is largely a motor load as in military aircraft. The high generated frequencies permit the use of very lightweight transformers and filter units in obtaining high d-c voltages for radio plate power supply. A complete complement of radio transmitting and receiving equipment for use with either dynamotors or single-phase 800-cycle supply has been designed by one radio manufacturer. In one a-c installation, on a DC-4 airplane, two 800-cycle auxiliary engine-driven alternators were used, mounted in the nacelles of two of the airplane's four engines. The rated capacity of each alternator was 7.5 kva.

21. Radio Shielding and Bonding in Aircraft. Intense electrical disturbances are set up in the radio receiving circuits by the electrical ignition system of the airplane engine, unless ignition shielding is provided. To obtain effective shielding, it becomes necessary to enclose the entire electrical system of the engine ignition in a high-conductivity metallic shield. This requires the provision of suitable metallic covers for the magneto distributing heads, for the booster magneto, for the ignition distributing wires running from the magnetos to the spark plugs, for the spark plugs themselves, for the ignition switch, and for the switch and booster magneto leads.

With the low-power ground stations in present use, it is necessary to utilize field intensities not appreciably greater than the prevailing static level. Hence very careful ignition shielding is essential. The use of u-h-f communication imposes additional requirements on the shielding efficiency.

Electrical disturbances may also be set up by any of the numerous electrical devices used in the modern airplane and by the periodic discharge of static voltages accumulated on isolated metallic parts on the airplane. The former is eliminated by direct shielding of the devices, coupled with filtering of connecting leads. The latter is eliminated by bonding all metallic cases, parts, and controls to the common ground formed by the airplane fuselage.

22. Precipitation Static. Oscillating corona discharge from points on the airplane to the surrounding atmosphere, occurring when the airplane flies through electrically turbulent air masses, leads to what is known as *rain*, *snow*, or *precipitation* static. Such static is often of sufficient intensity at the lower frequencies to mar reception even at short distances. The use of a shielded loop antenna (familiar in marine-radio direction finding) for receiving reduces this type of static in a material degree, probably because of the preponderance of electric field components in the h-f radiation of the static near its source of origin.

A second and more effective expedient for reducing precipitation static is to provide a discharge point well removed from the airplane antennas and to control the discharge by means of a resistor so that oscillating corona is minimized. This has been done by attaching a resistance cord (approximately 0.5 megohm) and wire to the tail of the airplane. The wire is 0.016 in. in diameter, the sharp point at its end forming a much more effective discharge point than the projections on the airplane from which corona discharge normally occurs.

23. Aircraft Radio Transmitters. Practically all transport airplanes employ multifrequency, crystal-controlled transmitters, generally 50 watts or higher with 100 per cent modulation. Simplex operation is used whereby transmission and reception is on the same frequency. The communication receiver is normally in stand-by position. Depressing the "press-to-talk" microphone button disconnects the receiver from the communication antenna, connects the transmitter in its place, and starts the transmitter dynamotor. Side tone is automatically provided.

The transmitter incorporates several pretuned crystal-oscillator and power-amplifier circuits with remote control means for connecting each set at will to a common set of r-f tubes and an associated modulator system. The remote control may be by means of a flexible shaft operating a ganged switch or it may be by electrical telephone-dial selection. Transmitters having up to 10 operating frequencies are in use. Plate modulation is general with a speech-frequency range of 300 to 3,500 cycles, frequencies below 300 cycles being suppressed mainly to reduce high audio interference levels set up by the propellers and engine exhausts. Dynamotor voltages for plate power supply are generally 1,250 volts or less.

Transmitters for private aircraft are set up for one, two, or three frequencies; 3,105, 3,120, and 6,210 kc; 3,105 and 6,210 kc; or 3,105 kc. For the smaller airplanes the transmitter operates on 3,105 kc and, with a companion range-beacon receiver, uses a dry-battery power supply system. The equipment is generally located to allow direct control.

With the advent of u-h-f operation, some simplification in channel requirements may be effected because of the elimination of change-over from day to night frequencies and vice versa.

24. Aircraft Radio Receivers. With the present setup of radio frequencies (see Table I), the full complement of receiving equipment on a transport liner includes the following: a l-f range-beacon receiver; a h-f communication receiver; an auxiliary receiver, usually of the all-wave type, for radio-direction finding and (primarily) for emergency use; and a 75-Mc marker-beacon receiver. These are gradually being supplemented with u-h-f equipment for use of the u-h-f airways facilities, blind-landing systems, etc. Itinerant aircraft generally carry only a range-beacon receiver which provides for weather-broadcast and range-beacon reception, for messages from airport traffic-control transmitters, and emergency messages from CAA l-f radiotelephone facilities.

All air-line receivers used at the present time are of the superheterodyne type. The range-beacon receiver has a sensitivity of about 3 μv and the communication receiver of about 1 μv for a signal/noise ratio of 2:1. The marker-beacon and landing-beam receivers are of fixed-sensitivity types with special provision for maintaining constant sensitivity under varying operational conditions. The u-h-f range-beacon and communication receivers have a sensitivity of approximately 5 μv.

The l-f range-beacon receiver is of the continuously variable type, controlled with a rotating flexible shaft which may be up to 30 ft. in length. Manual remote volume control is also provided. The frequency control is generally through a control crank, operated by the pilot, which is geared 4 to 1 to the flexible shaft; the latter has a ratio of 264:1 to the tuning condenser. Gear backlash effect is thus minimized. Provision is made for quick shift-over, generally by means of a relay and pretuned circuits, to 278 kc for use in traffic control communication. The receiver is arranged for dual output (to pilot and copilot) and, sometimes, for dual control.

The h-f communication receiver is of the multifrequency type with crystal control of the heterodyne oscillator in each of the channels. Both mechanical and electrical remote-control switching to preset frequencies are employed, as in the case of the companion multifrequency communication transmitters. In some cases the same control switches both the receiver and the transmitter simultaneously. Automatic-volume-control reception is used. Dual control and dual output for pilot and copilot are provided; the controls include provision for separate output level adjustment.

The auxiliary receiver is generally of the continuously variable tuning type but may employ a crystal "lock-in feature" at specific company communication frequencies. Normally the copilot uses this receiver for obtaining position checks when crossing other range-beacon courses or, in conjunction with a loop antenna, for obtaining directional bearings on range-beacon or entertainment broadcast frequencies. In emergencies this receiver may be switched by a normally sealed control to an emergency dry-battery power supply good for 4 hr. of continuous operation.

The marker-beacon receiver is a single-channel crystal-controlled superheterodyne type and is adjusted and controlled at 1,400-μv sensitivity. Three filters in the output are tuned, respectively, to 3,000, 1,300, and 400 cycles and feed separate, distinctively colored, lights on the airplane instrument panel. The first frequency corresponds to the modulating frequency used at cone and fan-type marker beacons; the other two correspond to the outer and inner approach markers of radio landing aids installations.

COURSE NAVIGATION AND POSITION DETERMINATION

25. Guiding Systems. Radio systems for guiding aircraft comprise two types: (1) aids for aircraft flying the established airways and (2) aids for aircraft flying over independent routes. The first is the more important in the United States. All commercial transport airplanes use fixed airways. The government aids to air navigation are being provided with the primary view of serving aircraft flying these airways.

An ideal system suitable for use by aircraft flying either fixed airways or independent routes, on land or on sea, is such that

1. The system shall give the pilot information to enable him to continue along a given route between any two points in a given service area when no landmarks or sky are visible. If he leaves the course, it should tell him how far off he is and to which side, should show him the way back to the course, and should inform him when he arrives at his destination.
2. The necessary directional service shall be available at all times and under all conditions, to all airplanes equipped to receive the service and flying within the area service.
3. The service shall be easily, positively, and quickly available to the pilot, with a minimum of effort on his part.
4. The radio equipment required on the airplane shall be simple, rugged, of light weight, and relatively inexpensive.
5. The ground equipment shall be as simple as possible. The radio frequencies, power, type of emission, and location of ground transmitting stations shall be such as to serve the needs with maximum efficiency and conservation of the limited ratio channels available.

26. Direction Finder on Airplane. One system employs a fixed-coil antenna, the plane of which is perpendicular to the longitudinal axis of the airplane. Zero signal is obtained in the receiving-set output as long as the airplane is pointing to the ground transmitting station. This is essentially a "homing" system and is subject to the limitation that a circuitous path is followed if heavy cross winds prevail. This is illustrated in Fig. 16 and applies only when compensating course corrections based on the indications of the magnetic compass are not periodically applied. The use of the zero-signal bearings is rendered much more flexible through the adoption of a rotating loop antenna on the airplane, but the system still has some defects. It lacks means for giving the pilot the sense of deviation from the course, the signal increasing from zero whether the airplane deviates to the left or to the right. Moreover, the use of a zero-signal indication is difficult under conditions of severe atmospheric disturbances or interference from other services.

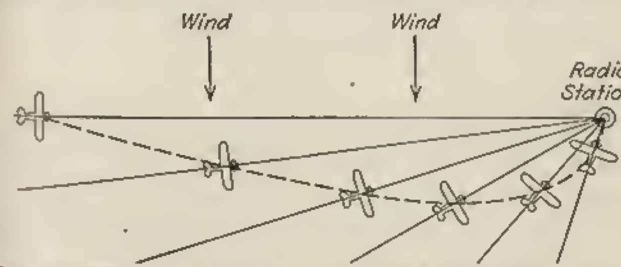


FIG. 16.—Effect of cross winds on path followed with direction finder.

To obviate these difficulties, the *Robinson direction-finding system* was developed. In this system, two crossed-coil antennes are used, one coil having its plane along the longitudinal axis of the airplane and the second having the plane perpendicular to this axis. The signal due to the second or auxiliary loop antenna is alternately added to and subtracted from the signal due to the first- or main-loop antenna. When on the course, since no voltage is then induced in the auxiliary coil, the two signals are of equal intensities. When off course to the left, for a given phase relationship between the two loop antennas, the sum of the two signals is greater than the difference; when off course to the right, their sum is less than their difference. Sharpness of course indication is directly dependent upon the ratio of effective height of the auxiliary loop antenna to that of the main antenna. A suitable automatic switching sequence enables the pilot to determine the orientation of his airplane with respect to the true course or bearing. The system as developed was of the fixed loop-antenna type.

In modern aircraft radio practice the Robinson direction finder has been replaced by equipment giving visual indication of the airplane heading relative to the course directed on the ground transmitting station. A number of commercial units have been introduced and are successfully employed for flying along independent routes and as adjuncts to the radio range-beacon system. These are generally modifications of the Robinson direction finder in which the main coil is replaced by a vertical antenna, since no directivity is required of this element and the auxiliary loop is made rotatable. Switching to the additive and subtractive positions is accomplished electrically and is performed at a rapid rate. The output signal of the receiving set is switched synchronously with the antenna system, so that it passes alternately in opposite directions through an indicating instrument of the zero-center type, thereby giving right and left indication of the heading of the airplane with respect to the desired course.

A circuit diagram of the earliest published arrangement[1] of this type and one which is similar in most of the essential details to many of the current commercial units is shown in Fig. 17. In this arrangement the tubes V_1 and V_2 are biased to cutoff by the bias battery C, passing current only when successive half-cycles of the switching frequency alternately make the grids less negative. The r-f voltage passed on from the coil L_1 (in the common plate circuit of V_1 and V_2) to the coil L_2 (connected to the input of a conventional receiving set) is thus alternately reversed. Voltage from a vertical antenna is also fed into L_2 in proper phase relation so that the loop-antenna voltage alternately adds to and subtracts from it. The amplified sum and difference voltage is detected and amplified and then passed through the current coil of an a-c electrodynamometer-type instrument. The field coil is excited by the switching frequency so that the zero-center pointer is deflected to the right, say, corresponding to the additive condition of the loop and vertical antenna voltages and to the left corresponding to the subtractive condition. The polar diagram indicating the response of the antenna system for the two conditions corresponding to varying directions of the airplane with respect to the transmitting station is shown in Fig. 18. The intersection of the two cardioid patterns corresponds to the zero-center or "course" position of the indicator. Whether the airplane is flying toward or away from the ground station is readily determined by noting whether the pointer deflects to the right or left or vice versa as the heading of the airplane is altered to the right or left of the course.

This type of direction finder is quite simple and may be used on any type of ground station, such as in the broadcast band. The same set may be used for the reception of weather-broadcast and range-beacon

[1] See reference to Dieckmann at end of section. For a description of commercial equipment, see *Electronics*, October, 1935.

signals. One important desirable improvement now receiving attention is the elimination of serious and erratic errors in the bearing obtained at night (also in daytime on the higher broadcast frequencies). A second project on which experimental work is in progress is the connection of

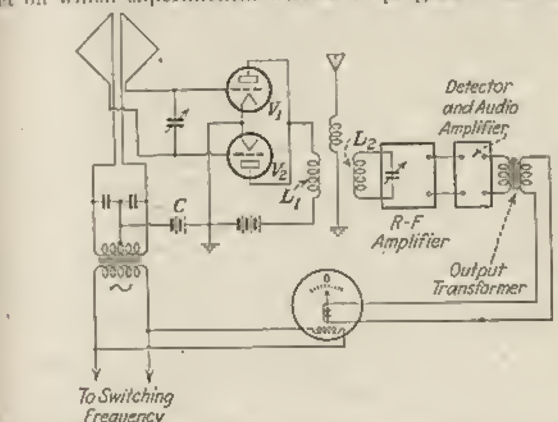


FIG. 17.—Schematic circuit diagram for visual-type airplane radio direction finder.

the course indicator to control the steering of the airplane, through use of the automatic pilot.

To make full use of the possibilities of a direction finder aboard aircraft, automatic indication of the direction of the tuned-in station is required. This has been accomplished in a number of commercial units through the use of a bidirectional motor system which drives the rotating loop antenna. The motor system replaces the bilateral pointer-type indicator and is arranged to drive the loop antenna clockwise corresponding to one cardioid pattern and counterclockwise corresponding to the second cardioid pattern. A little study will show that the system will be in stable equilibrium for only one course, which may be arranged to correspond to the forward direction. A bearing indicator, near the pilot, is attached to the loop-antenna driving shaft by means of a flexible drive and indicates the direction of the station correctly at all times. Automatic-volume-control reception is employed to render the direction-finder operation fully automatic.

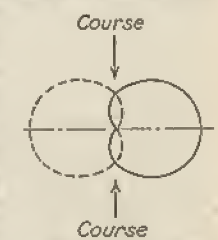


FIG. 18.—Crossed cardioid patterns for visual-type finder.

27. Direction Finder on the Ground. One system of navigational aids to aircraft is a direction-finding system, but with the direction finder located on the ground. Every airplane utilizing this system carries a radiotelephone (or radiotelegraph) transmitter and receiver. Permanent direction-finding stations are located at ground stations at strategic points. When an airplane desires to learn its position, it transmits a request on the airplane transmitting set, whereupon two or more of the ground direction-finding stations each determine the direction by observations upon the radio waves transmitted from the airplane. Triangulation then gives the position of the airplane, which information is transmitted to the airplane.

Five minutes is normally required between the time the request for a bearing is transmitted from an airplane and the time the bearing, as computed by two ground stations, is furnished the airplane. Obviously the system is best suited to long-distance operation over routes not too heavily congested, such as the transoceanic services.

A simple loop antenna may be used in conjunction with the receiving set required with this system, thereby giving the pilot additional directional or "homing" service to supplement the bearings furnished by the ground-station network. Even with this additional service, however, the airplane is not kept strictly on a given course at all times and is therefore not practicable where airplanes must fly over rigid airway routes. It has definite value, however, as an adjunct to the range-beacon system and is being used experimentally in the United States on the 2,900- to 6,600-kc band and at ultra-high frequencies. Completely automatic ground-station direction finders have been devised for this experimental service.

28. Rotating Radio-beacon System. A method of furnishing navigational aid to a flyer is the rotating radio beacon developed in England. This method employs a transmitter located at an airport, which has a loop antenna rotating at a constant speed of 1 r.p.m. A figure-of-eight pattern is thus rotated in space at a constant rate. A special signal indicates when the figure-of-eight minimum passes through north and also when it passes through east. A pilot listening to the beacon signal in the output of his receiving set can start a stop watch when the north signal is received and stop it when the figure-of-eight minimum reaches him. The number of seconds multiplied by six gives him his true direction in degrees from north. The stop watch may be calibrated directly in degrees, so that the position of the second hand, when the minimum signal is received, gives the bearing directly. The east signal is provided to overcome the difficulty in receiving the north signal when the airplane is north or south of the beacon, as on that bearing the signal strength is a minimum.

The receiving antenna is of a non-directional type. The receiving set may be used in the reception of weather-broadcast messages and other communications when not employed in direction determination. The system is capable of giving simultaneous service to any number of airplanes in any direction. Drift may be checked by determining positions, periodically, and correction may be employed. In the form described, the system has several inherent disadvantages. The service is intermittent and somewhat slow, requiring at least 30 sec. for each bearing. Since the determination of a minimum signal must be made, the system is particularly subject to interference and atmospheric disturbances.

Omnidirectional beacons have been developed in the United States which are based on the general principles of the foregoing system but which overcome its disadvantages. One installation, operating at 125 Mc, employs a five-element antenna system, either vertically or horizontally polarized, with four of the elements on the corners of a square and the fifth at the geometric center. A 125-Mc carrier, modulated 75 per

cent by a 10-kc modulation which in turn carries 60-cycle modulation, is applied to the central element; the radiation from this element is non-directional. Carrier power of the same frequency is fed to a small goniometer, which in turn feeds the four corner elements forming the directional antenna system. This will be seen to be quite similar to the simultaneous range-beacon and weather-broadcast system described in Art. 6, except that (since the goniometer primary windings are not keyed alternately) a single figure-of-eight radiation pattern is produced by the four corner antenna elements.

This figure-of-eight pattern is rotated in space 60 times per second by mechanical rotation of the small goniometer at this rate, and thus produces (at any point in space) a 60-cycle modulation of the non-directional carrier. The power fed to the four corner elements is adjusted so that the percentage 60-cycle modulation is 20 per cent. The phase of this modulation with respect to the 60-cycle modulation carried by the 10-kc subcarrier on the central radiation will obviously depend upon the direction from the transmitter and will go through 360 deg. as an airplane circles the transmitter.

Hence, in order to determine the bearing of the airplane with respect to the transmitter at any point in space, it is necessary to measure the phase difference between the variable- and fixed-phase 60-cycle modulation. This is done in one arrangement by separating them in the output of the receiving set (by 60-cycle and 10-kc filters) and applying them respectively to the field and armature windings of an electrodynamometer-type instrument such as is shown in Fig. 17. The pointer indicator of this instrument will read zero when the phase difference is 90 deg. A phase shifter connected between one filter unit and the corresponding instrument winding is used to adjust the actual phase to this value, its setting being then a direct measure of the station bearing. The indicating instrument then becomes a right-left indicator which enables the pilot to follow this fixed bearing to the station regardless of wind drift.

29. Radio Range-beacon System. Radio range beacons with aural and visual course indication have already been described in previous articles. Visual indication is more in keeping with airplane instrument practice, provides sharper course indication, eliminates the personal element in ascertaining the position of the airplane with respect to the course, and allows the use of a-v-c reception. The high level of atmospherics prevalent in the l-f range-beacon band led to the development of special mechanical-reed filter units to make visual course indicators practicable. The change-over to u-h-f operation renders the use of less selective electrical filters feasible. Several arrangements for obtaining visual course indication on aural-beacon signals have been tried but proved impractical.

The radio range beacon requires only a simple receiver aboard the airplane for its reception. The same receiving equipment is useful for other purposes, for example, receiving voice weather broadcasts, voice communication, etc. The system is simple to use by the pilot and permits him to fly along the established airways where all other aids to aviation are provided. Errors in course indication such as "night effects," bent and multiple courses, etc., which have been associated with the range beacons are present in all other systems but are, in general, not as easily recognized nor eliminated. The range-beacon system lends itself most readily to airways traffic control.

RADIO AIDS TO BLIND LANDING OF AIRCRAFT

30. Functions of an Instrument Landing System. The primary function of an instrument landing system is to provide continuous, positive three-dimensional guidance, whereby an airplane flying the airways under adverse visibility conditions (even zero ceiling, zero visibility) can find the proper runway on the airport and land safely on it. A secondary and, at present, more practical function is to expedite the landing maneuvers during poor visibility conditions so that more airplanes can land at an airport in a given time. The present practice of "stacking" airplanes over a control point limits the number of landings at even the largest airport to about 4 per hour. With the aid of an instrument landing system, this figure is expected to increase to about 10 per hour.

31. National Bureau of Standards System of Radio Landing Aids. The basic concepts of nearly all present radio landing systems were

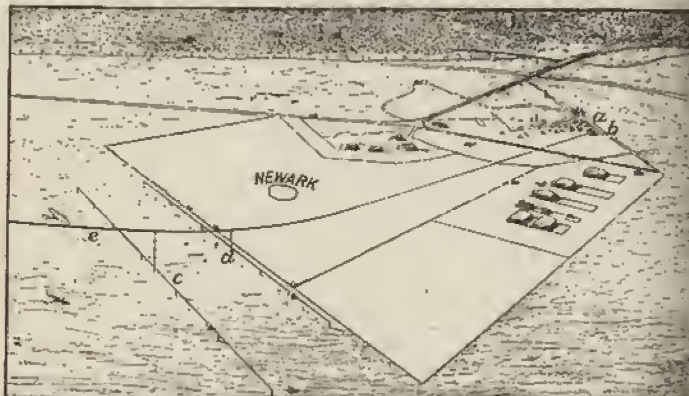


FIG. 19.—Radio landing system at Newark airport: (*a*) runway localizing beacon; (*b*) landing beam; (*c*) approach marker; (*d*) boundary marker; (*e*) spatial landing path followed by landing airplane.

demonstrated by the National Bureau of Standards at College Park, Md., in 1931 and at the Newark, N. J., municipal airport in 1933. The system included three elements to indicate the position of the landing airplane as it approached and reached the point of landing. Lateral position, given for the purpose of keeping the airplane directed to and over the desired landing-field runway, was secured by a small l-f (278-kc) visual-type radio range beacon. Approximate distance from this transmitter was given by a distance indicator operating from the automatic volume control in the beacon receiving set. Exact longitudinal position was given by two h-f (10-Mc) fan-type marker beacons located 1,500 ft. from the approach boundary of the airport and at the boundary. Vertical guidance was given by an u-h-f (91-Mc) landing beam which produced a curved gliding path for the landing airplane.

The runway-beacon course extended some 15 miles along the projection of the runway so that the pilot could orient himself along it from a

convenient point on the main range-beacon course or after passing over the main range-beacon tower. Two low modulating frequencies were used for distinguishing the overlapping radiation patterns of the runway beacon, and the airplane course indicator was designed to separate these frequencies by reed filters, to rectify them, and to apply the rectified voltages in phase opposition to a zero-center pointer-type microammeter. The marker beacons were of a type somewhat different from the present fan markers but produced essentially the same vertical sheet of radiated energy; the approach marker beacon was modulated by a h-f note and the boundary marker beacon by a l-f note. These were heard in the pilot's headphones.

32. Radio Landing Beacon. The method by which the suitable indication of absolute height above ground was obtained will be evident from

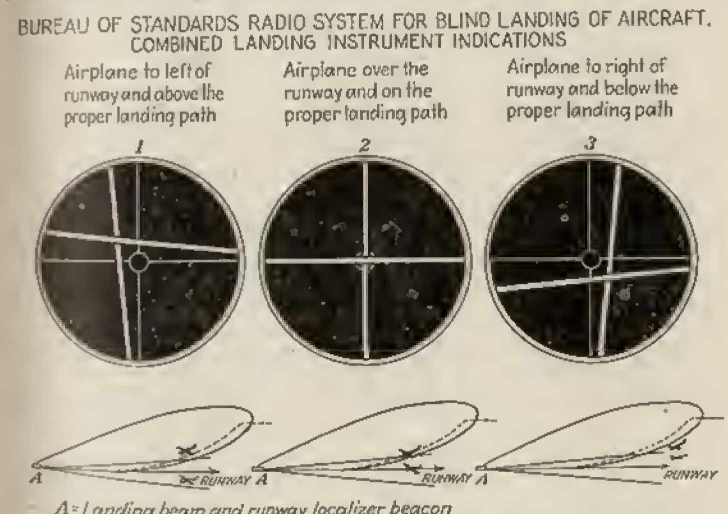


FIG. 20.—Combined-instrument indications for radio landing system.

Fig. 20, which shows the space radiation pattern in the vertical plane of the u-h-f radio landing beam. The polar pattern in the horizontal plane is of somewhat lower directivity. The airplane is therefore readily directed approximately along the horizontal axis of the beam by means of the course indications from the runway localizing beacon. It does not, however, fly along the inclined axis of the beam, but on a curved path whose curvature diminishes as the ground is approached. This path is a line of equal intensity of received signal below the inclined axis of the beam. The diminution of intensity as the airplane drops below the inclined axis is compensated by the increase of intensity due to approaching the beam transmitter. Thus, by flying the airplane along such a path as to keep constant the received signal intensity, as observed on a microammeter on the instrument board, the pilot comes down to ground on a curved line suitable for landing. If the airplane rises above this line of equal intensity of received signal, the microammeter deflection increases, while if it drops below this line the microammeter deflection decreases.

The landing-beam antenna system consists of a conventional directive antenna array. The vertical directive characteristic produced by the array operating in free space would be symmetrical about the horizontal plane, maximum radiation occurring in this plane. However, the presence of the ground, which acts as a perfect dielectric at these frequencies (91 Mc), modifies the vertical characteristic. At grazing incidence, *i.e.*, along the ground surface, the wave reflected from the ground cancels the direct wave to the receiving point, resulting in zero radiation. As the angle of elevation of the receiving point increases, there is an increasing difference in the distance traveled by the direct and reflected waves to reach the receiving point. The resultant phase difference produces increasing field intensity with increasing angle of elevation. When the phase difference is equal to half a period, the angle of maximum radiation is reached. It thus becomes evident how the landing beam shown in Fig. 20 is obtained.

Actually, there are a large number of lines of constant field intensity in the beam. It can be shown that the equation for these lines for the very low angles of elevation involved during a landing (less than 3 deg.) is a parabola. The particular parabola chosen to fit a given airport is a function of the transmitter power and the receiver sensitivity. Once chosen, it is essential that the position of this path in space does not vary. Horizontally polarized waves are used for the landing beam for this reason, since it has been determined that changes in the ground constants due to different weather conditions will have least effect on the landing path for this type of polarization. Also both the transmitter and receiver are designed to be of extreme simplicity to preclude the possibility of variation in the power output of the transmitter or in the receiver sensitivity.

33. Airplane Instrument. To simplify the indications used by the pilot, a combined instrument is employed for giving the runway-beacon and glide-path course indications. Two perpendicular reference lines are provided on the face of the combined instrument, the vertical reference line corresponding to the position of the runway and the horizontal reference line to the proper landing path. The pointers of the runway-course indicator and the landing-path indicator are arranged so that they cross each other, the former moving to the right or left of the vertical reference line and the latter above or below the horizontal reference line. The position of the point of intersection of the two pointers thus gives, through a single reading, the position of the airplane with respect to the runway and proper landing path. The instrument indications for several arbitrary positions of the airplane are given in Fig. 20. At 1 the airplane is to the left of the runway course and too high. At 2 the airplane is on the runway course and on the proper landing path. At 3 the airplane is to the right of the runway course and too low.

34. Modifications of System. There have been a number of modifications of this system which utilize the same basic principles. The *Lorenz system*, which is in extensive use abroad, utilizes a single transmitting set and antenna system for producing both the runway-beacon space pattern and the landing path. The antenna system comprises a vertical dipole radiator and two vertical reflector dipoles. The reflector dipoles are keyed, one to dashes and the second to dots, thereby producing two

corresponding space patterns which intersect in the horizontal plane and produce two beacon courses. The dashes and dots are interlocked producing a type of signal which may be converted for visual indication.

Since ultra-high frequencies are employed, the effect of the ground is to produce a vertical radiation pattern in which, for low angles of elevation, the field intensity is directly proportional to the angle of elevation starting from zero at the ground surface. A series of lines of constant field intensity having the shape of parabolas thus exists, as in the case of the regular landing beam, and may be used as landing paths.

Summarizing: the pilot follows a line of constant field intensity in the plane of intersection of the two beacon space patterns. On the airplane a single u-h-f receiving set is sufficient for reception of both the runway-course and landing-path indications.

The *Bendix system*, in this country, also employs a combined runway beacon and landing beam but with horizontally polarized waves. The *Air Track* system, developed by Washington Institute of Technology, houses the runway-beacon and landing-beam transmitters and antennas in a trailer and the marker-beacon transmitters on motor cycles so that the complete system becomes portable and may be moved readily to provide service upon any airport runway as determined by the existing wind condition.

35. CAA Indianapolis System. As a result of extensive tests of experimental installations of the various modifications described, the CAA through its Radio Technical Committee for Aeronautics set down detailed specifications on performance requirements for a radio landing system. These related to course sharpness of the runway-localizing beacon, freedom from bends, marker-beacon directivity patterns, landing-path shape, instrumentation, conventions as to instrument pointer deflections etc., and were based on the combined experience among the represented agencies of over 3,000 blind landings. A four-way installation was set up at Indianapolis which conformed with these specifications, and 10 similar installations are now undergoing service tests at 10 principal airports of the United States.

Figure 21 shows the installation of facilities at Indianapolis. The runway localizing beacons are of the two-course type (without sector identification), described in Art. 10, and operate on a frequency of 109.9 Mc. The rated power output is 300 watts. Because of the sharp radiation patterns used, the course is very sharp and maintains its direction within 0.1 deg. Bends in the course, which make it difficult for the pilot to maintain his orientation, are reduced to less than 0.15 deg., of negligible order. The marker beacons are of the fan type and operate on 75 Mc, the rated power output of the marker beacon transmitter is 5 watts. The landing-beam transmitters operate on 93.9 Mc and are of 300 watts rated power output. A complete monitoring system is provided so that an operator in the airport control tower can start up all the equipment corresponding to each landing direction and, from indicating instruments, may obtain information on the correctness of operation of the various units.

36. Control of Shape of Glide Path. An outstanding contribution of the Indianapolis setup consists of the control provided over the shape of the glide path. Pilot experience indicated that the glide path should be essentially straight from an altitude of 1,500 ft. at about 5 miles from the airport up to the airport boundary; from this point it is desirable that the path become slightly parabolic in shape, intersecting the runway surface at approximately 1 deg. This path allows the use of a constant rate of descent during landing but flares off toward the end, thereby eliminating excessive shock at the instant of contact.

The method of obtaining the desired path shape is indicated in Fig. 22. The landing-beam transmitting antenna is located at a considerable distance to one side of the runway and forward along the runway (see also Fig. 21), and its horizontal directive pattern is so adjusted that the

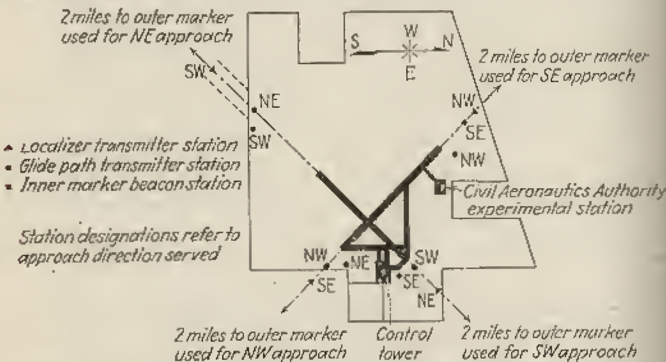


FIG. 21.—Indianapolis airport and instrument landing layout.

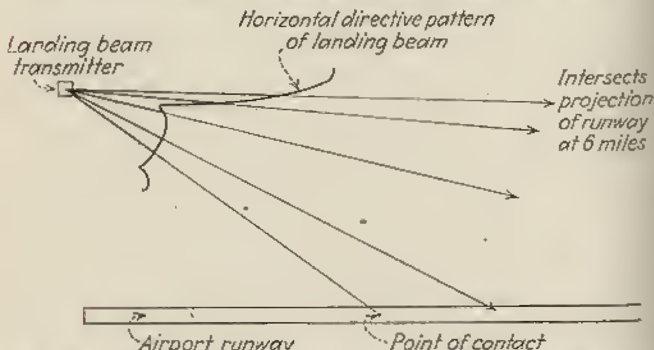


FIG. 22.—Method of controlling shape of glide path by horizontal directivity of landing-beam radiation pattern.

radiation to different points along the runway (and its projection) are of predetermined relative intensities. In this way the lines of constant field intensity in the vertical plane containing the runway may be made to have any desired shape.

37. The CAA-MIT Microwave System. This system employs very high frequencies, of the order of 750 to 3,000 Mc, for producing the runway-beacon, landing-beam, and marker-beacon patterns. Velocity-modulated transmitting tubes and horn-type radiators are employed.

By virtue of the highly directive patterns afforded, the glide path is of the equisignal type, being obtained by the overlapping of two patterns in the vertical plane so that a course making an angle of 3 deg. with the airport surface is produced.

The indicating system on the airplane combines the indications of the airplane bank, climb, and turn instruments with the indications afforded by the radio landing facilities. The indications are combined in a

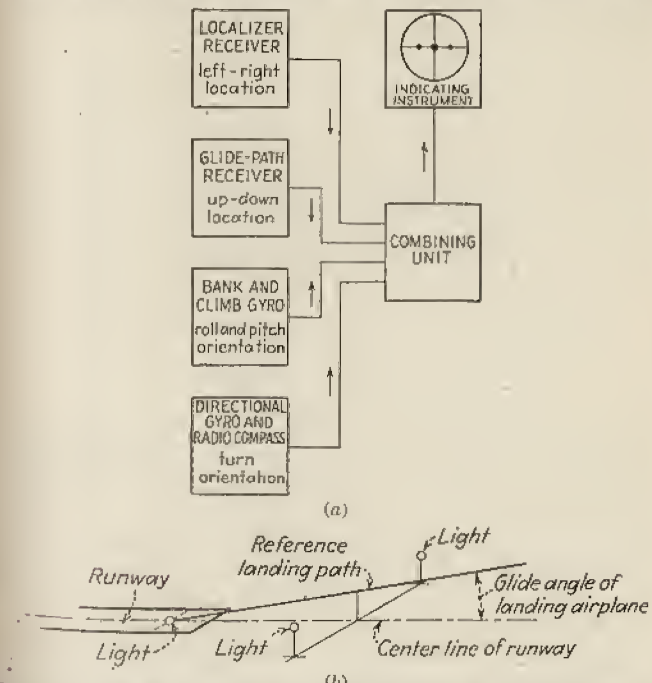


FIG. 23.—CAA-MIT microwave landing system: (a) combination of radio and airplane instrument indications; (b) three-light system on airport which gives equivalent visual indications.

cathode-ray indicator showing three spots; the central spot being controlled to the left-right or up-down by the radio aids and the other two spots by the gyro instruments (see Fig. 23a). The whole effect is to give exactly the same indications as when trying to fly a straight path visually by following three lights on a landing field oriented as shown in Fig. 23b. The indicating system may be applied to any three-element radio landing system and is similar in a number of respects to the Sperry "Flightray" developed at about the same time.

38. Army Landing System. A system of radio landing aids developed by the U. S. Army Air Corps utilizes a visual-type radio direction finder for lateral guidance of the landing airplane along the airport runway, two u-h-f marker beacons for longitudinal guidance, and a sensitive-type barometric altimeter in combination with these other elements for vertical guidance. The marker beacons now used in the CAA Indianapolis system are based on the type developed by the Army for its landing system.

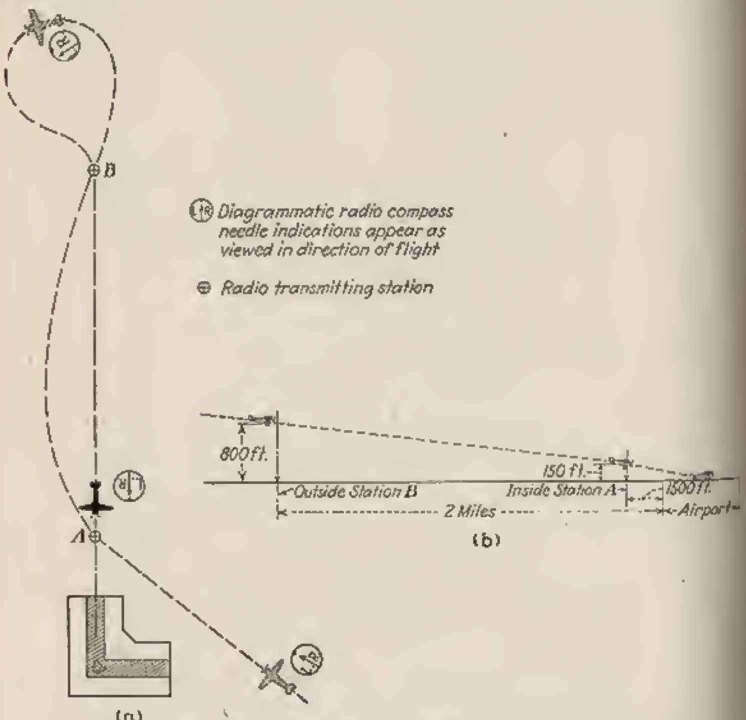


FIG. 24.—Army Air Corps radio landing aid.

The operation of this system is best understood by reference to Fig. 24. A and B are ground transmitting stations located along the projection of the center line of the airport runway. They send out tone-modulated transmissions suitable for use of the radio compass on the airplane. The power rating of these transmitters is approximately 50 watts, and the antennas used are vertical masts approximately 30 ft. high. Station A is placed approximately 1,500 ft. from the approach end of the landing field and station B about 2 miles from it. At each of the stations there is also located a low-power marker beacon operating on a frequency of about 60 Mc and using a half-wave transmitting antenna oriented along the direction of approach of the landing airplane.

Upon reaching the general vicinity of the airport through the use of the main radio range beacon, the pilot tunes his radio direction-finder receiver to station A and, upon reaching it, tunes to station B. He flies back and forth between these two stations as many times as is necessary to establish his course along the projection of the airport runway, setting his directional gyroscopic compass to the value found for that course. The necessity for this maneuver is apparent from a study of Fig. 16 in Art. 26. To compensate for possible departure from the true course due to cross winds, it is essential that the pilot determine exactly the required angle of crabbing of the airplane into the wind. This is particularly important in the case of narrow approaches to the airport with hazards located alongside of the approaches.

Upon establishing the proper course, the pilot makes a final approach to the landing field. The sensitive barometric altimeter is corrected to the barometric pressure obtaining on the ground, as determined by radio information, and is then relied upon in combination with the other flight instruments to maintain the airplane in a glide such that the altitude is approximately 800 ft. over station B and 150 ft. over station A. Continuation of this glide results in contacting the airport surface.

The exact point of contact is not so definite as with the first system described, depending upon the usual errors in the barometric altimeter, errors in determining the angle of glide of the airplane under varying load and air conditions, and errors in estimating the magnitude of the component of the existing wind along the runway. For this reason the system is safely applicable only to the larger airports and may be used only as an approach system at small airports. Advancements in the art have improved the practicability of this system. A dual automatic direction finder incorporating two loop antennas tuned to stations A and B, respectively, has been developed by Lear Developments, Inc., to provide automatic simultaneous indication of the bearings to the two stations. This simplifies the approach procedure. The accuracy of the vertical guidance may also be improved through the use of a radio altimeter indicating absolute height above ground. With these devices the system becomes quite accurate and easy to follow.

39. Dingley Induction-type Landing System. This is a non-radio system utilizing the magnetic field surrounding two horizontal multiconductor cables to establish in space a path of constant electromagnetic field intensity which coincides with the desired landing path. The cables are laid on or below the ground surface on either side of the airport runway and its extension. A large loop formed by grounded cables is excited by a 500-cycle alternator in given phase, and smaller loops, formed by cables grounded at other points, are excited by the same alternator in opposite phase. In this way it is possible to control the intensity of the magnetic field so that up to a point, when approaching the airport, a line of constant intensity is parallel to the ground, while beyond this point it assumes any desired angle contacting the runway surface.

The system provides lateral and vertical guidance of the airplane. The equipment required on the airplane includes two collecting coils perpendicular to each other and forming angles of 45 deg. with the horizontal when the airplane is in normal flight, two tuned 500-cycle amplifiers, and a crossed-pointer instrument. With this arrangement it is possible for the pilot to determine his lateral position with respect to the runway as well as his vertical position with respect to the glide path.

The system is outstanding in the simplicity of ground and airplane equipment required. Its installation at a modern airport, in the form described, involves such practical considerations as right of way for the cables, cost of installation, etc. It has been tried successfully at the Naval Air Station, Lakehurst, N. J., using a length of cable of 9,000 ft. with a glide path angle of about 5.5 deg.; the course indications from a radio range beacon were used in the initial orientation of the airplane to enable it to get on to the landing path.

ABSOLUTE ALTIMETERS

One of the most important navigational factors in aviation is that of altitude. Both in point-to-point flight and during landing, reliance upon altitude with reference to sea level has proved a continuous burden. Altimeters for indicating the absolute height above ground fall into three classifications; the sonic altimeter, the capacity altimeter, and the reflection altimeter. Of these only the reflection altimeter has proved successful enough to warrant commercial sale.

40. Sonic Altimeter. In this method the time taken by sound to reach the ground and return to the airplane is measured. Knowing the velocity of sound, the height of the airplane above ground may be determined. In a model developed by the General Electric Co., two horns are employed: one, driven by an electric trip relay and plunger, sends down the sound wave, and the other receives it back again after reflection from the ground. An instrument, which is started by the emitted wave and stopped by the reflected wave, records all heights above 50 ft., while below 50 ft. the pilot uses his headphones. At 50 ft., the echo comes back $\frac{1}{10}$ sec. after the emitted sound is sent out, at 5 ft. it comes back $\frac{1}{100}$ sec. later. A sound-delay filter is used in the output of the receiving horn so that the whistle and the echo do not blend into one sound until the airplane is at some point below 5 ft. This indication may be used effectively by the pilot during a landing.

An experimental unit based on this principle and developed by the Bell Telephone Laboratories, Inc., provides visual indication of the height above ground down to a few inches. In this system the received signal automatically starts the transmitted signal, so that the frequency of occurrence of the emitted sound increases with decreasing altitude. An arrangement of neon lights is used for obtaining the visual indication.

41. In the capacity altimeter the distance from the ground is measured by detecting the change in the electrical capacity between two plates on the airplane as the airplane approaches the ground. In one arrangement this capacity is made a part of a resonant circuit, coupled to an extremely stable r-f oscillator. A vacuum-tube voltmeter records the voltage developed across a portion of the resonant circuit. The circuit is adjusted so that the voltmeter-indicating instrument reads zero when the airplane is at any height above 100 ft. The gradual increase in capacity as the airplane approaches the ground serves to bring the resonant circuit into closer tune with the oscillator frequency, the voltmeter indication increasing accordingly. The indicating instrument, once calibrated, serves to indicate true height above ground. Since the capacity between the two plates is practically unchanged at altitudes greater than of the order of 100 ft., the field of usefulness of the capacity altimeter is limited to landing operations only.

42. Reflection Altimeter. In the reflection altimeter, altitude above ground is indicated by sending a radio wave to the ground and timing the interval required for it to reach the ground and return to the airplane after it has been reflected from the ground. The equipment used comprises a transmitter, a receiver, and a frequency meter operating a "terrain-clearance" meter. The transmitter is varied in frequency from 420 to 445 Mc and return at the rate of 60 times per second by means of a small condenser in the oscillator circuit (see Fig. 25). The rate of

change of frequency is thus 3×10^9 cps. The transmitter output is radiated downward by a half-wave horizontal doublet mounted on one of the lower surfaces of the airplane, and the receiver is connected to a second half-wave horizontal antenna similarly mounted and arranged for minimum coupling to the transmitting antenna. The direct and reflected waves picked up by the receiving antenna are fed to the detector and the detector output is amplified and fed to an electronic frequency meter. It will be evident that the frequency of the signal output from the detector is equal to the instantaneous frequency difference existing between the direct and reflected waves and is directly proportional to the

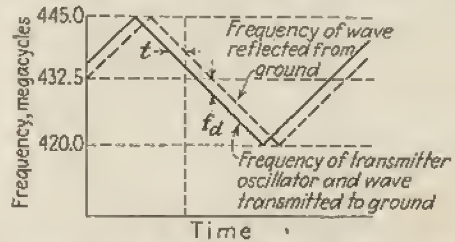


FIG. 25.—Principles of operation of Western Electric reflection altimeter.

height above the ground. Thus the frequency meter may be calibrated directly in feet.

For any altitude H, the frequency difference f_d is given by the expression

$$f_d = 3 \times 10^9 \times \frac{2H}{C}$$

where C is the velocity of light in the same units as H. Thus assume that H is 0.5 mile; the frequency difference is then 16,200 cps. This amounts to approximately 6 cycles per foot.

The upper range of the "terrain-clearance" indicator is arbitrarily limited to 5,000 ft. although the frequency meter is capable of responding to frequencies corresponding to about 15,000 ft. The reason for the upper limit is to obtain reasonable sensitivity of indication within the most useful range. If desired, a range-change switch could be used to provide indications over such altitude ranges as 0 to 500, 0 to 1,000, 0 to 5,000, and 0 to 15,000 ft.

As the radiation toward the ground is directed over a rather large area, the altitude above ground indicated will be the average of the elevation covered by a cone of about 30 deg. Thus buildings and other small obstacles will affect the readings only at low altitudes. Clouds will have negligible effect on the readings since the reflection coefficient of radio waves from clouds is only about 0.01 per cent of that from land.

FUTURE DEVELOPMENTS

43. Radio aids of the future will probably include radio control of airplanes, radio weather teletype, and radio aids for collision prevention. By radio control of airplanes is meant the harnessing of the various radio indicating instruments to the automatic pilot so that the airplane will automatically follow a predetermined course and the pilot's duties will correspond more nearly to that of the captain of a gyro-controlled ocean liner. The various instrumentalities for accomplishing this are available and have already been tested in isolated experiments.

Radio weather teletype would provide the pilot with printed weather reports and other messages which would facilitate operations and furnish reference records. The CAA has demonstrated such a service to be practicable and will probably begin experimental operation in the near future.

Radio aids to furnish the pilot information on the presence in his vicinity of other airplanes or of obstacles (such as mountains) would be of inestimable value. Very little has been done toward the provision of such aids, partly because of the complexity of the problem.

References

Books:

Fassbender, H.: "High-frequency Technique in Aircraft," Verlag Julius Springer, Berlin, 1932.

Morgan, H. K.: "Aircraft Radio," Pitman Publishing Corporation, New York, 1939.

Redfath, P. H., and J. M. Coburn: "Airline Navigation," Pitman Publishing Corporation (in preparation).

Publications:

Aero Digest, aircraft radio number, September, 1939.

Alford, A., and A. G. Kandocan: Ultra-high-frequency Loop Antennas, *A.I.E.E. Tech. Paper* 40–45, January, 1940.

Bowles, E. L., W. L. Barrow, W. M. Hall, F. D. Lewis: The CAA-MIT Microwave Instrument Landing System, *A.I.E.E. Tech. Paper* 40–44, January, 1940.

Dellinger, J. H.: Applications of Radio in Air Navigation, *Mech. Eng.*, **49**, 29–32, January, 1927.

Diamond, H.: The Cause and Elimination of Night Effects in Radio Range-beacon Reception, *Bur. Stand. Jour. Research, Research Paper* 513, **10**, 7–34, 1933; *Proc. I.R.E.*, **21**, 808–832, 1933.

———: Performance Tests of Radio System of Landing Aids, *Bur. Stand. Jour. Research, Research Paper* 602, **11**, 463–490, October, 1933.

———, and G. L. Davies: Characteristics of Airplane Antennas for Radio Range Beacon Reception, *Bur. Stand. Research*, May, 1931; *Proc. I.R.E.*, **20**, February, 1932.

———, and F. W. Dunmore: A Radio Beacon and Receiving System for Blind Landing of Aircraft, *Bur. Standards Jour. Research, Research Paper* 238, **5**, 897, 1930; *Proc. I.R.E.*, **19**, 585, April, 1931.

———, and ———: Experiments with Underground Ultra-high-frequency Antennas for Airplane Landing Beam, *Bur. Standards Jour. Research, Proc. I.R.E.*, December 1937.

———, and F. G. Gardner: Engine Ignition Shielding for Radio Reception on Aircraft, *Bur. Stand. Jour. Research, Research Paper* 158, **4**, 415–424, March, 1930; *Proc. I.R.E.*, **18**, 840–861, May, 1930.

———, W. S. Hinman, Jr., F. W. Dunmore, and E. G. Lapham: Upper-air Weather Soundings by Radio, *A.I.E.E. Tech. Paper* 40–47, December, 1939.

Dieckman, M.: Report on the Status of a Direct-indicating Airplane Radio Direction Finder, July, 1927, *Deutsche Versuchsanstalt für Luftfahrt*, Berlin.

Dingley, E. N.: An Instrument Landing System, *Communications*, **18**, 7, June, 1938.

———: A True Omnidirectional Radio Beacon, *Communications*, **20**, 5–6, 35, January, 1940.

Dunmore, F. W.: A Course Indicator of Pointer Type for the Visual Radio Range-beacon System, *Bur. Stand. Jour. Research, Research Paper* 336, **7**, 147–170, 1931; *Proc. I.R.E.*, **19**, 1579–1605, 1931; A Method of Providing Course and Quadrant Identification with the Radio Range-beacon System, *Bur. Stand. Jour. Research, Research Paper* 593, **11**, 309–325, 1933.

Engel, F. H., and F. W. Dunmore: A Directive Type of Radiobeacon and Its Application to Navigation, *Bur. Stand. Sci. Paper* 840, 1923.

Espenschied, L., and R. C. Newhouse: A Terrain Clearance Indicator, *Bell System Tech. Jour.*, **18**, 222, January, 1939.

Fischer, H. B.: Remotely Controlled Receiver for Radio Telephone Systems, *Proc. I.R.E.*, **27**, 264–269, April, 1939.

Gill, T. H., and N. F. S. Hecht: Rotating-loop Radio Transmitters and Their Application to Direction Finding and Navigation, *Jour. I.E.E. (London)*, **66**, March, 1928.

Haller, G. L.: Constants of Fixed Antennas on Aircraft, *Proc. I.R.E.*, **26**, 415, April, 1938.

Harrison, A. E.: Geographical Separation of Radio Range Stations Operating on the Same or Adjacent Frequencies in the 200–400 Kc Band, *Report* 4, Safety and Planning Division, CAA, January, 1938.

Hinman, W. S., Jr.: A Radio Direction Finder for Use on Aircraft, *Bur. Stand. Jour. Research, Research Paper* 621, **11**, 733–741, 1933.

Holloman, G. V.: Automatically Controlled Blind Landing, *S.A.E. Jour.*, **42**, 13, June, 1938.

Hepke, H. M.: Snow Static Effects on Aircraft, *Communication and Broadcasting Eng.*, **4**, 7, July, 1937; *Proc. I.R.E.*, **27**, 301, April, 1939.

Jackson, W. E., A. Alford, P. F. Byrne, and H. B. Fischer: Development of the CAA Instrument Landing System at Indianapolis, *A.I.E.E. Tech. Paper* 40-43, December, 1939.

———, and H. I. Metz: Development, Adjustment and Application of the Cone-type (Z) Marker, *Report* 16, Safety and Planning Division, CAA, July, 1938.

———, and D. M. Stuart: Simultaneous Radio-range and Telephone Transmission, *Proc. I.R.E.*, **25**, 314, March, 1937.

Kear, F. G.: Phase Synchronization in Directive Antenna Arrays with Particular Application to the Radio Range Beacon, *Bur. Stand. Jour. Research, Research Paper* 581, **11**, 123–139, 1933.

Kear, F. G., and G. H. Wintermute: A Simultaneous Radiotelephone and Visual Range-Beacon for the Airways, *Bur. Stand. Jour. Research, Research Paper* 341, **7**, 261–287, 1931; *Proc. I.R.E.*, **20**, 478–515, 1932.

Kramar, E.: A New Field of Application for Ultra-short Waves, *Proc. I.R.E.*, **21**, 1519–1531, 1933.

Kramar, E.: Rotary Radio Beacons, *Aero Digest*, April, 1938, p. 38.

———, and W. Hahnemann: Ultra-short-wave Guide-ray-beacon and Its Application, *Proc. I.R.E.*, **26**, 17–44, January, 1938.

McKeel, P. D., J. M. Lee, and H. I. Metz: Development of an Improved U-h-f Radio Fan Marker, *Report* 14, Safety and Planning Division, CAA, July, 1938.

Murphy, W. H.: Space Characteristics of Antennae, *Jour. Franklin Inst.*, **201**, 411, 1926; **203**, 289, 1927.

———, and L. M. Wolfe: Stationary and Rotating Equi-signal Beacons, *Jour. S.A.E.*, **19**, 209, September, 1926.

Nance, H. H.: Wire Communication Aids to Air Transportation, *Bell System Tech. Jour.*, 11, 462–476, July, 1932; *Elec. Eng.*, **51**, 492–496, July, 1932.

Nelson, E. L., and F. M. Ryan: Provision of Radio Facilities for Aircraft Communications, *S.A.E.* preprint (St. Louis meeting), February, 1930. See also numerous papers in *Bell Lab. Rec.* and *Bell System Tech. Jour.* by these and other authors.

Norton, K. A.: Summary of statement before the FCC Television Hearing, Jan. 15, 1940.

Pratt, H.: Apparent Night Variations with Crossed-coil Radio Beacons, *Proc. I.R.E.*, **16**, 625, May, 1928.

Rettenmeyer, F. X.: Some Problems of Aviation Radio, *R.C.A. Rev.*, **1**, 113, April, 1937.

Sandretto, P. C.: Some Principles in Aeronautical Ground-radio Station Design, *Proc. I.R.E.*, **27**, 5–11, January, 1939.

Smith, S. B.: The Night Performance of Marconi-Adcock Direction Finder, *Marconi Rev.*, No. 50, September-October, 1934.

Smith-Rose, R. L.: Radio Direction Finding by Transmission and Reception, *Proc. I.R.E.*, **17**, 425–478, March, 1929.

Stuart, D. M.: Circuit Design of Low-frequency Radio-ranges, *Report* 8, Technical Development Division, CAA, November, 1939.

SECTION 18

ANTENNAS

By Edmund A. Laport[1]

INTRODUCTION

The transmission and reception of electromagnetic waves used for radi communication are accomplished by radiators and collectors exposed space and known as *antennas*. An antenna is a device composed of system of one or more linear conductors, usually of large electric dimensions, from a fraction to several wave lengths, which is used to couple a h-f a-c generator or receiver to space. Between the transmittin and the receiving antenna there is a combination of earth, water, air and ionospheres which constitute the mediums in which electromagnet waves are propagated. The action of the waves in traversing thes mediums is very complex at best, being dependent upon many know and other unknown factors. Prominent among the known factors ar the transmitting frequency, the radiation characteristics of the transmit ting antenna, the orientation of the path of transmission in the earth magnetic field, the time of day and the conditions of daylight and dark ness along the path, the season of the year, solar activity, the electric characteristics of soil or water in the immediate vicinity of the antenn as well as along the path of the surface waves, the immediate condition of ionization of the atmosphere at various levels, the distance betwe transmitter and receiver, and the characteristics of the receiving antenn

1. Antenna Terminology. The following terms are used in this work

1. *Meter-amperes.* In general, this means $\int i\,dl$, where i is the r-m-s curre in an elementary length of the antenna, dl. The integration is perform over the entire length of the exposed (radiating) parts of the radiator. View geometrically, this is the area of a plot of r-m-s antenna current in amper against distance along the antenna measured in meters. The directions the currents must be considered.

2. *Doublet.* A differential of antenna length, short enough to be co sidered to have uniform current throughout its length.

3. *Dipole.* A linear conductor with a full half wave of in-phase curren distributed throughout its length. A half-wave oscillating element.

4. *Self-impedance.* The impedance of a single radiating element in th absence of any influences from other radiators, as measured at a curre antinode. The ratio of the impressed voltage and the antinode current.

5. *Mutual Impedance.* The circuital equivalent of radiation couplin Mathematically expressed, it is the negative ratio of the induced potential the base (or the current antinode) of a second radiator to the base curre (or antinode current) of the first radiator.

6. *Harmonic.* Any natural frequency of oscillation of a system express as a number which is the multiple of the fundamental frequency. Not to confused with overtones.

[1] RCA Victor Co. Ltd., Montreal, Que.

7. *Antenna Tuning.* The act of resonating an antenna system to some frequency other than a natural frequency by means of reactive devices.

8. *Antenna Loading.* Lumped reactances connected in the antenna system for the purpose of antenna tuning.

9. *Distributed Loading.* Units of reactance added at small electrical intervals along a conductor for the purpose of smoothly modifying the natural distributed constants of the system. Pupinization.

10. *Node, or Nodal Point.* In a standing wave system, the points of either zero or minimum potential or current.

11. *Antinode.* In a standing wave system, the points of maximum potential or current.

12. *Vertical Polarization.* A wave orientation such that all the lines of electric force lie in planes perpendicular to the ground plane.

13. *Horizontal Polarization.* A wave orientation such that the lines of electric force are parallel to the ground plane.

14. *Elliptical Polarization.* A field of force having both vertically and horizontally polarized components.

15. *Reflector.* Conductor or conductors so disposed with respect to a radiator as to react upon the latter in a manner which transforms the radiation pattern by suppressing radiation in its direction while reinforcing it in the opposite direction.

16. *Antenna Array.* A multiplicity of radiating elements disposed in any manner whatsoever for the purpose of molding the space characteristic in some desired fashion.

17. *Space Characteristic.* A means for describing the over-all radiation characteristics of an antenna system. Usually refers to a geometrical solid in spherical coordinates with distance from the origin proportional to the radiation intensity in any direction. Radius vectors may be proportional to field intensity or to power. Power flow by radiation in any direction is proportional to the square of field intensity.

18. *Fundamental Frequency.* The frequency at which the impedance of an antenna at a current antinode is minimum. The lowest frequency of oscillation of an antenna.

19. *Fundamental Wave Length.* The length of the space wave emitted by an antenna oscillating at its fundamental frequency.

20. *Mode of Operation.* The ratio of the operating wave length to the fundamental wave length; also, the ratio of the fundamental frequency to the operating frequency.

21. *Electrical Length.* The length of a standing wave in any linear system expressed in degrees or radians. The electrical length of a wire is its actual length in terms of wave lengths and fractions thereof multiplied by 360 deg. Valid only in systems with sinusoidal current distribution.

22. *Effective Height.* The height h obtained from the following equation:

$$h = \frac{\epsilon d}{1.25fI}$$

where h = effective height in meters
ϵ = measured field intensity in microvolts per meter
d = distance in kilometers from the antenna to the point where ϵ is measured
f = frequency in kilocycles
I = antenna current at the point where the antenna is energized.

Note. d must be small enough so that the effect of attenuation is absent, and great enough to be beyond the limits of the induction field.

23. *Antenna Resistance.* The total dissipative component of the antenna impedance measured at the point where power is introduced.

24. *Radiation Resistance.* The ratio of the total power radiated by an antenna and the square of the current at some reference point in the system, usually the point where power is introduced, or a current antinode.

25. *Oscillating Wire.* A linear conductor containing a standing wave of oscillatory energy.

2. Radiation from Linear Conductors. The existence of a field of force in either electromagnetic or electrostatic form represents a storage of energy in space. Faraday originated the descriptive method of picturing a field in terms of lines of force and lines of equal intensity which formed the basis for subsequent curvilinear geometry which is now more or less familiar to electrical engineers. In ordinary electrical engineering it is customary to concentrate a field as much as possible and to prevent stray lines of force, known as *leakage flux*, from reaching any considerable distance from an electrical device. In antenna design, however, the opposite case is desired. Here as much as possible of the energy of the field is made to be stored in space which is far removed from the conductor. The linear straight conductor is the most satisfactory practical device for producing distant fields.

In ordinary electrical devices the energy of the electric or magnetic field is returned to the parent circuit when the charge or current that produced it is removed. The field collapses. It takes time for a field to be propagated from one point to another in space, so that its formation or disappearance at any point is not coincident with the events in the conductor which produce it. The finite rate of propagation of electric and magnetic fields, 3×10^8 meters per second, causes events in the field of force to lag behind the events in the parent circuit by a time dependent upon the distance from the circuit. This fact is of no great interest ordinarily, but in connection with antennas it is of primary importance and forms the basis of all radiation phenomena. Assume for simplicity a straight wire which is charged. The electric field has been established out to a very great distance. If the charge be removed suddenly, the collapse of the field will return the stored energy to the circuit after a suitable time interval. If, on the other hand, the charge on the wire be instantly reversed, a field of the opposite polarity forms near the wire before the energy stored in space for the previous charge can return to give up its energy to the wire. The original field becomes detached and manifests itself as a free wave of electric energy traveling in space. When, instead of an instantaneous reversal of the charge on the wire, there is a gradual reversal at a rapid rate under the stimulus of a h-f generator, some of the energy of the field very near to the wire returns to the circuit before the charge reverses, but a large amount of energy in the more distant fields is unable to return before reversal occurs and becomes a detached field of force, an electromagnetic wave. That portion of the field which returns its energy to the circuit is known as the *induction field* and the detached portion as the *radiation field*. The energy lost by radiation is represented in the impedance of the circuit by the radiation resistance.

Radiation of energy takes place from linear conductors which are electrically unbalanced. When it is desired to prevent radiation, two parallel conductors are placed very close together electrically and equal and opposite charges are distributed identically along the conductors. To produce radiation, the spacing between conductors is increased and the balance of charges upset more and more. The ultimate in this direction is that of the familiar simple antenna, a single straight wire which is completely unbalanced.

3. Fundamental Radiation Formula.[1] Dellinger's derivation of the fundamental radiation formula for an antenna is of the following form:

$$H_t = -\frac{\omega \int_0^l i dl}{10cd} \cos \omega\left(t - \frac{d}{c}\right) - \frac{\int_0^l i dl}{10d^2} \sin \omega\left(t - \frac{d}{c}\right)$$

where H_t = instantaneous magnetic field intensity in gilberts per centimeter
ω = 2π times frequency of oscillating current in the wire
i = instantaneous current at any point in the wire
l = length of the wire in centimeters
t = time in seconds
c = velocity of propagation of light (3×10^{10} cm per second)
d = distance from the wire in centimeters perpendicular to the wire.

The first term is known as the *radiation-field term*, and the second as the *induction-field term*. These two terms are of equal magnitude where $d = \lambda/2\pi$.

After converting the radiation term into the most practical units for engineering usages, we have

$$E_d = 377 \frac{\int_0^l I\,dl}{\lambda d} \quad \text{(for free space transmission)}$$

where E_d = millivolts per meter (field intensity)
I = r-m-s current in amperes in each elementary length dl
λ = wave length in meters
d = distance from antenna (in normal direction) in kilometers
$\int_0^l I\,dl$ = total meter-amperes of system.

For a half-wave dipole in free space, with sinusoidal current distribution,

$$E_d = \frac{60I}{d} \quad \text{(for free space transmission)}$$

where E_d = volts per centimeter (field intensity)
I = r-m-s current at the antinode
d = distance in centimeters.

The field intensities in directions other than normal to the wire depend upon the length of the wire and the distribution of currents in it.

4. Current and Potential Distribution in Straight Wires. The action of an oscillating wire as an antenna depends upon the current and potential distribution in the wire. These distributions in turn are dependent upon the manner in which charges are propagated in the wire under various conditions of excitation by a h-f generator. If an uncharged wire be connected to a source of h-f energy, charges move from the generator into the wire, travel along the wire, and, after an interval of time depending upon the length of the wire and the velocity of propagation of the charges, arrive at the distant end. If the end of the wire is an open circuit, as most antennas are, there will be a transformation of energy at the end which causes the potential there to double and the current to become zero. The high potential at the end, due to the accumulation of charges which continue to be supplied by the generator, causes another wave of energy to be propagated from the open end back to the generator.

The mechanism of the production of standing waves of current and potential on a linear conductor may be studied analytically by referring

[1] Dellinger, J. H., Principles of Transmission and Reception with Antenna and Coil Aerials, *Bur. Standards Sci. Paper* 354, 1919.

to any good text on reflections in transmission lines. A very elementary introduction is given here, however, to establish a physical picture of this important phenomenon. Consider Fig. 1, which is a wire which may be connected to ground through a battery B by closing the switch S. At first the switch is open, and the wire is at zero potential. Upon closing the switch, a current flows from the battery into the wire, and the wire becomes charged. The charging process is not instantaneous, because time is required for charges to travel the full length of the wire. The current flow from the battery into the wire persists, not only for the duration of the movement of charges to the end of the wire, but until the wave of charges is reflected from the open end back to the battery. During this interval the battery supplies charges as it would to a line of infinite length. It is only when the reflected wave of charges arrives at the battery that the very finite length of the wire is manifest, at which time the excess charge on the entire length of the wire gives it a potential higher than that of the battery. When the wire is finally charged to battery potential, the total energy of the additional charges which compose the reflected wave must be eliminated from the system. After reflection, the wire is positive with respect to the battery, so that it may be said the battery is negative with respect to the wire and that it now starts to charge the wire with negative charges. The same process repeats itself until the original positive charges are neutralized and the wire is charged negatively. This continues cyclically. Owing to circuit losses, radiation, etc., there is a gradual consumption of the excess energy of the system, and each reflection is weaker than the one preceding. When the excess energy is consumed completely, the wire reaches steady state with a uniform potential throughout its length equal to that of the battery.

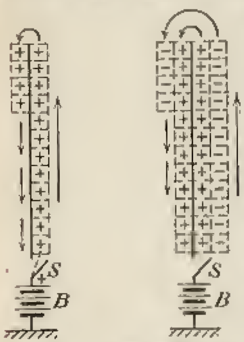


FIG. 1.—Charged wire.

When an a-c generator is used to energize the wire, the same process takes place, but, when the wire is "tuned" to the generator frequency, the reflected energy arrives at the generator when it is reversing its polarity, in which case the energy of the reflected wave is absorbed by the generator and is not re-reflected. Thus, in the typical antenna problem, the characteristic current and potential distribution is the result of a simple reflection—a wave of charges moving from the generator toward the end of the wire, and the reflection from the end back to the generator.

In the steady state both potential and current vary harmonically in time, but their maximum values vary with their position in the wire. In the simple straight-wire antenna the variation of potential and of current along the wire is very nearly cosinusoidal and sinusoidal, respectively, when measured from the open end of the wire.

There are important relationships between the potential and current distribution, on which the impedance of the antenna depends. By solving the case of a simple reflection in a wire, for example, a wire one-third wave length (or 120 deg.) long, on the assumption that there is no energy dissipated in the system during the propagation of the initial wave of

energy imparted to the wire from the generator and during its reflection from the open end, we obtain the vector relationships between current and voltage for several equidistant points along the wire as shown in Fig. 2. If electrical degrees are measured from the open end, it is seen that, up to a distance of 90 deg., the current vectors are in advance of the voltage vectors and the impedance of the antenna for lengths less than one-quarter wave length is a pure capacitive reactance. When the wire is longer than one-quarter wave length, as in the example, the wire impedance becomes a pure inductive reactance. If this were

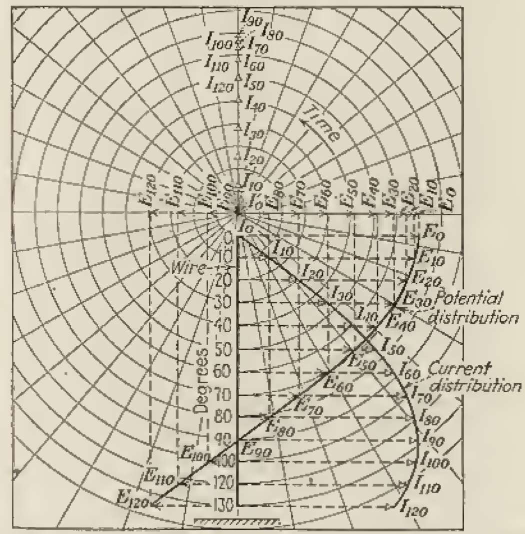


FIG. 2.—Vector relations between current and voltage in a wire 120 deg. long.

continued for several quarter-wave lengths, it would be seen that for odd quarter-wave lengths the impedance of the antenna would be capacitive reactance, and for even quarter-wave lengths it would be inductive reactance. At exactly 90 deg. and odd multiples of 90 deg. (potential nodes) the impedance would be zero, while for even multiples of 90 deg. (current nodes) the impedance would be infinite. By plotting out the current and potential vectors against their position along the wire, in two-dimensional rectangular coordinates, it is found that the potential varies cosinusoidally and the current sinusoidally.

Now there cannot exist in nature a dissipationless system. Waves of charges propagated in a wire suffer some attenuation. We know there are Joulian losses in the wire as well as loss of energy through radiation, especially in an antenna wire which is an efficient radiator. Working out the case of a simple reflection in a 120-deg. antenna wire on the basis of a considerable power loss in the wire, we get the vector diagram of currents and voltages shown in Fig. 3. Between 0 and 90 deg. of length the current vectors lead those of potential, but there is now a component of potential in phase with the current, so that the impedance in this range is resistance and capacitive reactance. At 90 deg. the potential vector is in phase with the current vector, at which point the antenna impedance is pure resistance. Beyond the 90-deg. (quarter-wave) point the voltage vectors swing into a leading position, and the antenna impedance becomes resistance and inductive reactance. This continues up to the half-wave point (not shown in our example), at which place the current vectors come into phase with the voltage. Here again the antenna impedance becomes a finite pure resistance, but of a very high value. In Fig. 3

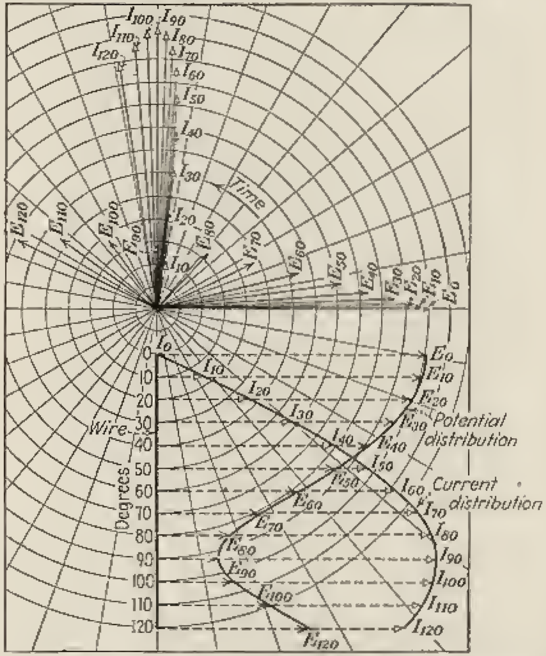


FIG. 3.—Currents and voltages in a wire of 120 deg. when appreciable power loss occurs.

the several vectors are plotted in rectangular coordinates against their position along the wire. It can be seen plainly that the potential distribution is not cosinusoidal, especially in the vicinity of the node. Voltage passes through a minimum, accompanied by a rapid change of phase, but does not become zero as in Fig. 2. If the wire were made a half-wave length long, so that the current would pass through its node, it would be seen that the current also passes through a minimum value, but not zero. The example of Fig. 3 is greatly exaggerated so as to clearly show the problem. In antenna systems the energy lost is so small with respect to the energy stored in them (very low attenuation of the traveling waves in the wire) that the current distribution is very nearly sinusoidal.

The radiating characteristics of an antenna depend upon the current distribution. When calculating radiation patterns for simple wire antennas, the assumption of sinusoidal current distribution is fully justified. The complex circuital impedance of an antenna, however, is the result of the true current and potential distributions which are not simple harmonic functions of distance along the wire.

5. Current and Potential Distributions in Linear Conductors with Attenuation.[1]

$$E_l = E_r\sqrt{\sinh^2 \alpha l + \cos^2 \beta l}\,\angle \tan^{-1}\,(\tan \beta l \tanh \alpha l)$$

$$I_l = \frac{E_r}{Z_0}\sqrt{\sinh^2 \alpha l + \sin^2 \beta l}\,\angle \tan^{-1}\left(\frac{\tan \beta l}{\tanh \alpha l}\right)$$

E_l and I_l are the voltage and current, respectively, at any point in the antenna wire which has the distance l from the open end. E_r is the voltage at the open end. αl is the attenuation constant in nepers (hyperbolic radians) per unit length, and βl is the wave-length constant in circular radians per unit length of the wire. Z_0 is the characteristic impedance of the wire. In an antenna this factor has no true scientific significance, but for many practical purposes a value can be placed upon it which has engineering significance.

6. Current Distribution in Antennas of Various Practical Forms. Radiation phenomena are usually studied in terms of the electromagnetic field, which is associated with the antenna currents. In matters involving space characteristics, field intensities, etc., the basis of reference is usually the current distribution. In many forms of antennas to be found in practice, there are numerous departures from the simple conditions which produce sinusoidal current distribution. In a single-wire T, current along the vertical portion is distributed as a partial sinusoid and can be calculated as a real part of the equivalent vertical wire. The current in the flat-top sections is linear, very nearly, if each branch is less than 30 deg. long. The current at the top of the vertical is divided equally between the two branches and tapers to zero, or nearly zero, at the ends of the T branches (see Fig. 4).

FIG. 4.—Current in simple T antenna.

The current distribution in a single-wire inverted L has also been shown to be nearly sinusoidal, as was assumed from theory.[2]

Non-sinusoidal distributions occur in systems that have non-uniform constants per unit length, such as fan, umbrella, and many other forms of multiwire antennas. Irregularities in the distributed L and C of the antenna are sources of reflections and lead to very complicated distributions. With the gradual disappearance of such systems, however, no particular attention need be directed to the matter here.

Large capacities at the end of a wire, such as insulator caps, rain shields, corona shields, outriggers, are equivalent, in their effect upon the current distribution, to an elongation of the wire.

[1] EVERITT, W. L., "Communication Engineering," Chap. VI, McGraw-Hill Book Company, Inc., New York, 1932.

[2] WILMOTTE, R. M., Distribution of Current in Transmitting Antennas, *Jour. I.E.E.* (London), June 1928; PIERCE, G. W., "Electric Oscillations and Electric Waves," McGraw-Hill Book Company, Inc., New York, 1920.

7. Antenna Potential and Potential Distribution. In the design of antenna insulation, potential magnitude and distribution must be calculated. Potential distribution can be calculated under the same conditions that current distribution can be, which is principally those cases where the distribution is very nearly cosinusoidal. The actual voltage at the feed point is the product of the antenna current and the antenna *impedance*. The potentials at other points in the system are obtained from the potential distribution with respect to the potential at the feed point for a given power input.

8. Reactances of Linear Conductors. The reactance of a linear conductor is given by the following formula:[1]

$$X = -\sqrt{\frac{L}{C}} \cot \omega l \sqrt{LC} = -Z_0 \cot G$$

where X = reactance in ohms, either positive or negative, depending upon angle $\omega \sqrt{LC}$

L and C = microhenries and microfarads per unit length of wire. (For calculating, refer to footnotes.[1,2])

$\omega = 2\pi \times$ frequency

Z_0 = characteristic impedance. (Values range from approx. 750 ohms for vertical wire antennas to as low as 200 ohms for uniform cross-section tower radiators.)

G = electrical length of conductor.

NOTE: This formula, based on sinusoidal current distribution, is unreliable for values of G close to 90 deg. and its multiples, and the error increases as G becomes very large.

9. Radiation Resistance. Useful data are shown in Figs. 5, 6, 7, and 8.

RESISTANCE OF STRAIGHT VERTICAL ANTENNA FOR DIFFERENT VALUES OF WAVE LENGTH OBTAINED BY INDUCTANCE AT THE BASE

λ/λ_0 ratio of wave length to natural wave length	G, deg.	R, radiation resistance, ohms
1.00	90	36.57
1.12	80.4	26.40
1.21	74.4	21.70
1.31	68.8	17.65
1.43	64	14.28
1.57	57.3	11.62
1.74	51.7	9.10
1.97	45.7	6.92
2.24	40	5.19
2.62	34.4	3.78
3.14	28.7	2.58
3.93	23	1.65
5.26	17.1	0.90
7.85	11.5	0.30
15.70	5.73	0.082
31.42	2.87	0.01

[1] Radio Instrument and Measurement, *Bur. Standards Circ.* 74.

[2] GROVER, F. W., Methods, Formulas and Tables for Calculation of Antenna Capacity, *Bur. Standards Paper* 568.

For the theory and calculation of radiation resistance, see:

PISTOLKERS, A. A., The Radiation Resistance of Beam Antennas, *Proc. I.R.E.*, March, 1929.

BECHMANN, R., On the Calculation of Radiation Resistance of Antennas and Antenna Combinations, *Proc. I.R.E.*, August, 1931.

HANSEN, W. W., and J. G. BECKERLEY, Concerning New Methods of Calculating Radiation Resistance, Either with or without Ground, *Proc. I.R.E.*, December, 1936.

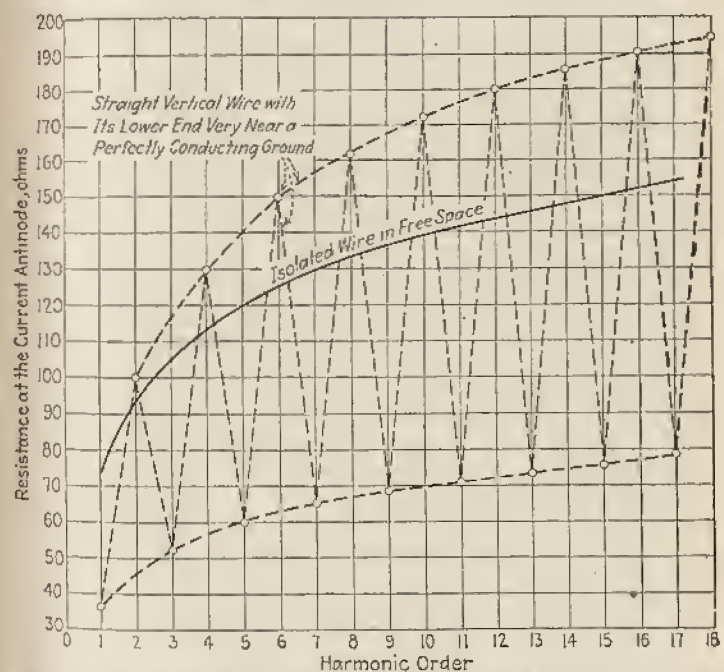


FIG. 5.—Radiation resistances referred to a current antinode for straight wire in free space and for straight vertical wire with lower end close to perfect earth.[1]

ALFORD, A Discussion of Methods Employed in Calculations of Electromagnetic Fields of Radiating Conductors, *Electrical Communication*, July, 1936.

PIERCE, G. W., "Electrical Oscillations and Electric Waves," Chap. IX, McGraw-Hill Book Company, Inc., New York, 1920.

SCHELKUNOFF, S. A., A General Radiation Formula, *Proc. I.R.E.*, October, 1939.

SIEGEL, E., and J. LABUS, Feldverteilung und Energieemission von Richtantennen, *Hochfrequenz Technik und Elektroakustik*, Band 38, Heft 6, 1932.

10. Self-impedance of an Antenna. The impedance of an antenna, as seen from the point where power is introduced, is usually complex. The resistive component is made up of the radiation resistance referred to

[1] LEVIN, S. A., and C. J. YOUNG, Field Distribution and Radiation Resistance of a Straight Vertical Unloaded Antenna Radiating at One of Its Harmonics, *Proc. I.R.E.*, May, 1926.

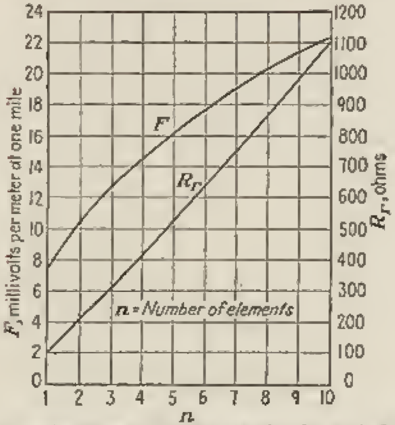


FIG. 6.—Radiation resistance at current antinode and field intensity at 1 mile radiated for a vertical array of colinear cophased dipoles.[1]

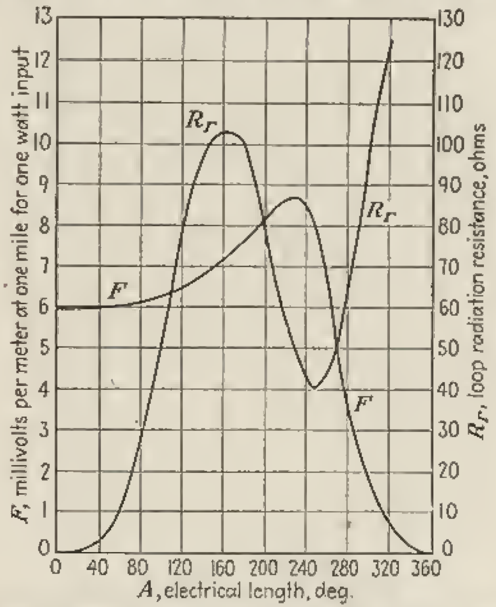


FIG. 7.—Radiation resistance and field intensity from simple vertical antenna over perfect earth.[1]

[1] BROWN, G. H., A Critical Study of the Characteristics of Broadcast Antennas as Affected by Antenna Current Distribution, *Proc. I.R.E.*, January, 1936.

the feed point, the conductor resistance, and the ground resistance. The reactive component is determined by the characteristic impedance of the antenna, the electrical length, and the influence of any top, distributed or base loading. Any distributed capacitance due to base insulators, protective gaps, drain coils, and any other devices that attach to the system near the feed point transform the true antenna impedance to a

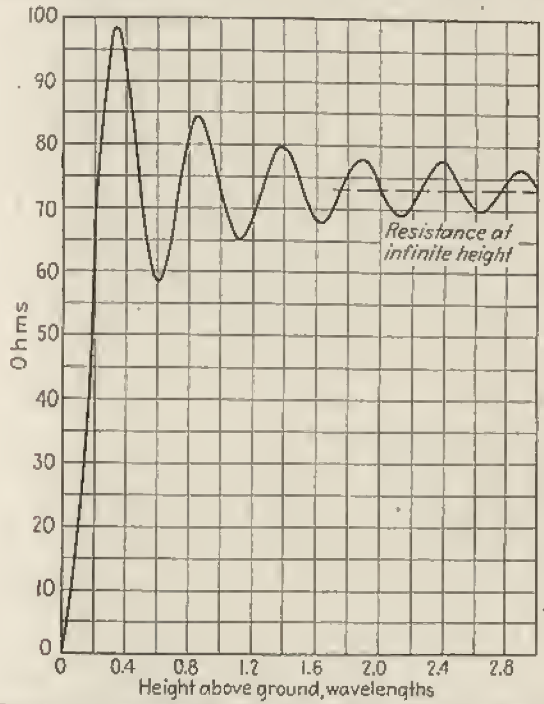


FIG. 8.—Radiation resistance at current antinode for horizontal dipole over perfectly conducting earth.[1]

new impedance which is the load actually seen by the generator at that point. We may call this point

$$Z_a = R_a \pm jX_a$$

The power input to the system is

$$W = I_a^2 R_a$$

and the potential across the load is

$$E \text{ (volts)} = I_a Z_a$$

[1] CARTER, P. S., Circuit Relations in Radiating Systems and Their Application to Antenna Problems, *Proc. I.R.E.*, June, 1932.

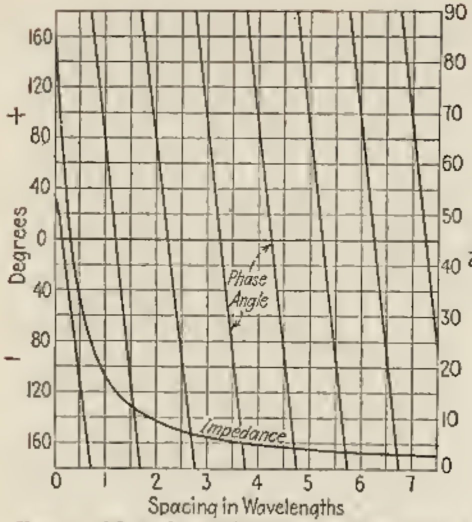


Fig. 9.—Mutual impedance for parallel dipoles.

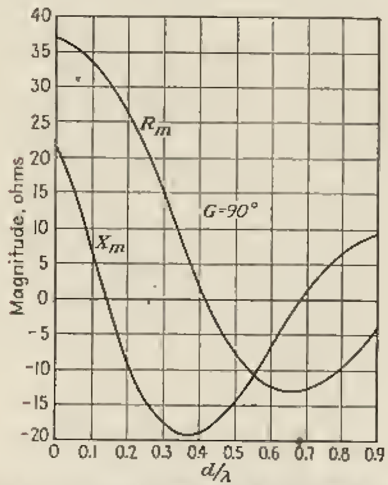
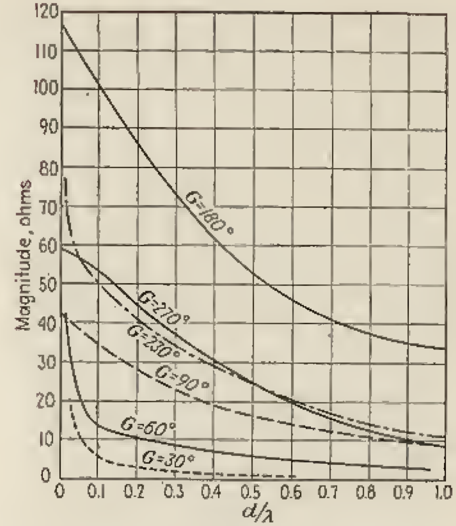


Fig. 10.—Resistive and reactive components of mutual impedance between two quarter-wire vertical radiators.



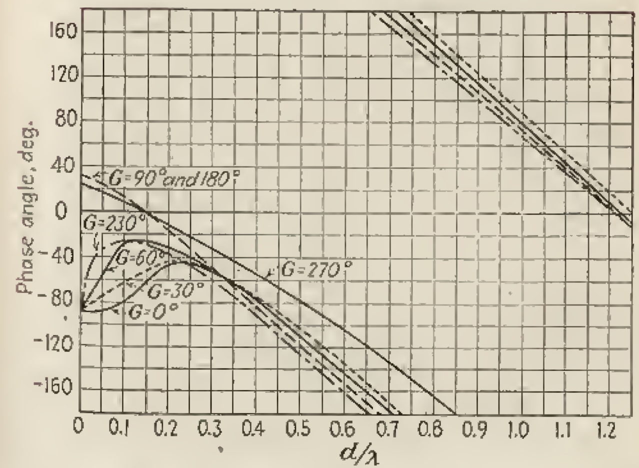


Fig. 11*a* and *b*.—Magnitude and phase angle of mutual impedance vector for identical vertical radiators.

11. Mutual Impedance. Whenever two conductors are disposed in space so that there is appreciable interchange of energy through radiation coupling, the circuital conception of mutual impedance is introduced to enable this reaction to be predicted and manipulated by the convenient methods of ordinary network theory. Mutual impedance is a vector quantity which may appear in any of the four quadrants. It is derived through the Maxwell field equations by an extension of a method of Kliatzkin and Pistolkors where the Poynting vector is integrated over the surfaces of the radiators. Mutual impedance must be reckoned with quantitatively in all directive antennas. There are reproduced in Figs. 9

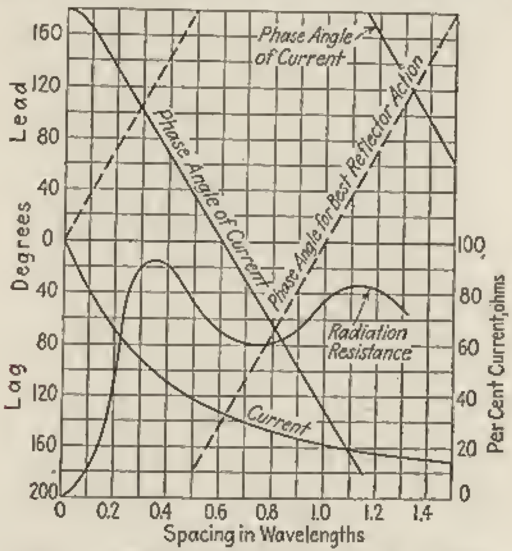


Fig. 12.—Conditions for array of parallel dipoles.

to 11 the data of most frequent practical value. For further information, consult the literature.[1]

12. Radiation of Electromagnetic Waves from Antennas. Effective application of antennas to practical communication problems makes use of special radiation characteristics made possible by the disposition of radiators, their length, current distribution, current phase, and amplitude relations. The radiation characteristics are so intimately associated with the physical and statistical conditions of wave propagation that the whole subject of propagation belongs with the subject of antennas. In all antenna applications maximum effectiveness requires objective control of the distribution of energy radiated into space.

[1] Carter, P. S., Circuit Relations in Radiating Systems and Their Application to Antenna Problems, *Proc. I.R.E.*, June, 1932; Brown, G. H., Directive Antennas, *Proc. I.R.E.*, January, 1937; Pistolkors, A. A., The Radiation Resistance of Beam Antennas, *Proc. I.R.E.*, March, 1929.

All radiation control is due to wave interference, and the space characteristics of antennas and arrays result from interferences between the fields produced by all the infinitesimal portions of all the radiators when

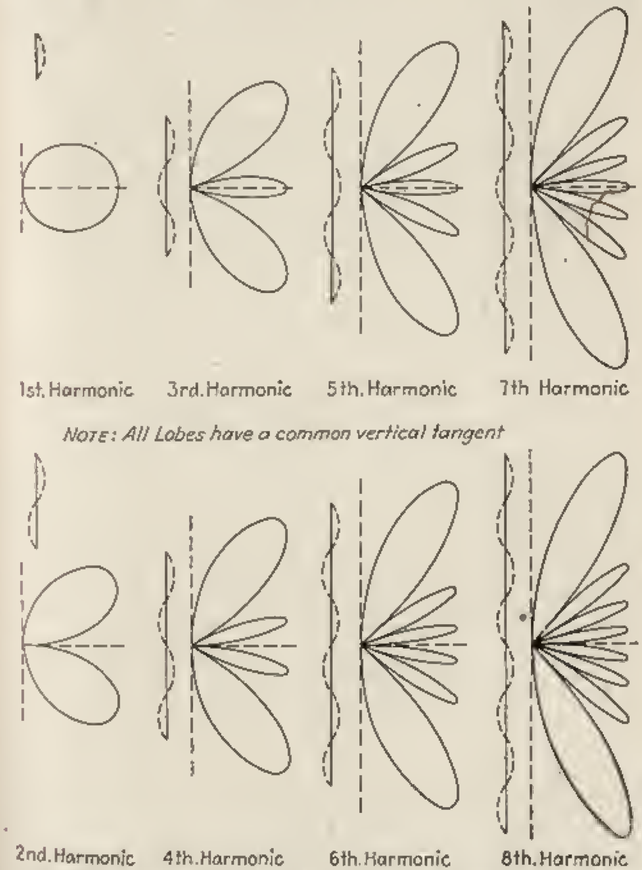


Fig. 13.—Polar diagrams of relative field strength distribution for straight wire antennas in free space with standing waves of current as shown, corresponding to the operation at various harmonics of the fundamental frequency of the antenna.

currents flow in them. For grounded antennas interferences result from wave reflections from the ground (image radiations), and for this reason the electrical constants of the earth have an important influence on the radiation patterns.

In dealing with radiation patterns it is customary for reasons of practicality to employ plots of relative or absolute field intensities, in convenient units, to depict the magnetic field distribution at various points on the surface of a hypothetical sphere or hemisphere when the antenna is located at its center. Three dimensions are thus required to specify

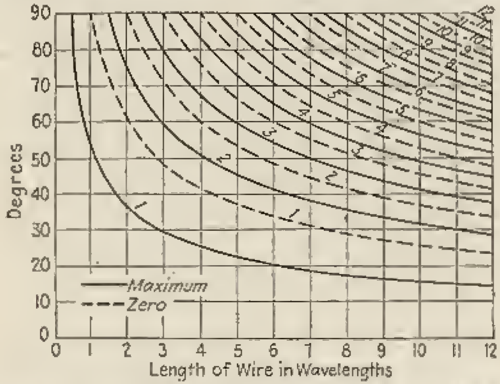


FIG. 14.—Angles at which nulls and maxima occur for the various patterns of Fig. 13.[1]

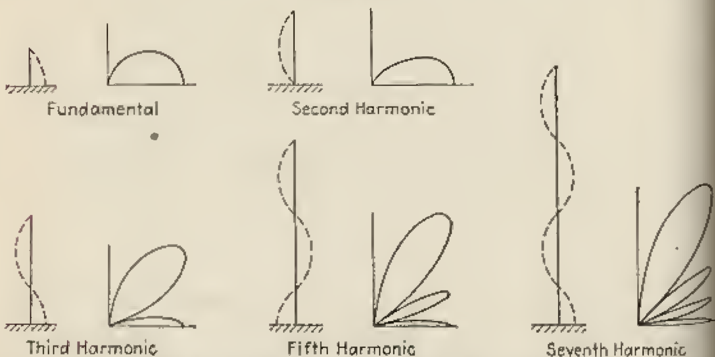


FIG. 15.—Polar patterns of field strength distribution from vertical grounded antennas with current distributions as shown.

space characteristics except in those cases where there is axial symmetry, when two-dimensional figures suffice. Three-dimensional data are shown as paper or plaster models, or as a family of two-dimensional curves.

Figures 13 through 18 show various basic forms of radiation patterns which are employed singly or in combination for radiation control.

[1] CARTER, HANSELL, and LINDENBLAD, Development of Directive Transmitting Antennas for RCA Communications, Inc., *Proc. I.R.E.*, October, 1931.

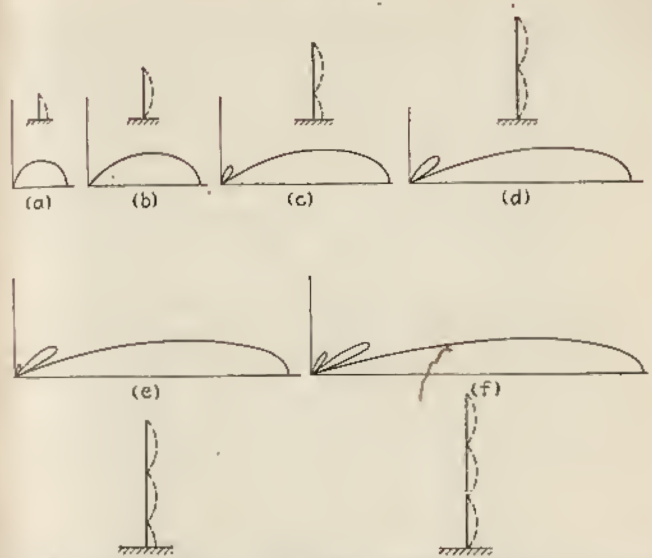


FIG. 16.—Polar patterns of the field strength distribution for vertical antenna over perfect earth when the currents in successive dipole sections are cophased. Note that minor lobes never exceed the horizontal tangent to the major lobe.

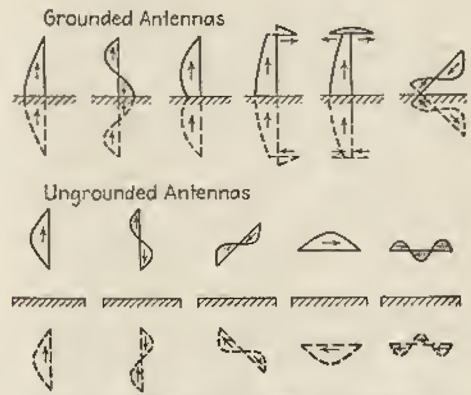


FIG. 17.—Electrical images of antennas.[1]

[1] TERMAN, F. E., "Radio Engineering," McGraw-Hill Book Company, Inc.

13. Calculation of Vertical Radiation Patterns from Vertical Antennas above Perfectly Conducting Ground. A vertical doublet of infinitesimal length in free space produces a field which has a magnitude proportional to the sine of the angle θ from the doublet axis. In a vertical quarter-wave antenna with sinusoidal current distribution, integration of the influences of all the doublets throughout its length gives a distribution

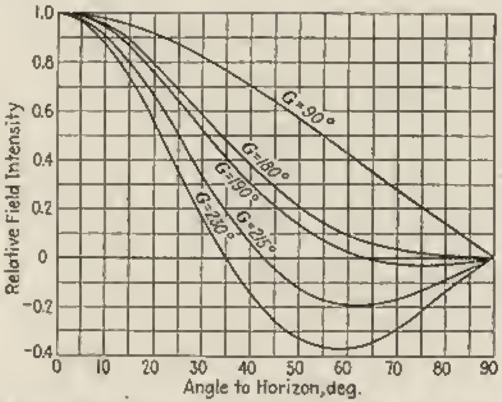


FIG. 18.—Vertical field strength distribution for vertical antennas over perfectly conducting earth, for various electrical heights G.[1]

only slightly flattened with respect to that for a doublet, and it has the equation

$$f(\theta) = \frac{\cos (90^\circ \cos \theta)}{\sin \theta}$$

For a vertical dipole above perfect earth,

$$f(\theta) = \frac{\cos (90^\circ \cos \theta) \cos (H \cos \theta)}{\sin \theta}$$

where H is the height of the current antinode in degrees above the reflecting surface.

For any vertical grounded antenna with sinusoidal current distribution having a total electrical height G,

$$f(\theta) = \frac{\cos (G \cos \theta) - \cos G}{\sin \theta \, (1 - \cos G)}$$

14. Directivity Diagrams in the Horizontal Plane for Two Identical Vertical Radiators. Two radiators, synchronously excited with equal currents, produce interference patterns which vary with the separation of the radiators and the relative phase of the radiator currents. Furthermore, owing to the influence of the mutual impedance upon the two radiators, the field intensities obtained will at some points exceed, and at other points be less than, for the same power input to one radiator alone.

[1] BROWN, G. H., *loc. cit.*

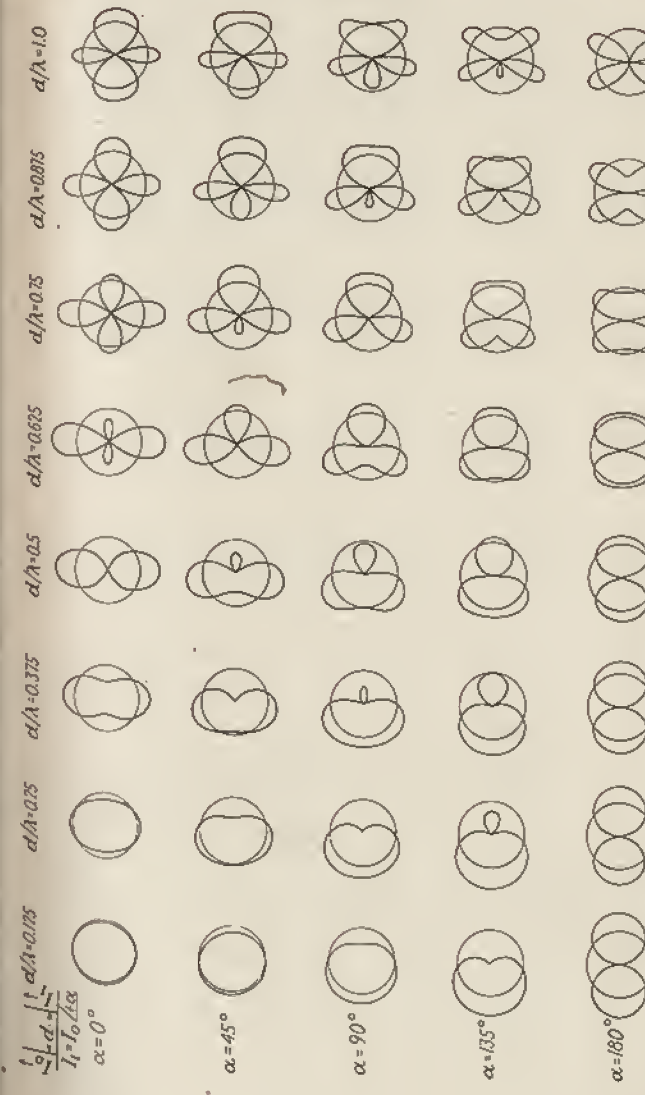


FIG. 19.—Horizontal patterns for two radiators fed with equal magnitude currents taking mutual impedance into account. Circles indicate relative fields from one radiator alone with same power input.[1]

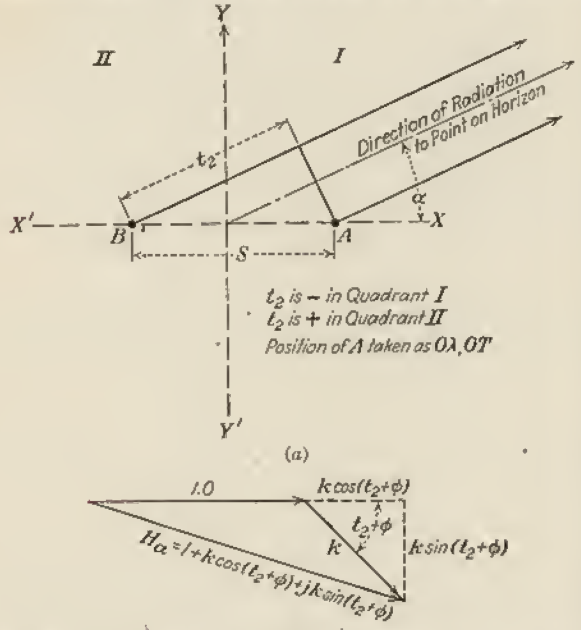
[1] BROWN, G. H., Directional Antenna, *loc. cit.*

The directivity patterns obtained for spacings and phasings over all values of practical importance, drawn to a scale which shows the relative field intensity obtained with the same power in a single radiator, are of great importance in antenna calculations. Figure 19 shows such directivity patterns.

15. Calculation of Directivity Patterns. For two radiators having equal currents, in terms of parameters $A\lambda$ and BT, where T = time phase

$$f(\alpha) \doteq [\cos(\pi A \cos d + \pi B)]$$

For identical radiators with unequal currents, the resulting pattern in terms



(b)

FIG. 20a and b.—Data used in calculating directivity diagrams.

of relative field intensities can be obtained from Fig. 20 and from

$$f(\alpha) = 1 + k\cos(t_2 + \phi) + jk\sin(t_2 + \phi)$$

where α = angle from the axis of the radiators
k = current ratio for identical radiators or the horizontal field ratio for dissimilar radiators
t_2 = angular difference in path length which $= -S\cos\alpha$
S = radiator separation in electrical degrees
ϕ = phase difference of I_B with respect to I_A.

All such patterns are symmetrical with respect to the axis of the array.

Where three or more radiators are used with arbitrary spacings, phasings, and current ratios, the combinations become so great that formulation must usually be worked out for the particular case at hand.

16. Three-dimensional Radiation Patterns for Array of Two Identical Radiators. Frequently two-element arrays are employed for directional transmission, in which case it becomes necessary to know the space characteristics. A rough preliminary investigation of the three-dimensional distribution of field intensities in the horizontal plane, the vertical plane through the radiators, and the vertical plane broadside to the radiators may be quickly made in the following manner: From the patterns of Fig. 19 select the horizontal pattern corresponding to the separation and phase difference to be used. From this pattern the

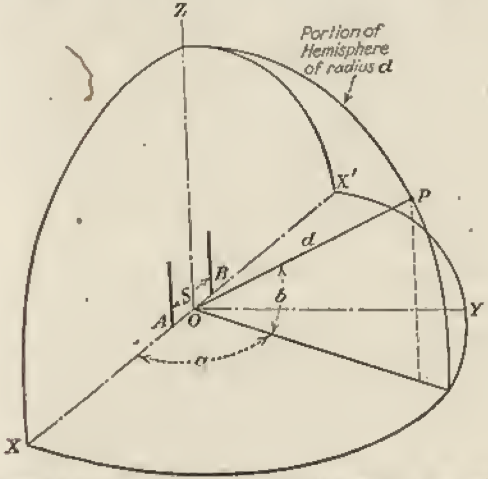


FIG. 21.—Geometry for calculating space characteristic of two-element array.

vertical-plane distribution through the radiators may be found by multiplying the upper half of the pattern (which lies above the X-axis) by the polar characteristic in the vertical plane for *one* of the antennas. When the radiators are grounded quarter-wave elements, it is merely necessary to multiply the radius vector at any angle by the cosine of the angle. (Perform this for the entire 180 deg.) Familiarity with this method enables one to estimate the vertical-plane distribution immediately by inspection of the horizontal pattern. For the vertical plane broadside to the array, the distribution is the same as for a single antenna. Thus we have the following:

1. The distribution in the plane through the radiators is the same as for the upper half of the horizontal pattern multiplied by the characteristic for one radiator.
2. The broadside plane distribution is the same as the vertical pattern for one radiator.

3. Where there is suppression of radiation in line with the radiators in the horizontal pattern, there will be one or more lobes of high-angle radiation in that direction in the vertical pattern. The shape of the high-angle lobes will depend upon the vertical-plane pattern for one antenna.

4. Where suppression of radiation occurs broadside to the array in the horizontal pattern, there will be proportionate suppression at all vertical angles.

5. Where there is a maximum of radiation in line with the radiators in the horizontal plane, there will be a flattening of the pattern over that of a single radiator, in the vertical plane.

6. When the ratio of currents in the two radiators is other than unity, the angles of maxima and minima occur at the same place, but the nulls and peaks are less pronounced. As the current ratio approaches zero, the pattern approaches that of a single antenna.

To determine the entire space characteristic, it is best first to calculate the horizontal pattern as described above and then to calculate a series of vertical-plane patterns at various angles to the array axis. As the horizontal pattern is symmetrical with respect to the axis x–x', so the space characteristic is symmetrical with respect to the X–Z plane through the radiators. Furthermore, the space characteristic in the half space above the X–Y plane is symmetrical with that in the half space below the X–Y plane. The geometry for use with the following formula for the space characteristic in the upper half space, for grounded antennas and perfectly conducting ground, is given in Fig. 21.

$$H_{ab} = [1 + k \cos (t_4 + \phi) + jk \sin (t_4 + \phi)].$$
$$[1 + \cos t_1 + j \sin t_1] \cdot \left[\frac{\cos (90° \sin b)}{\cos b}\right]$$

(Multiply only the scalar values of each factor.)

In this equation H = field intensity in arbitrary units at an angle (a) measured horizontally with respect to the line through the radiators and the angle (b) above the horizon

k = current ratio (equal to or less than unity)

t_4 = total phase difference, in degrees, between radiations from A and B in the direction (a) (b)

$$t_4 = |(S' \cos a \cos b)|$$

and (s' is the spacing between radiators in electrical degrees)

t_1 = $-2h \sin b$, where h is the height of the current antinode above ground, in electrical degrees

ϕ = initial phase difference between I_B with respect to I_A

NOTE. When the radiators are exact quarter-wave elements, the second factor becomes constant and can be ignored. When the radiators are considerably less than one-quarter wave length in height, the second factor can be ignored and the third factor simplified to cos b.

The above equation is restricted to those cases where the physical length of the radiators does not exceed one-half wave length, though the height of radiators above ground is not restricted.

To obtain the vertical distribution pattern for one radiator, ignore the first factor and use only the second and third.

17. General Solution for the Space Characteristics for Any Array of Antennas Disposed in Any Manner in Three Dimensions. In view of the special nature of the general solution for extended antenna arrays, we shall not attempt to condense this important subject in this work but shall merely refer the interested reader to the references below.[1] Extended antenna arrays are extensively applied in h-f directive transmission and are of great engineering importance at the present day.

BROADCAST ANTENNAS, 550 TO 2,000 KC

18. Prevailing Types of Broadcast Antennas. The old-fashioned forms of antenna construction, familiar for many years, are still largely used but deserve no particular attention from present-day engineers because they are rapidly being replaced by more efficient radiators.

Broadcast antennas may be classified as follows:

1. The high vertical single-wire antenna, suspended from a triatic between self-supporting steel towers (widely spaced), and having a fundamental frequency lower than the operating frequency.
2. The high single-wire T antenna, being similar to A, but with a relatively short T flat top, and operating above its fundamental frequency.
3. The guyed cantilever steel tower, having a height somewhat greater than one-half wave length, the tower itself forming the antenna conductor.
4. The self-supporting (slender) steel tower, having a height from one-quarter to more than one-half wave length, the tower itself being the antenna conductor.
5. The single-wire vertical antenna suspended along the axis of a self-supporting treated-wood tower, and operating, in general, at a frequency much higher than its fundamental.[2]
6. Directive antenna arrays of two or more vertical elements, designed either to get more advantageous coverage where population distribution is irregular, or to reduce interference in the directions of other stations that may be on the same channel.

19. Progress in Antenna Improvements. The low multiwire with a large L or T flat top was the ordinary form of radio antenna for many years and was used until recently for broadcasting. Since 1927 there has been a rapid development in broadcast antennas, and their form has been greatly modified.[3] The results sought are reduction of high-angle radiation for the reduction of fading and greater efficiency giving larger service areas for a given power input. The present commercial importance of broadcasting justifies a considerable investment in an improved radiator.

In terms of their characteristic current distributions and relative linear dimensions several types of broadcast antennas are represented in Fig. 22. As the height of the antenna increases, the position of the current antinode is raised above ground, which causes the high-angle radiation to decrease and the low-angle radiation to increase. The effect of antenna height (in terms of electrical degrees) on the relative distribution of field intensity for five different antenna lengths is demonstrated in Fig. 18, where these data are plotted in rectangular instead of the more usual polar

[1] FOSTER, R. M., Directive Diagrams for Antenna Arrays, *Bell System Tech. Jour.*, April, 1926; SIEGEL, E., and J. LABUS, Feldverteilung und Energieemission von Richtantennen, *Hochfrequenz Technik und Elektroakustik*, Band 38, Heft 6, 1932; SOUTHWORTH, G. E., Certain Factors Affecting the Gain of Directive Antennas, *Proc. I.R.E.*, September, 1930; BAEUMLER, M., K. KRÜGER, H. PENDL, and W. PFITZER, Radiation Measurements of a Short-wave Directive Antenna at the Nauen High-power Radio Station, *Proc. I.R.E.*, May, 1931.

[2] EPPEN, F., and A. GOTHE, Über die Schwundvermindernde Antenne des Rundfunksenders Breslau, *E.N.T.*, Band 10, Heft 4, 1933.

[3] BROWN, G. H., and H. E. GIHRING, General Considerations of Tower Antennas for Broadcast Use, *Proc. I.R.E.*, April, 1935; CHAMBERLAIN, A. B., and W. B. LODGE, The Broadcast Antenna, *Proc. Radio Club Amer.*, November, 1934; LAPORT, E. A., Improved Efficiency with Tower Antennas, *Electronics*, August, 1934.

coordinates. The portions of the curves shown as negative field intensities indicate radiations in a secondary (high-angle) lobe in which the direction of the electric field is reversed. For a straight vertical antenna, it is seen that, when the height of the current antinode exceeds one-quarter wave length above ground, the high-angle lobe forms rapidly and soon assumes a value unsatisfactory for broadcasting use because of fading. For this type of antenna the 190-deg. length is about the maximum permissible.

Since the previous edition of this handbook, tower radiators have been thoroughly proved in, but this involved a change from the original cantilever guyed structures, through the trial of broad-based self-supporting towers, and finally to the uniform cross-section guyed or self-supporting radiators. The latter have reached the stage of optimum electrical performance, reliable mechanical design, and moderate cost. During this period most stations have constructed modern radiators and have retired the supported-wire antennas, and many have installed directive arrays of two to three radiators for minimizing interference and better covering of local areas. Where airline routes have limited antenna

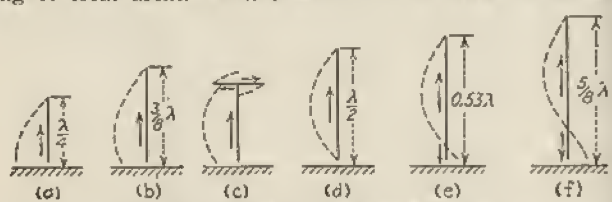


FIG. 22.—Types of broadcast antennas.

heights, top-loaded and sectionalized antennas have been employed to attain high efficiency and fading reduction.[1]

In the United States and Canada antenna heights have been specified by regulations, the heights being worked out in harmony with the general objectives of allocation and efficient utilization of facilities.

20. Ground Systems for Antennas. The importance of the ground terminal for a radiating system cannot be overemphasized. If there existed such a thing as a perfectly conducting earth, any sort of a firm connection to the earth would suffice for a terminal. Soils, and even salt-water marsh, at best are poor conductors at radio frequencies. The ground system used with an antenna must make the best possible contact with existing ground substances as found at a station site. A few years ago it was thought that a ground system had only to extend outward as far as the limits of the induction field of the antenna. The major function of the ground system as a reflecting surface for the down-coming waves from the antenna is now generally recognized, and for this purpose a ground system must extend outward for a considerable distance. There

[1] BROWN, G. H., A Critical Study of the Characteristics of Broadcast Antennas as Affected by Current Distribution, *Proc. I.R.E.*, January, 1936; BROWN and LEITCH, The Fading Characteristics of the Top-loaded WCAU Antenna, *Proc. I.R.E.*, May, 1937; BROWN, G. H., A Consideration of the Radio-frequency Voltages Encountered by the Insulating Material of Broadcast Tower Antennas, *Proc. I.R.E.*, September, 1939; MORRISON and SMITH, The Shunt-Excited Antenna, *Proc. I.R.E.*, June, 1937; GIHRING, R. F., Notes on Broadcast Antenna Developments, *RCA Rev.*, April, 1937; FITCH and DUTTERA, Measurement of Broadcast Coverage and Antenna Performance, *RCA Rev.*, April and July, 1938, and April, 1939.

have been many temporary theories and practices regarding the configuration of the conductors in the ground system, but there is now, broadly speaking, a convergence of preference for the radial system with an effective earth termination for each wire. Recent studies have further proved the need for a large number of very long radials. The more nearly a system of wires approaches a continuous metallic sheet of great extent, the better it is as a ground system.

The work of Brown[1] on the theoretical and experimental study of ground systems has established definite criteria for their design. Broadly summarized, a radial system of 120 radial wires approximately one-half wave length long approaches very nearly the characteristics of an ideal ground terminal, as shown by Figs. 23 and 24.

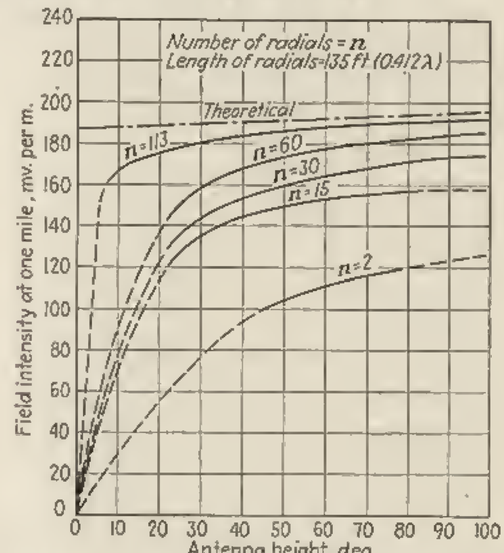


FIG. 23.—Variation of field intensity with antenna height and number of 0.412 wave length radials.

21. The Counterpoise. Where a buried ground system cannot be employed, a counterpoise is frequently required as a high-capacity ground terminal. In general the same considerations which apply to radial ground systems apply also to counterpoises. Where extremely high electric fields exist near the base of a radiator, a small counterpoise will help to reduce the potential gradients in imperfect dielectrics, such as soil or wood, and thus decrease losses. Where ground systems of adequate length are impractical, such as in l-f radio range antenna systems, relatively small radial counterpoises provide a stable ground terminal and consequently stable radio range courses. Roof antenna

[1] BROWN, G. H., Ground Systems as a Factor in Antenna Efficiency, *Proc. I.R.E.*, June, 1937.

installations which employ counterpoises of adequate area and whi clear other small structures usually found on the roofs of such building have shown good performance.

For certain u-h-f antennas elevated counterpoises employing a surfa of 4-in. mesh wire screen on a metallic framework have been extensive used. Typical applications are for fan-marker, cone-of-silence marke and u-h-f four-course radio range antennas for use on the airways.

22. Antenna Measurements. Antenna measurements are theore cally simple, but skill and experience are required, together with goo instruments, to attain accurate results. It is for this reason that t

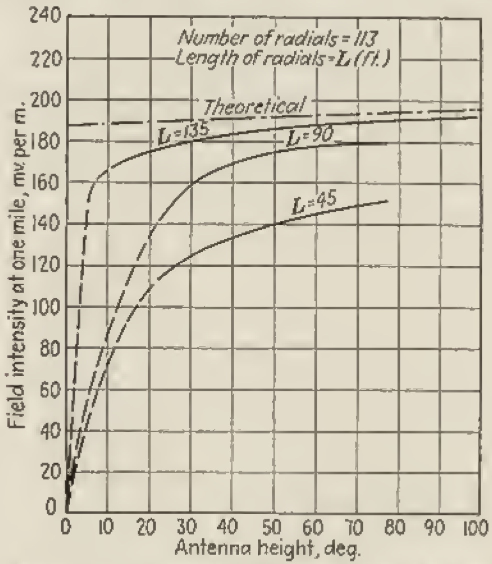


Fig. 24.—Variation of field intensity with antenna height and length of radia in a 113-wire ground system.

FCC specifies that such measurements, to be submitted to it for approva must be made by a qualified person with approved instruments of know accuracy. The practical difficulties of measurement increase with th frequency.

Resistance. Approved methods of measuring antenna resistance ar described in Radio Instruments and Measurements, *Bull.* 74, of the U. S Bureau of Standards. (This may be obtained from the Superintende of Documents, Government Printing Office, Washington, D. C.) Fo low-impedance antennas ordinary precautions may suffice, but for hig impedance antennas of the order of several hundred ohms extreme ca and occasionally special methods are employed for reliable results. I is well to repeat measurements two or three times with new setups befor certifying the accuracy of the data.

Radio-frequency bridges are now available and, with proper manipulation, lend themselves well to antenna measurements. Resistance and reactance can both be measured directly. For low-impedance antennas, straight bridge methods are useful. For moderately high impedances a substitution method, where the unknown impedance is connected in parallel with one arm of a balanced bridge, is required. One serious drawback to bridge measurements in antenna circuits is the susceptibility to inaccurate balances when static levels are high and when there is interference from other stations in the locality. In the latter case power methods of measurement are essential.

A simple method for measuring antenna resistance and reactance of a very high-impedance antenna is the following: Using an ordinary wavemeter of the precision-absorption type equipped with a thermoammeter and a calibrated condenser, adjust the wavemeter to the desired frequency and bring the oscillator into tune at this same frequency. Couple the wavemeter to the oscillator until full scale deflection of the ammeter results. One side of the wavemeter (the shield side) should be grounded to the regular antenna ground system. Note the setting of the variable condenser and the exact meter reading at resonance. Then connect the antenna downlead to the ungrounded side of the wavemeter as shown in Fig. 25 and retune the wavemeter for maximum current. Note the condenser setting and the new meter reading for this condition.

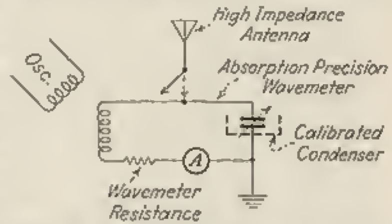


Fig. 25.—Antenna measuring circuit.

The lower the antenna resistance, the lower will be the ammeter reading with the antenna attached. Also, if the antenna has an inductive reactance at the particular frequency, the capacitance of the wavemeter will have to be increased to restore resonance, and vice versa.

By substitution, known standard values of resistance and reactance in series are connected in parallel with the wavemeter to reproduce the same series of adjustments and readings as observed, first with the wavemeter alone and then with antenna attached to it. The resistance and reactance values which reproduce the antenna values precisely are equal to those of the antenna.

The precautions to be observed in using this method are as follows: The oscillator must be of sufficient power output and regulation as to be unaffected by the presence of the wavemeter; the standards of impedance used for substitution must be essentially free from stray capacitance when arranged for use, for small values of stray capacitance can seriously disturb the accuracy of the results; the readings and adjustments before and after adding the shunt impedance must exactly duplicate those observed in the process of measuring the antenna.

If the wavemeter resistance is accurately known, the unknown antenna impedance, in terms of resistance and reactance, can be calculated.

It is customary, in view of certain difficulties in making antenna measurements, to ensure greater accuracy by making a series of such measurements over a considerable range of frequencies. Individual errors are averaged out by drawing a smooth curve through the values as plotted out in graphical form. Antenna resistance and reactance measurements over a wide band of frequencies are often invaluable in analyzing the action of an antenna, as well as for predetermining the proper circuit constants to be used for tuning it or matching its impedance to a given transmission line.

Reactance. When not measured directly with an r-f impedance bridge, antenna reactance can be measured by resonating the antenna at the desired frequency, using a calibrated inductor or capacitor. At resonance, the antenna reactance is equal and opposite in sign to that of the tuning device.

Fundamental Frequency. Connect the antenna directly to ground through an r-f current instrument of adequate sensitivity, and couple a variable-frequency oscillator lightly to the system by proximity only. Search for the lowest frequency at which resonance is indicated by a maximum current.

Direct Power Input Measurement. This measurement is dependent upon an accurate measurement of antenna resistance and the use of a

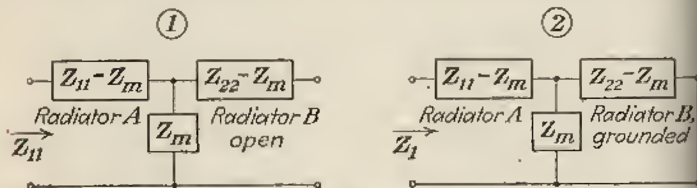


FIG. 26.—Equivalent circuit of two radiators coupled by radiation.

ammeter of suitable accuracy located at the point where resistance was measured. The power is the product of the antenna resistance and the square of the entering current. In all probability direct-reading wattmeters will be commercially available during 1941.

Mutual Impedance. Mutual impedance can be measured only indirectly. Where conditions permit, the method is to measure the self-impedance Z_{11} of one radiator (assuming both radiators to be identical) with the second radiator first open-circuited, and again Z_1 when the latter has been grounded. From these two impedance measurements the mutual impedance is calculated. Where more than two radiators are employed, such a measurement is required for every combination of radiators taken two at a time, with the other radiators open-circuited so that they do not affect the pair under measurement by reradiation.

The equivalent circuit of two radiators coupled by radiation is shown in Fig. 26. From this figure

$$Z_m = \sqrt{Z_{11}^2 - Z_1 Z_{11}}$$

The accuracy of this method is enhanced by tuning out the self-reactance of each radiator before measuring Z_1 in terms of $R_1 + jX_1$. When this is done

$$Z_1 = R_1 + jX_1 = \frac{R_{11}^2 - R_m^2 + X_m^2}{R_{11}} - j\frac{2R_mX_m}{R_{11}}$$

With R_{11}, R_1, and X_1 measured, the two terms above can be solved simultaneously for R_m and X_m.

23. Vertical Radiator Self-impedances. Resistance and reactance measurements on a 400-ft. vertical uniform cross-section tower having a base insulator capacitance of 30 $\mu\mu$f are shown in Fig. 27. The resistance and reactance of a slender tubular steel mast are shown in Fig. 28.

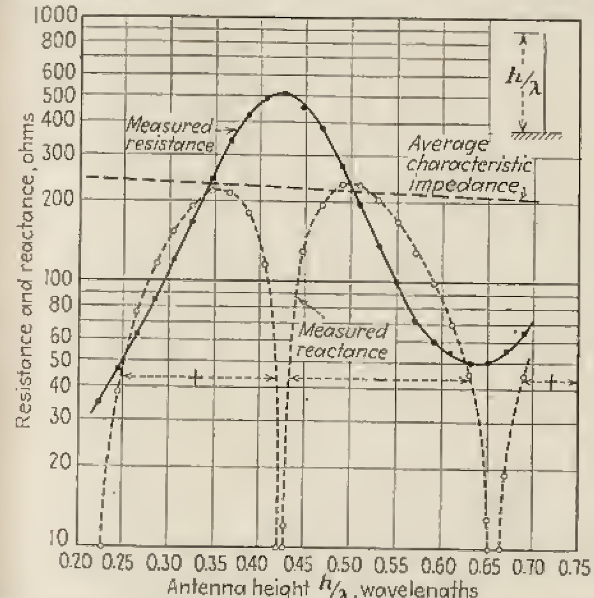


FIG. 27.—Resistance, reactance, and characteristic impedance for WWJ uniform cross-section 400-ft. tower.

24. Coupling and Feeder Circuits Used for Broadcasting. Antennas are either fed directly from the transmitter, or at a distance from the transmitter by using some form of radio-frequency transmission line. The following types of lines are used, listed roughly in order of their numbers in service:

Type	Approximate Z_0
4-wire open line with 2 opposite wires grounded	235
Concentric tubular lines	70
2-wire open balanced line	600
3-wire open balanced line (middle wire grounded)	600
4-wire balanced line	315
1-wire open line with ground return	500

Transmission lines require equipment, suitably adjusted, to be capable of transforming the antenna impedance to the characteristic impedance line. A line, terminated in its characteristic impedance Z_0, provides an unidirectional flow of power from the transmitter to the antenna without the losses due to reflections of energy in the system. For single-end transmission lines, simple T, Pi, or, more usually, L networks of reactive elements are used to match the antenna impedance to the line impedan Where balanced lines are used, the balanced to single-end impedan matching transformation is usually accomplished by inductivity coupl circuits. In any case the adjustment of the terminal network for te mination of the line is another case where skill and proper instrume are needed.

Figure 29 shows the simplest coupling circuits, their theory for feed single radiators, and where the phase shift through the network is immat rial. In directive antennas where a given impedance match must b made with a specified phase shift, a three-element network is requir

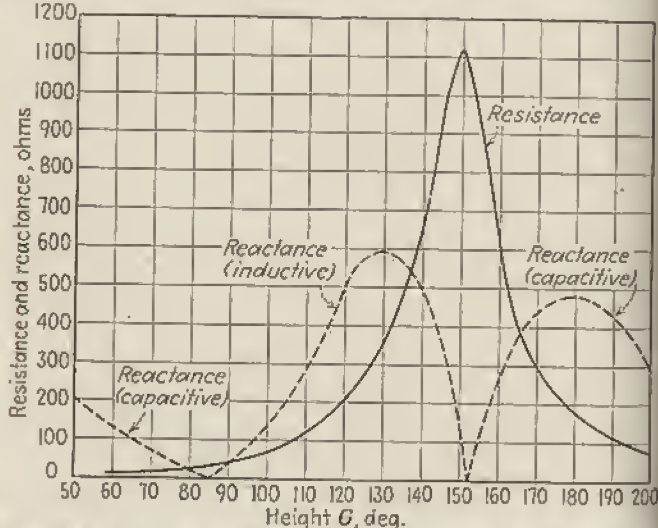


FIG. 28.—Measured resistance and reactance of 150-ft. vertical tubular ma antenna having a base insulator capacitance of 85 $\mu\mu$f.

The values of the circuit elements are calculated after the anten impedance and the characteristic impedance of the line have been mea ured. The required reactances are then set to specified values, and min corrective adjustments are made to obtain a perfect match. A r impedance bridge is very convenient for this purpose.

For the balanced line, an impedance bridge, being a grounded dev for one of the unknown terminals, is less useful. The following meth of adjusting terminal impedances for balanced lines is simple, accura and rapid and requires a minimum of equipment.

Calculate or measure Z_0. Calculate the line current for any giv power from $I_0 = \sqrt{W/Z_0}$, and the voltage across the termination und this same condition from $E_0 = \sqrt{WZ_0}$. If a tank-circuit terminati is used, with inductive coupling to the antenna circuit, choose a val of capacitance across the line which has a reactance of the order of ha

that of the value of Z_0 (arbitrary) at the operating frequency. This capacitance must be accurately known so that its reactance can be determined and used to find the proper current through this condenser at resonance when the termination is correct. $I_c = E_0/X_c$.

Knowing now the values of I_0 and I_c, we can take their ratio, I_c/I_0. Now, by inserting matched ammeters in series with the line at the entrance to the termination and in series with the capacitance leg of the

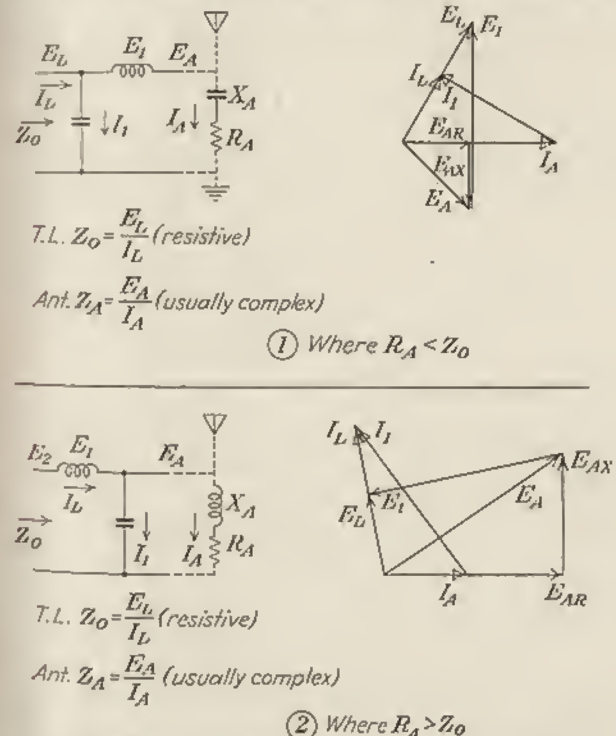


FIG. 29.—Coupling circuits and vector relations.

tank circuit, we know that, when the proper termination has been reached, the ratio I_c/I_0, previously calculated, must prevail. To obtain a termination that has unity power factor, the tank circuit must be close to an antiresonant adjustment. After inserting the ammeters in circuit as shown in Fig. 30, apply power (the amount is unimportant at this stage, because the ratio I_c/I_0 is independent of power) and adjust the primary inductance until the line current I_0 is minimum, at which time the values I_0 and I_c are observed and the ratio calculated. If the ratio is too high, the coupling is too loose or the impedance of the antenna circuit is too high. If the former, add more coupling turns or otherwise tigh coupling between tank and antenna circuits; if the latter, decrease t reactance of the antenna circuit by increasing inductance for a capacit antenna or decreasing capacitance for an inductive antenna (both ass ing that the antenna is not resonant). Retune the tank circuit a again measure the ratio of currents. If the ratio of currents is too h the opposite procedure is followed, *i.e.*, the coupling is reduced or t antenna impedance increased, or both. By discrete steps and by car adjustments, the exact ratio I_r/I_0 is quick attained. The final test of correctness is measure simultaneously the current in bo ends of the line. These currents should equal.

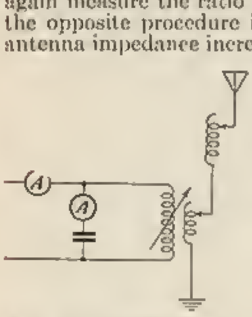


FIG. 30.—Antenna adjustment.

25. Adjustment of a Directive Anten Array for Broadcasting. The proper design a directive array includes the calculation the desired radiation pattern in terms of re izable field intensities, which in turn depe upon the amplitude and phase relations b tween the radiator currents, the self- mutual impedances, and the power input. T input impedance to each radiator must b be calculated from the operating condit of the array, which permits the division of power between the radiat to be found.

The next step is the design of the transmission lines and their coupl networks which effect the energy transfer at each radiator and wh match the radiator input impedance to that of the line with the ex phase shift which, with that due to the time of propagation over the li will bring each radiator current to its precise amplitude and phase. Th is done as a preliminary step only.

The adjustment of such a system to realize a specified performa requires that the above sequence of conditions be reproduced physical Thus the first step is to measure the self- and mutual impedances after t radiators are constructed from the design data and to recalculate t input impedances to the radiators from these data. Then the ex coupling networks are synthesized and constructed,[1] and adjustme made to the previously calculated values instrumentally, with utm accuracy. The performance of the system is then verified by measur the currents and their phases[2] and finally by measurement of the radiat which consists usually of a plot of a number of field-intensity meas ments made on the mile circle. Here is a fascinating problem requir the finest technique of theoretical calculation and physical measureme

MARINE TRANSMITTING ANTENNAS

26. Limitations to Shipboard Antennas. There has been little cha in the design and construction of shipboard antennas for the reason th there is little choice available. The limited space and the presence stacks, derricks, etc., place severe limitations on the mechanical arran ment of the antenna. For that reason, shipboard antennas have been but slightly modified in many years. The outstanding change is the gradual abandonment of multiwire forms for the single wire.

[1] LAPORT, E. A., Graphical Network Synthesis, *Broadcast News*, January, June, 1 BROWN, G. H., Directive Antennas, *Proc. I.R.E.*, January, 1937.

[2] MORRISON, J. F., Simple Method of Observing the Current Amplitude and Ph Relations in Antenna Arrays, *Proc. I.R.E.*, October, 1937.

Large vessels in the passenger business now have several transmitters in their radio rooms. For long-wave ship traffic at moderately high powers the ship's antenna has a very short electrical length, which gives a nearly uniform distribution of voltage throughout its length. Such antennas must be insulated equally at all points.

For intermediate-wave operation, ship antennas have fairly good characteristics and efficiency. The antennas on the larger ships have fundamental wave lengths somewhere near the intermediate marine band, so that they operate essentially as quarter-wave systems. For h-f telegraph operation it is now quite the usual practice to use the main ship antenna, operating it at or near one of its harmonics.

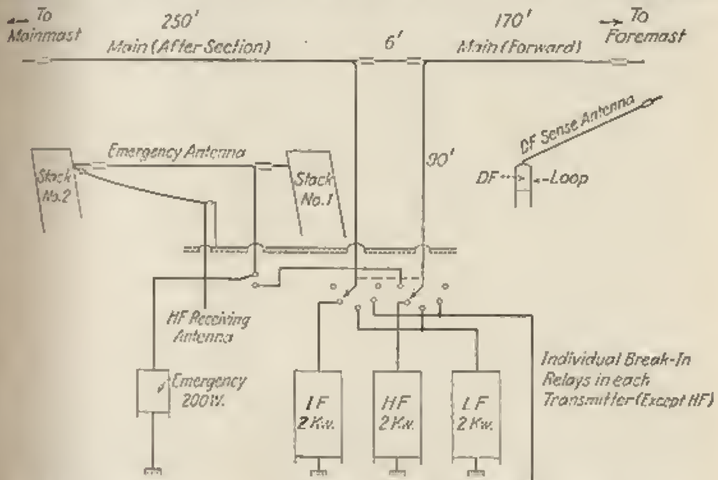


FIG. 31.—Marine antennas (SS. "Washington" and SS. "Manhattan").

A system used on some of the best-known American ships is that shown in Fig. 31. This not only permits utilizing the main antenna for all the marine frequencies but provides a convenient means for simultaneously operating the short-wave and intermediate-wave transmitters. For long-wave operation both halves of the antenna are connected in parallel.

Ships having commercial telephone services usually employ separate half-wave dipole antennas fed by terminated transmission lines. These are mounted anywhere on the ship where there is space, frequently using the stacks for support, and often suspended by means of insulator strings from the triatic of the main antenna. Where a half-wave dipole antenna is used, it is necessary to have a different antenna for each operating frequency.

27. Antenna Characteristics. Shipboard-antenna characteristics vary over extremely wide ranges because of differences in mechanical forms and dimensions, the effects of other conductors in the field on the antenna, the nearness of stacks, etc. For example, antenna resistances in the intermediate- and long-wave bands range from 3 to 10 ohms. Static capacitances range from 400 to 1,200 $\mu\mu$f. Fundamental wave lengths range from 200 to 500 meters. It is difficult to specify typical antenna characteristics beyond these figures.

28. Construction of Ship Antennas. The essential mechanical requirements for an antenna design are extreme ruggedness and reliability under all the severe weather conditions met at sea. Heavy phosphor-bronze stranded cable is employed for the triatic, preferably for the entire antenna. The use of an inverted L or a T is principally determined by the layout of the ship and the location of the radio room with respect to the antenna. Regular ship-rigging construction is employed. The antenna must be easily lowered and raised. On some ships the antenna must be lowered to permit operating the derricks.

The essential electrical requirements are as follows:

1. A maximum of antenna size for a given available space.
2. Maximum possible clearance of ship's rigging, bridge, stacks, etc., to reduce losses by induction.
3. Liberal high-voltage insulation throughout the length of the antenna including the deck insulator.
4. Avoidance of sharp points, broken strands, or V-bends which would become corona discharge points.
5. Positive firm electrical connections between different sections of the antenna and at the entrance bushings.
6. The use of a single-wire system.
7. The avoidance of the use of hemp guys and stays at points of high potential gradients near the antenna wire and insulators where rapid deterioration due to burning would result.

29. Shipboard Receiving Antennas. A separate wire receiving antenna is now common practice on shipboard for short-wave reception. For intermediate- and long-wave reception the main transmitting antenna is quite generally used, connection of the receiver to the antenna being made through a break-in keying relay when the transmitter is not actually transmitting, and to ground when the transmitter is exciting the antenna.

The sense antenna used in conjunction with the direction finder is a separate wire and used only for that purpose.

Broadcast receiving antennas may occupy any remaining space available on the ship.

NON-DIRECTIVE ANTENNAS FOR H-F TRANSMISSION

30. Types of Antennas in Current Use. Antennas for the circular diffusion of energy at high frequencies approach very nearly the fundamental ideal forms. For a given form of antenna for a given performance the mechanical size is proportional to the transmitting wave length, and when this becomes comparatively short, the mechanical aspects of the problem become very simple.

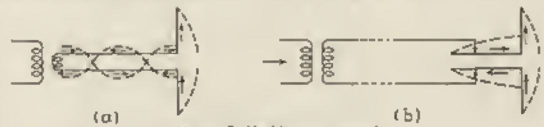
1. A fundamental and widely applied form of h-f transmitting antenna is the half-wave dipole. It can be employed in a variety of ways by changing its orientation in space and its position with respect to ground. When located in hypothetical free space, its electrical values are constant; but, when located within a few wave lengths of real earth, as in practice, they are influenced by orientation and position.

When placed vertically with respect to the surface of the earth, a half-wave dipole transmits vertically polarized fields in every direction. When mounted horizontally, the radiated field is horizontally polarized in any direction perpendicular to the antenna wire, while it is vertically polarized in the directions of the wire. In intermediate directions the fields will have both vertically and horizontally polarized components, a state called *elliptical polarization.* These conditions have a bearing upon the propagation characteristics of radiation in different directions.

2. A second fundamental type of h-f transmitting antenna is a straight wire operated at one of its harmonics. Where one antenna is used for both l-f and h-f transmission, as on shipboard, we have a case where, at high frequencies, the antenna may be several times the length of a half wave. If such an antenna is vertical, the radiation is uniform in all horizontal directions but of rapidly varying intensity in the vertical plane. The general characteristics were discussed and described in a preceding section.

3. A third important type of non-directional antenna for h-f transmission is the vertical wire with the current in adjacent dipole sections cophased. Instead of the current-distribution characteristic of the antenna operating at a natural harmonic, where the current in each successive half-wave section is reversed in direction, this antenna has currents all flowing in the same direction. This is achieved by using antiresonant coils or networks at each current node in the system except the extreme ends. A vertical antenna of this type produces a high degree of radiation concentration at angles close to the horizontal, a characteristic of great value in efficient long-distance transmission.

31. Feed Methods for H-f Antennas. 1. *Pure Current Feed.* A balanced current-feed system for energizing a divided half-wave dipole



NOTE: Dotted lines represent current distribution on antenna and feed lines.

FIG. 32.—Current feed for half-wave dipole. (*a*) For balanced line; (*b*) for balanced terminated line.

is shown in Fig. 32 in two forms, where (*a*) is for the use of a balanced oscillating transmission line throughout and (*b*) is for a balanced terminated transmission-line system, the termination being made by means of connections to proper points on a balanced quarter-wave transformation section.

2. *Pure Voltage-feed System.* Three forms of the pure voltage-feed system are shown in Fig. 33, where (*a*) is the balanced system using resonant line feeder, (*b*) the unbalanced system using resonant line feeder, and (*c*) a balanced system with balanced terminated transmission line, the antenna impedance being matched to that of the line by means of a resonant line transformer. In the case (*b*) the feed line can be a concentric tubular system, the antenna being connected to the inner conductor.

3. *Voltage Feed from Terminated Concentric Transmission Line.* A method of voltage feeding a half-wave dipole from a terminated concentric transmission line is represented schematically in Fig. 34. Here the concentric line is made to have a characteristic impedance equal to the radiation resistance of the antenna at the current antinode (73.2

ohms, if the antenna is several wave lengths above ground). A wire one-quarter wave length long, projects beyond the end of the outer conductor parallel and close to the extension of the inner conductor which continues on to become the antenna. If there is essentially zero-radiation resistance due to the opposed quarter-wave sections, these act as a transformer to transfer the radiation resistance at the current antinode of the

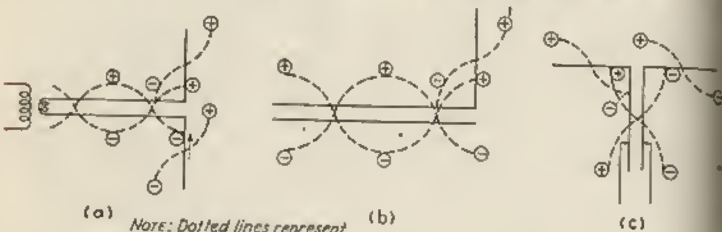


FIG. 33.—Voltage-feed systems (see text p. 663).

antenna to across the end of the concentric line, thus effectively terminating the latter.

4. *Terminated Transmission Feed.* At high frequencies it is possible to obtain a satisfactory termination of an open-balanced transmission line by connecting the extremities of the line directly to the antenna wire as shown in Fig. 35. The connections are made symmetrically to those points on the antenna which show an impedance as nearly as possible like that of the characteristic impedance of the transmission line. In spreading the wires of the feed line to bridge the proper impedance in the antenna, there results a change in the characteristic impedance of the line in that portion which makes a perfect termination

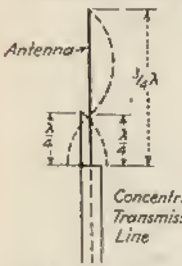


FIG. 34.—Use of quarter-wave wire as matching transformer.

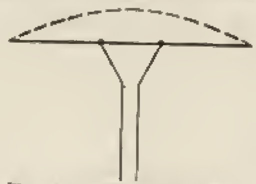


FIG. 35.—Method of connecting line to radiator.

theoretically impossible, though satisfactory practical adjustments are obtained. For optimum line balance, exact symmetry of connection is required. The location of the connection points is critical. The adjustment is dependent upon the location of the antenna with respect to ground and other conductors, the effects of insulator caps, etc.

5. *Other Methods of Terminating Open-wire Transmission Lines in Antennas by Means of Networks.* There remain several methods for

terminating a balanced transmission line in an antenna by means of networks of inductance and capacitance. The antenna has a certain complex impedance when viewed from any given feed point. To match this impedance to the line impedance; a suitable transforming network is designed.

DIRECTIVE H-F TRANSMITTING ANTENNAS

In this branch of engineering we find the antenna art at its best. Unhampered by serious mechanical obstacles, full advantage may be taken of electrically long radiators, and extended arrays of many such radiators, for obtaining a very high degree of radiation concentration in a desired direction. Present-day h-f directive antennas project a beam of electromagnetic energy which is analogous in fact to the beam of a searchlight.

Out of the unlimited variety of possible forms of antenna arrays which are suitable for use in directive radio transmission, experience has brought about a selection of a few types which have exceptional electrical performance and which at the same time have other advantages such as low initial and maintenance costs, ease and stability of adjustment, and physical ruggedness. It seems that each of the major commercial radio engineering organizations of the world has evolved a system of its own. We find such distinctive systems as the Marconi-Franklin beam, the Telefunken "pine-tree" antenna, the SFR-Chireix-Mesny diamond-grid radiator, the A.T. & T. Co.-Sterba antenna curtain, and the RCA broadside, and harmonic-wire end-fire projectors.

The principles of modern directive antenna arrays are easily grasped, once the mechanics of wave interference are understood. However, the detailed design of any one of these systems is an engineering task of formidable proportions. Final adjustments and corrections after erection must be kept to a minimum, because of the great difficulties of making even minor changes once the rigging is complete. In design work of this sort experience plays a prominent part. The theoretical aspects of design have been discussed in a number of papers, of which some are listed in the bibliography.

32. Gain of Directive Antennas. When the radiant energy (which, with a simple antenna, would be widely diffused in space in every direction) is collected and focused into a narrow unidirectional beam by a directive array, there is a gain in effective power of transmission in the favored direction. Gain is usually reckoned in comparison with the field intensity from a single half-wave dipole located at the mean height of the array. On this basis some present-day directive arrays have gains as high as 22 db or a power gain of 158. Increases in gain result from increases in the radiation area of a broadside array, and with the length of a harmonic wire array.

33. Typical H-f Directive Antennas. The following description of typical directional antennas does not exhaust the various types but is representative:

1. *The RCA Model A Broadside Antenna.*[1] The schematic electrical circuits are shown in Fig. 36. The system consists of a large number of vertical pairs of colinear wires arranged in a plane and energized from a feed bus (transmission line) running through the middle. The feed bus

[1] CARTER, HANSELL, and LINDENBLAD, Development of Directive Transmitting Antennas for RCA Communications, Inc., *Proc. I.R.E.*, October, 1931.

has the series inductance and the parallel capacitance neutralized so as to have the characteristics of infinite phase velocity. All the radiators are thus energized in the same phase, and the direction of maximum transmission is normal to the plane of the radiators. In this system the over-all length of the radiators is 0.225 wave length, the spacing between

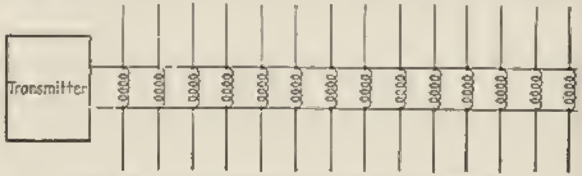


Fig. 36.—RCA model A broadside antenna.

radiators is 0.125 wave length, the maximum length of bus on each side of a feed point is 1.5 wave length, and the volt-ampere ratio between bus and radiators is 5. Such a system can have any desired length with progressive improvements in gain and directivity. Another identical

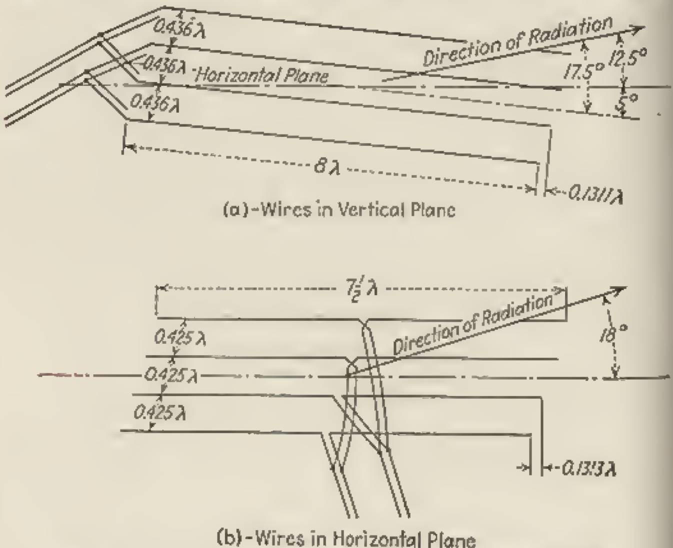


Fig. 37.—RCA model B and C harmonic wire antennas.

array in a second plane can be used as the reflector for unidirectional transmission. Gain with one bay with directly energized reflector is approximately 10 db.

2. *RCA Models B and C Harmonic Wire Antennas.* The geometry of these antennas is shown in Fig. 37. It was seen in Fig. 14 how the ampli-

tude and direction of the major radiation lobe changed as the length of the wire was increased. In this system, where each radiating wire is 8 wave lengths long, the major lobe has an angle of 17.5 deg. to the wire, and all secondary lobes are of relatively low amplitude. By using another radiator parallel to it, spaced 0.872 wave length and energized in opposite

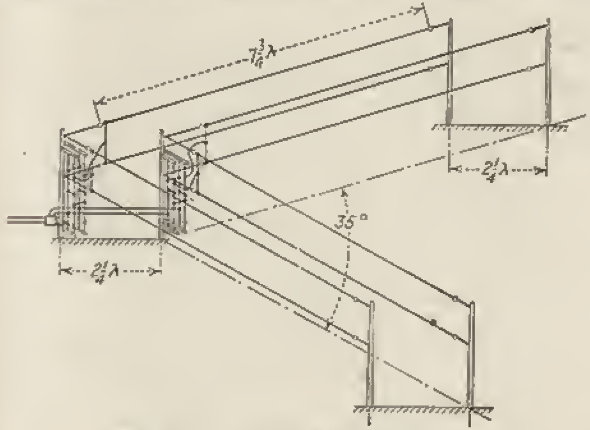


Fig. 38.—RCA model D antenna.

phase, one side of the forward and one side of the backward radiation lobe are eliminated. By adding two more such wires as reflectors (making now four parallel radiators spaced 0.436 wave length and staggered 0.131 wave length), the backward lobe is eliminated and the radiation

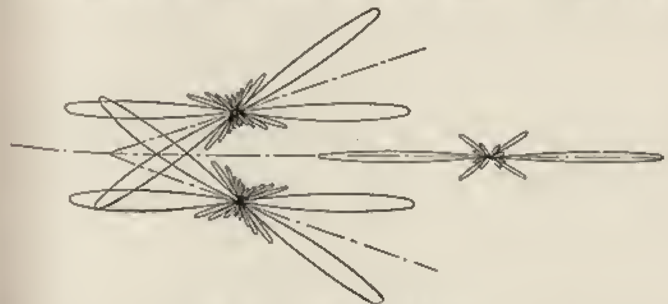


Fig. 39.—Pattern of Fig. 38.

concentrated in one very sharp forward lobe. In the model B the wires lie in a plane vertical with respect to the ground and transmit vertically polarized waves. In the model C the wires lie in a horizontal plane and radiate horizontally polarized waves. With these antennas the gains over a single half-wave dipole are approximately 12 and 12.4 db, respectively.

3. *RCA Model D Antenna.*[1] The layout of the model D project (one bay) is shown in Fig. 39. In this system, two major radiation lob (one from each side of the V) have a common direction and reinfor each other while the other two lobes are canceled, as in Fig. 39. By ad ing another V to the rear as a reflector, the backward lobe of Fig. 39 removed, giving a very sharp unidirectional beam of radiation. It ha a gain of 16 db for one bay. With two sections the gain increases approximately 19, with three to nearly 21, and with 4 to approximat 22 db. The last figure is a power ratio of 156 over that for a single hal wave dipole. In practice the point of the beam is focused at appro mately 14 deg. above the horizon.

The reference[1] contains a complete engineering and theoretical trea ment of the development of these antennas.

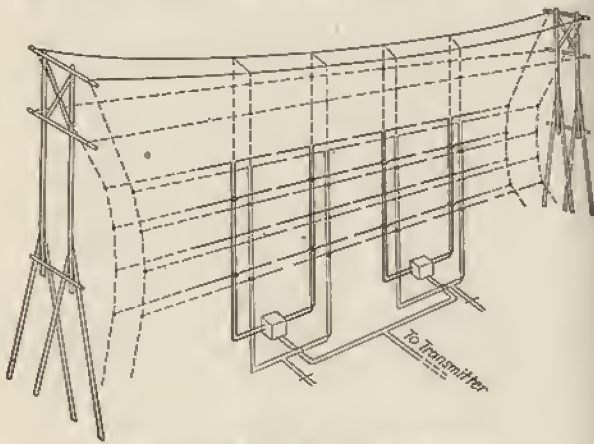


Fig. 40.—Telefunken directional antenna.

4. *The Telefunken Directional Antenna.* The arrangement of th antenna is shown in Fig. 40.[2] It consists of 64 horizontal dipoles in t vertical planes of 32 each. In each of the two planes there are four lin of eight dipoles end to end. The two planes are separated one-quart wave length, and the second (reflector) is energized by radiation from t first. The dipoles are voltage fed from the potential antinodes balanced resonant transmission lines, uniphasing being obtained attaching each successive pair of dipoles to alternate wires of the tran mission line. As with all horizontally polarized wave systems, there zero electric intensity along the ground, but the beam peaks in the vicin of 10 deg. above the horizontal, with a secondary lobe of 25 per cent p intensity maximum at 45 deg. The horizontal pattern as measured shown in Fig. 41.

5. *T. Walmsley Antenna of the British Post Office.* In Fig. 42 a shown the elements of the Walmsley beam antenna. The radiat

[1] *Ibid.*
[2] Baumler, Kruger, Pendl, and Pfitzer, *Proc. I.R.E.*, May, 1931.

arranged as shown produce a bidirectional beam broadside to the array, which usually consists of 48 energized vertical pairs. As a reflector a curtain of insulated half-wave dipoles is placed one-quarter wave length behind the array, excited by the backward radiation. A unidirectional beam is obtained in this manner. Owing to the lower current amplitudes in the reflectors as compared with those in the directly energized radiators,

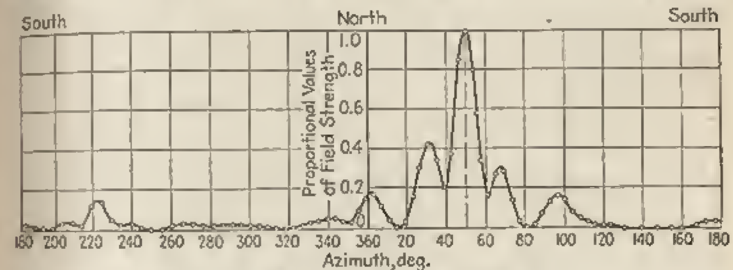


Fig. 41.—Horizontal pattern of Fig. 40.

there is not a complete suppression of backward radiation, and there is a backward lobe with an intensity 22 per cent of that of the forward beam.

6. *Marconi-Franklin Beam Antenna.* This antenna system, one of the first employed for high-speed short-wave point-to-point communication, consists of a front curtain of vertical radiators, each consisting of several

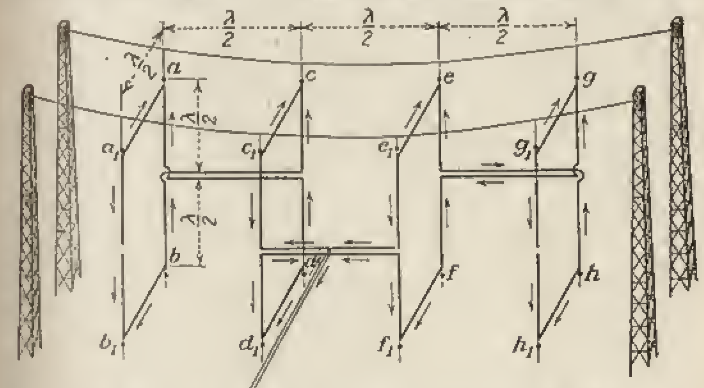


Fig. 42.—Walmsley beam antenna of British Post Office.

cophased dipoles in series, and another curtain of reflecting wires of the same construction situated one-quarter wave length to the rear. There are twice as many reflectors as radiators. The reflectors are radiation excited. In plan view, two reflectors and one radiator form the points of an equilateral triangle. Cophasing of successive radiating dipoles is obtained by winding the intermediate half-wave sections (wherein the currents are reversed) into a small non-radiating coil or web. Reflectors are energized by radiation from the front radiator curtain. A two-bay array has a gain of approximately 18 db.

7. *Chireix-Mesny (French) Beam.* Another early type of directive antenna for short waves is that used in France, shown schematically in Fig. 43. Each dipole section forms one side of a square. The currents in all the diagonals have cophased vertical and horizontal components. A similar reflecting sheet is placed one-quarter wave length behind the radiator sheet and is energized by radiation to give an essentially unidirectional pattern broadside to the plane of the radiators.

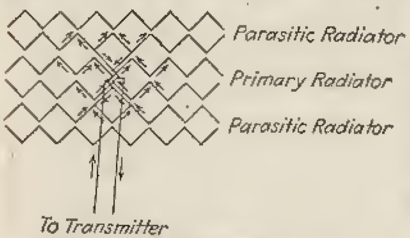


Fig. 43.—Chireix-Mesny beam.

8. *Bell System-Sterba Directive Antenna Array.* This system, used for some time in the transatlantic telephone service on short waves, is a barrage antenna employing a front curtain of several vertical radiators spaced one-half wave length, with uniphased currents, and a similar reflector curtain directly excited by transmission lines. One arrangement of an antenna of this type is shown in Fig. 44, together with transmission lines, phasing devices, protective items, and sleet-melting circuits. The unit element in this array, as shown, is a *panel* 1.5 wave lengths high and 0.5 wave length wide. The current distribution for one type of panel is shown in Fig. 45. The crossovers constitute balanced non-radiating lines, while currents in all the verticals are uniphased. Radiation from

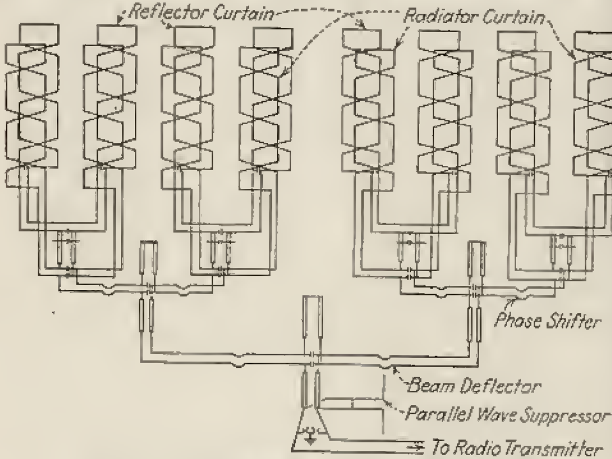


Fig. 44.—Barrage antenna of vertical radiators.

the unbalanced horizontal wires at top and bottom is reduced to negligible proportions by having equal and reversed current areas, the current nodes occurring in the middle of these horizontals. In the typical design (two bays supported by three steel towers), gains of approximately 20 to 23 db are achieved.

34. Loop-type Directive Transmitting Antennas. The principal use of loop-transmitting antennas has been in connection with radio beacons for guiding ships and aircraft. Some applications are described in Section 17 of this handbook.[1]

35. Mechanical Design of Directive Antenna Arrays. The mechanical design of a directive array for high gains is as remarkable as the electrical design. Dimensions of electrical portions must be rigorously correct and must remain so, even under conditions of severe wind and ice loading. High-gain broadside projectors are complicated webs of conductors and supporting wires, and rigging them is a specialty cultivated only by experience. The long-wire projectors are simpler, mechanically, and therefore cost less for a given gain.

Self-supporting steel towers and also guyed wood masts are used for support. General practice is to locate the active portions of the antenna at a mean height of the order of 1 wave length or more. Antennas employing vertical radiators composed of several colinear half-wave sections require towers sometimes approaching in height those used for broadcasting applications. Tower designs often include a cross arm of sufficient length to permit hanging the radiator curtain from one end and the reflector curtain from the other.

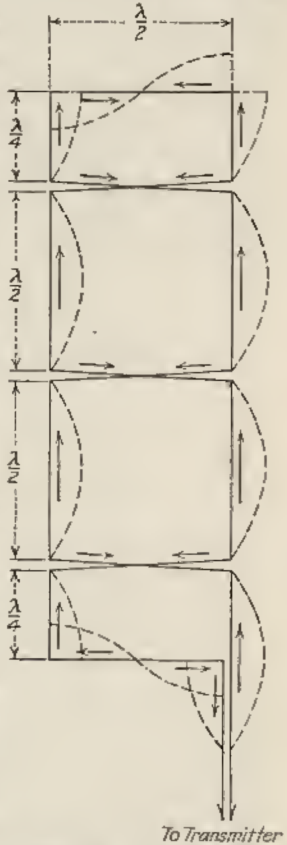


Fig. 45.—Current distribution in barrage antenna panel.

The rigging is always made up of wires, the supporting wires being broken into very short electrical lengths by insulators so that they have negligible electrical influence. Main supporting wires, usually in the form of catenaries, are under great tension and are so maintained by counterweights and anchors. Means for equalizing tensions in all parts of the rigging are important.

[1] Diamond, H., and F. G. Kear, A Twelve-course Radio Range for Guiding Aircraft with Tuned-reed Visual Indication, *Proc. I.R.E.*, June, 1930; Pratt, H., Field-intensity Characteristics of Double-modulation Type of Directive Radio Beacon, *Proc. I.R.E.*, May, 1929; Chinn, H. A., A Radio Range Beacon Free from Night Effects, *Proc. I.R.E.*, June, 1933; Diamond, H., On the Solution of the Problem of Night Effects with the Radio-range Beacon System, *Proc. I.R.E.*, June, 1933.

Insulation of the radiators with tension-type low-capacity insulators without metallic caps is practical with modern ceramic materials. Compression-type insulators assembled in strings have been used widely for this purpose also. Breakup insulators in the rigging are usually of the compression type. The voltage at the potential antinodes of the radiators depends upon the power transmitted, and the number of radiators depends upon the power transmitted and the number of radiators in the array. Liberal insulation tolerances are necessary.

Ice accumulation on the array is minimized by sleet-melting provisions, whereby large currents at commercial frequency are circulated through the conductors whenever there are ice-forming conditions. To pass heating currents through the wires when the antenna is in service requires by-pass circuits of very high impedance to the high frequencies and very low impedance to 60 cycles. Antiresonant networks or the equivalent transmission-line stub circuit fulfill this requirement.

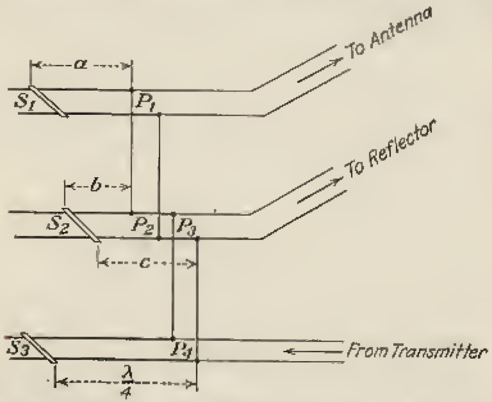


Fig. 46.—Transformer made up of transmission-line section.

The orientation of an antenna of high directivity is a matter of precise surveying. The peak of the beam is pointed along a great circle to the reception point. By adjusting the relative phases of various bays of an array, the direction of the beam can be controlled within a few degrees.

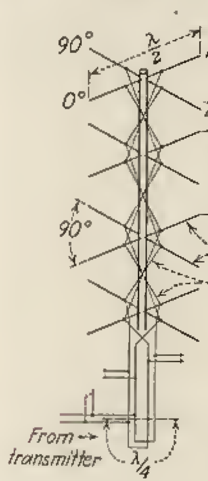
Transmission lines for transferring power to the antennas are of both concentric and open-wire types. The latter are cheaper and are extensively used. Transmission-line sections are also employed as transformers for obtaining proper relative phases and amplitudes of currents in the various conductors. An example of such a transformer circuit is shown in Fig. 46.[1] With the several types of antennas, switching means are often provided whereby the reflector and radiator screens may be interchanged electrically, thus reversing the beam 180 deg.

36. Horizontal Rhombic Antenna Used for Transmission. The rhombic antenna (see Art. 46) has been successfully applied for transmission. As a bidirective radiator, with its distant end open, it performs much in the manner of the RCA Model D (V) antenna. When terminated in its characteristic impedance, the terminal resistance absorbs the energy of one directional lobe (50 per cent of antenna power), making the system unidirectional with broad frequency response, hence desirable for many applications. Its low cost is a further advantage.[1]

[1] Carter, Hansell, and Lindenblad, *op. cit.*

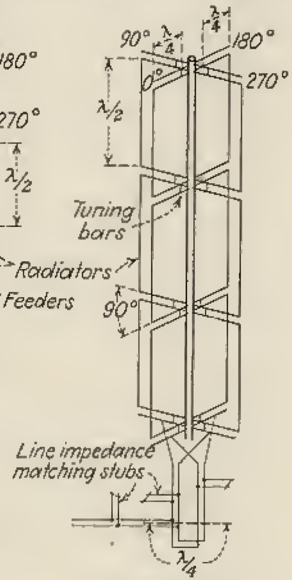
ULTRA-HIGH-FREQUENCY ANTENNAS

Antenna design technique for u-h-f applications is an art of its own. The dimensions involved are such as to permit the construction of rigid



4-Element horizontal turnstile antenna

Fig. 47.



3-Element vertical turnstile antenna

Fig. 48.

Fig. 47.—Arrangement of the horizontal turnstile antenna for ultra-high frequencies.

Fig. 48.—Arrangement of the vertical turnstile antenna for ultra-high frequencies.

structures, and one finds rod and pipe used instead of wires for the radiators and feeders. The multitudinous and rapidly growing applications of ultra-high frequencies to communication and broadcasting, including television, have proved a fertile field for ingenious electrical and mechanical designs for radiating systems. In this field we also find the frequent necessity to employ the principles of directive antenna design to obtain non-directive transmission from a number of radiators, functioning cooperatively. In general u-h-f antennas are located at points

[1] Foster, D., Radiation from Rhombic Antennas, *Proc. I.R.E.*, October, 1937.

of maximum accessible height, such as the top of a tower, high building, or mountain, and many design problems are imposed by the situation.

37. Turnstile Antenna. The turnstile antenna of Brown[1] is one widely used for non-directive transmission with relatively high gains due to low-angle concentration of energy. Arrangements of this type antenna for both horizontally and vertically polarized transmission are shown in Figs. 47 and 48. A mechanical advantage offered by the turnstile antenna is its all-metallic construction. Conductor potentials and radiated field intensities are nil at the vertical axis which permits the use of a metal pipe

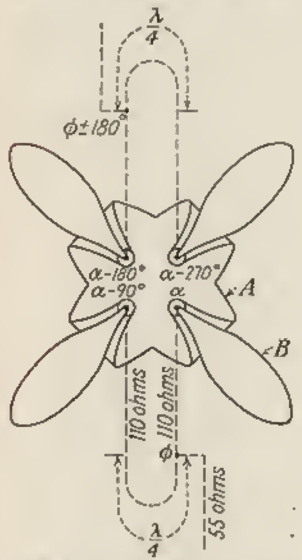


FIG. 49.—Schematic arrangement of wide-band television antenna developed for the Empire State Building.

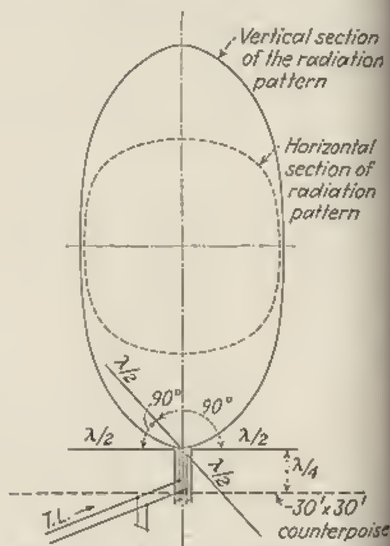


FIG. 50.—CAA cone-of-silence marker antenna system.

as a central supporting member. A circular field is produced by exciting opposite radiators in opposite phase and the quadrature conductors in quarter phase. The phasing is done by the feeders which form a part of the system design.

38. Horizontal Turnstile Antenna Using Ellipsoidal Radiators for Wide-band Television Transmission. A single-stage horizontal turnstile antenna employing ellipsoidal radiators of proper proportions was developed for high-definition television transmission from the Empire State Building in New York City.[2] This requires the essentially uniform transmission of side frequencies having a band width of more than 1

[1] BROWN, G. H., The Turnstile Antenna, *Electronics*, April, 1936.
[2] LINDENBLAD, N. E., The Television Transmitting Antenna for Empire State Building, *RCA Rev.*, April, 1939.

per cent with respect to the carrier frequency. Of still greater importance was the requirement that the input impedance to the antenna transmission line be substantially constant at all the frequencies within the video spectrum of the transmitter. The basic arrangement for this purpose is shown in Fig. 49.

39. Cone-of-silence Marker Antenna. For the purpose of providing a positive indication when an airplane passes over a radio range station instead of the negative one obtained by the cone-of-silence of the four-course l-f radio range antenna, there was developed another application of a single-stage horizontal turnstile antenna for the transmission of a vertical beam of energy at 75 Mc. The antenna is located one-quarter wave length above a horizontal metal counterpoise composed of 4-in.

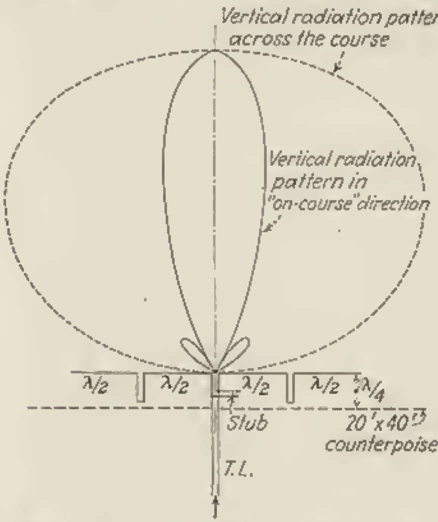


FIG. 51.—CAA fan-marker antenna system.

square mesh wire on a structural steel framework. The counterpoise, acting as a reflecting screen of permanent electrical characteristics reinforces the vertical transmission of energy from the antenna system. Figure 50 shows the arrangement of the cone-of-silence marker antennas as currently used on the airways in the United States.[1] This is already being superseded by improved designs giving greater vertical directivity.

40. Fan-marker Antenna. For use in determining fixed points of location and course identification in conjunction with navigation with l-f four-course radio ranges, directive antennas are used to transmit a thin fan of radio energy at 75 Mc, through which a plane flies and identifies his position. In this system as currently used in the United States and Canada, four cophased dipoles are disposed horizontally in the direction of the radio range course on which it is located, over a counter-

[1] JACKSON and METZ, Development, Adjustment and Application of the Z-Marker, *Report* 16, Bureau of Air Commerce, July, 1938.

poise similar to that used for the cone-of-silence marker beacon anten
This system is shown in Fig. 51.[1]

41. Ultra-high-frequency Four-course Radio Range Antennas. Th application of four-course u-h-f radio ranges to the airways continu under development in 1940 and is classed as one of the most importa projects for the immediate future for airways use. Improved cou stability and very much lower cost for antenna and equipment are t principal gains expected.

Interlocked figure-of-eight patterns produce the four courses by th familiar A-N keying. The ease of orienting the antenna makes the u of a goniometer unnecessary for course alignment. Course squeezi and bending, however, are not yet achieved at this writing but are capabl of development as the need arises. Waves of one polarization only a essential for this purpose. Current developments employ pure horizont polarization, so that horizontal loop radiators are employed with eve effort made to avoid any leakage radiation which is vertically polarize

The trend is also toward the development of two-course u-h-f rad ranges with aural course identification.

ANTENNAS FOR RECEPTION OF ELECTRIC WAVES

42. Non-directional General-purpose Receiving Antennas. Th ordinary receiving antenna for general purposes is a single wire, of leng more or less proportional to the wave lengths to be received but usual only a small fraction of these wave lengths in physical length. It tak all the conventional forms, inverted L, T, or vertical. In some cases th antenna is resonated for reception of a particular wave length, but mo commonly it is aperiodic by being terminated at the point where receivi apparatus is located in a resistance. One or more receivers of high-inpu impedance are bridged across the terminating resistance, and selectivi is obtained in the receiving apparatus.

For optimum reception for waves arriving from some preferred dire tion, account must be taken of the wave tilt and the wire must be oriented as to bridge the greatest potential difference in space which giv a maximum voltage across the terminating resistance. It is well know that any antenna that is not a simple vertical has some inherent directi ity, though it may be very small. Where absolute non-directivity unessential, advantage should be taken of the various simple means f obtaining optimum response to waves coming from preferred directio Of these, one is to incline the wire at an angle normal to the wave tilt the vicinity of the receiving site, and another is to locate the wire abov any other wires or metallic structures in the vicinity. Field-intensit measurements have shown that the field intensity under or near overhe wires and metallic structures falls to a small fraction of its free-spa value when these conductors form *apertures* which are smaller than wave length in dimensions. However, local electrical noise is not sim larly influenced. To obtain a favorable signal/noise ratio, it becom important to have the antenna high above any other parasitic conducto in the vicinity.

43. Directive Receiving Antennas. Except for mobile stations a home-broadcast reception, there are few cases where some degree directive discrimination at the receiver is not desirable or even necessa

[1] Development of an Improved UHF Radio Fan Marker, *Report* 14, Bureau of A Commerce, July, 1938.

In the fixed point-to-point services, highly directive receiving antennas are used for both long- and short-wave reception.

There are four main types of relatively high directivity receiving antennas, as follows:

1. The loop (frame) antenna which can be rotated, or the fixed crossed-loop system with rotating radio goniometer. With these the directivity is adjustable by the operator. They are usually employed as direction finders.[1]
2. The directive antenna array which is the same as that used for directive transmission. Used for the fixed services, on high frequencies.
3. The long folded-wire types of which the Bell System-Bruce rhombic antenna is an example. Used for high frequencies in the fixed services.
4. The long-wire transmission-line type of antenna known as the Beverage, or wave, antenna. Used for l-f and h-f reception in the fixed services.

44. Loop Antennas. This form of antenna is well known to the art and is described and explained in almost every publication on elementary radio. Its response is of a very low order, requiring a very high gain receiver. Its small mechanical dimensions make it a useful device for some portable applications, such as military field sets and field-intensity meters. Its constant electrical characteristics and its independence of ground have special value in the latter application. However, its principal application is in direction-finding apparatus, which is discussed elsewhere in this handbook.

The response in the maximum directions is very broad, but the minima are very sharp. When used in direction finders, the signal is adjusted for a minimum which can be determined with great accuracy, especially when the loop is balanced to ground. A loop, in conjunction with a vertical wire antenna, produces a unidirectional response which enables one to determine the exact direction of the arriving waves. Without this auxiliary vertical "sense" antenna, the loop has two responsive directions 180 deg. apart and can therefore give errors of this order in cases where there might be some doubt concerning the relative geographical positions of transmitter and receiver, as with ships at sea.

45. Directive Transmitting Antennas Used for Reception. In certain communication systems, such as the Telefunken, Marconi, and Société Radio Française, the receiving antenna is a duplicate of that used for transmission. With extended arrays there results a directional discrimination comparable with that at the transmitter. Thus static and interfering signals or disturbances originating in unfavored directions are essentially eliminated from the receiver. Extended arrays also give a limited measure of diversity effect (discussed more fully under its proper title) which tends to level out fading variations. Any of the directive arrays already described could be used for reception, provided they are properly oriented and polarized.

Some of the special problems in connection with reception may be briefly outlined as follows:

1. The arrival of a multiplicity of waves from the same transmitter, which have definite time differences as well as different angles of arrival.
2. All the components of a wave group have individual variations in intensity and relative phase, so that their group influence is highly variable. The

[1] Smith-Rose, R. L., Radio Direction Finding by Transmission and Reception (with extensive bibliography), *Proc. I.R.E.*, March, 1929; Palmer, L. S., and L. L. K. Honeyball, The Action of Short-wave Frame Aerials, *Proc. I.R.E.*, August, 1932.

result is familiarly known as *fading*, which may be uniform for a small ba of frequencies (such as those composing a modulated signal), or non-unifor The latter, called *selective fading*, produces serious distortion of telepho signals.

3. It has been discovered that signals which fade do not fade in exac the same manner or at exactly the same time at different geographical po tions. This latter, now known as the *diversity effect*, has been utilized in RCA system of diversity reception, to be described.

4. Atmospheric disturbances, as well as interfering signals, are reduc in the same degree as the directivity is increased in a favored directio thus providing improved signal/noise ratios. This advantage falls dow however, when the disturbances originate in the direction of the desir signals.

5. High gain is often required to override receiver noise.

It is plain that these problems are peculiar to the reception end of communication circuit. Adapting a transmitting array to receptio may partially satisfy problem 1 if its horizontal and vertical directivi are high enough to give a sensible reduction of those minor componen of the wave group which are more harmful than useful. A transmitti array seldom is of sufficient geographical extent to give much spa equalization of signal by diversity. Furthermore phase differences co tinue to exist between the currents in the system due to the various wa components, so that comparatively little improvement in fading obtained in this manner. From the standpoint of gain, the transmitt type of antenna is perhaps equal to the special types developed for recep tion purposes. Transmitting antennas, being generally of the resonan conductor type, suffer rather high reradiation losses when used fo reception.

46. Folded-wire Receiving Antennas.[1] A very simple and effectiv type of receiving antenna has been developed by Bruce and his cowork of the Bell System, known as the *rhombic antenna*. This antenna h several useful intermediate forms between an electrically long vertic wire and the horizontal rhomboid, or diamond. Among these are th tilted wire, the vertical inverted V, and the vertical diamond. Th application of any one of these forms must take into account the polariza tion of the incoming waves, the direction and the wave tilt, the frequenc range to be covered with one antenna, and the available space.

Three forms of this antenna are shown in Fig. 52. In (*a*) is a vertic inverted V which has bidirectional response. In (*b*) is the same antenn equipped to absorb completely in a terminating resistance all energ received from a backward direction, giving unidirectional response the receiver. Both (*a*) and (*b*) are vertically polarized. In (*c*), for hor zontally polarized waves, terminated to give unidirectional respons there are in effect two opposed V sections of the type of (*a*) and (*b*). Fo any wave direction, there exists a wire length l which will give maximu response. This occurs when the wire length is one-half wave lengt greater than its projection upon the line representing the wave directio in the plane of the antenna. The horizontal rhombic antenna (*c*) ha zero response along the ground, and the peak of the directive patter can be focused at the vertical angle which corresponds to the incomin wave direction by suitably proportioning the antenna dimensions.

[1] Bruce, E., Developments in Short-wave Directive Antennas, *Proc. I.R.E.*, August 1931; Bruce, E., A. C. Beck, and L. R. Lowry, Horizontal Rhombic Antennas, *Proc. I.R.E.*, January, 1935.

In the design of a horizontal rhombic antenna there are three variables, the length of a side l, the angle ϕ, and the height above ground H. The

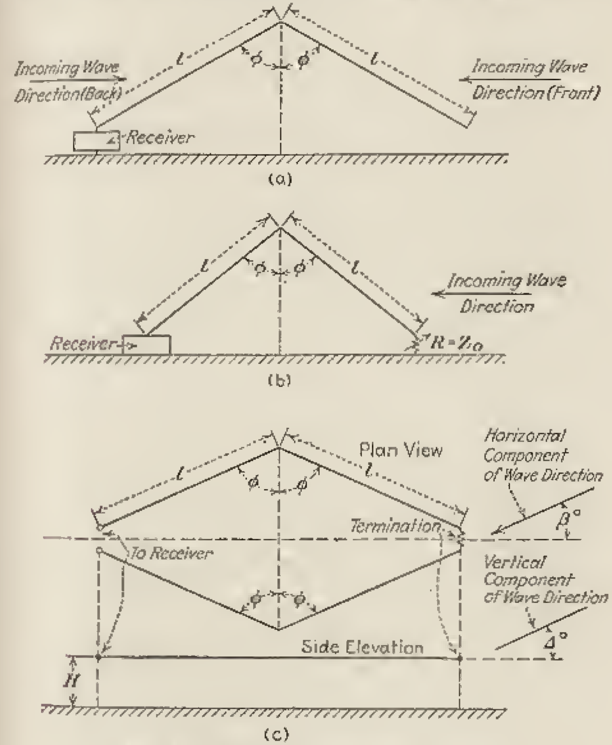


Fig. 52.—Rhombic antennas of Bruce.

lowest practical height is when

$$H = \frac{\lambda}{4 \sin \Delta}$$

The value of the angle ϕ is obtained when

$$\sin \phi = \cos \Delta$$

For maximum gain the value of l is found from the equation

$$l = \frac{\lambda}{2 \sin^2 \Delta}$$

With this value of l the peak of the major directivity lobe may not fall at the desired angle corresponding to the wave direction. Where the received wave direction is unstable or where maximum signal to noise discrimination is sought, the length is adjusted to focus the point of the beam in the wave direction. This occurs when the length is shortened to

$$l = \frac{0.371}{\sin^2 \Delta}$$

The greater the length, the greater the range of frequencies which can be efficiently received on one antenna.

The main axis of this antenna is oriented in the great-circle direction of the associated transmitting station.

The proper value of the terminating resistance for back-wave suppression is determined experimentally. Impedance measurements of the antenna are made at the receiver end with trial values of resistance at the termination. The proper termination is that which gives the flattest impedance-frequency characteristic. One might make a preliminary determination of the order of the terminal resistance by making a rough calculation of the characteristic impedance of the antenna as a transmission line of parallel wires.

Finally the output terminals of the antenna are connected through a termination network to a transmission line running to the receivers. The terminal impedance is matched to that of the transmission line. Accurate balance to ground must be maintained in the antenna system, as well as in the transmission line, if it be of the open-wire type.

47. Multiple-unit Steerable Antenna.[1] The receiving rhombic antenna can be made directive in the vertical plane by altering its length-width proportions. The angle of arrival of waves changes from time to time; and there are groups of waves arriving simultaneously with different angles of incidence, any one of which may be of dominant magnitude. To take advantage of selecting the dominant wave group for optimum reception, experiments were carried out with rhombic antennas which were mechanically adjustable in length and width so as to obtain a "steerable" antenna, responsive to various angles of arrival as desired.

The same ends were later obtained by electrical steering in the multiple-unit steerable antenna system. The multiple-unit steerable antenna (known as MUSA) is part of the multiple-unit steerable antenna system of short-wave reception, an elaborate and highly developed method of selecting and combining in proper phase the ever-changing multiple wave groups arriving at the receiving location. The antenna comprises a multiplicity of rhombic antennas, arranged in line on the great circle bearing to the transmitting station. Each antenna is a directive responsive device which feeds its energy into a concentric transmission line where it is brought back to the receivers. A long line of such antennas provides extreme space diversity, and the cumulative energy collected over a continuous expanse of as much as 2 miles of antennas, properly phased out, reaches large values. In this system the dominant wave group is selected and the others rejected. By virtue of space diversity, sharp directional characteristics due to the antennas, together with the selective phasing of the multiple wave

[1] Friis and Feldman, A Multiple Unit Steerable Antenna for Short-wave Reception, *Proc. I.R.E.*, July, 1937.

groups in the receivers, unusual signal/noise ratios are obtained, and fading is effectively equalized.

In this system the receiver plays as important a part as the antenna array in obtaining the desired performance. It is through the medium of the phasing of the various individual lines from the antenna elements and again through the phasing of the branches of the receiver that the array is given its continuously variable control of the two or more vertical directional lobes. In its commercial form this phasing is accomplished automatically. The entire vertical plane is explored continuously, and automatic phasing causes the antenna response pattern to follow that of the angle of arrival of the dominant wave groups from moment to moment. It is the complete receiving system, then, and not the antenna alone, which achieves directional steering.

48. The Beverage (Wave) Antenna.[1] This type of antenna, one of the earliest effective directive receiving systems to be used commercially, is a long transmission line. It is named after its inventor, H. H. Beverage but is also called the *wave antenna*. A long open-wire transmission line pointed in the direction of a down-coming wave, has a high degree of exposure to the horizontal component of the wave front, which induces in the line a continuous series of e.m.fs. that are propagated along the wires in the form of a traveling wave. A wave front sets up a wave in the wire which starts at the distant extremity (in the direction of the arrival of the space wave) which is propagated toward the home end where a receiver is situated. In addition, the entire wire receives energy from the down-coming wave, so that the effects are cumulative at the receiver and a relatively large amount of energy is extracted from the space wave for energizing the receiver. The antenna functions only where there is an angular difference between the direction of the wire and the incidental direction of the space wave. This condition is suitably met in practice due to natural conditions, since finite earth conductivity causes a wave traveling in space near the surface to be tilted forward at a considerable angle. Thus a long transmission line parallel to the surface of the ground has a workable inclination with respect to the wave front. This applies to vertical polarization.

The Beverage antenna has many useful forms which are specially adapted to long-wave reception, to short-wave reception, to bidirectional and unidirectional selectivity, for vertical and horizontal polarization, etc. A thorough treatment of these is impossible here, and detailed data must be obtained from the original and subsequent papers on the subject.

For long waves the antenna construction is very similar to ordinary open-wire telephone lines. The antennas may be located at a considerable distance from the station and coupled to the receivers by transmission lines. The Beverage antenna is directive in the line of its orientation and is made unidirectional by terminating the distant end in a resistance equal to the characteristic impedance of the line. Thus energy collected from a wave in the backward direction is completely dissipated without producing any influence in the receiver. Directivity may be sharpened by using two or more antennas in an array. This has been done in the system shown in Fig. 53 which is used for transatlantic telephone recep-

[1] Beverage, H. H., C. W. Rice, and E. W. Kellogg, The Wave Antenna, *Trans. A.I.E.E.*, February, 1923; Beverage, H. H., and H. O. Peterson, Diversity Receiving System of RCA Communications, Inc., for Radio Telegraphy, *Proc. I.R.E.*, April, 1931; Bailey, Austin, S. W. Dean, and W. T. Wintringham, Receiving System for Long-wave Transatlantic Radio Telephony, *Proc. I.R.E.*, December, 1928.

tion on long waves. One of the several forms of the antenna which is used in this application is that which couples the receiver to the end of the antenna that is nearest the transmitting station. A two-wire line is used to achieve this in the following manner: Waves arriving from the preferred direction act upon the two wires in parallel to ground, and the induced wave of energy in the wire travels to the distant end where it

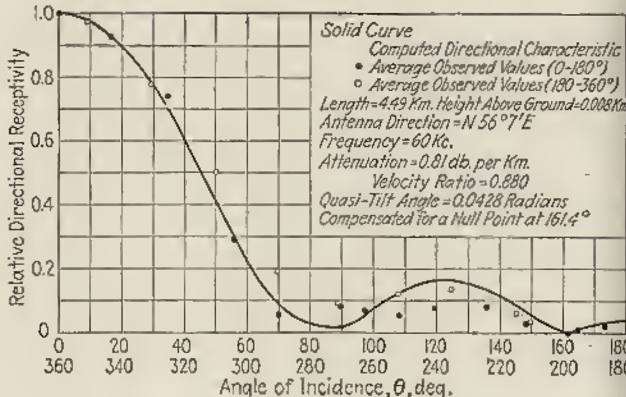


FIG. 53.—Directivity of Beverage antenna.

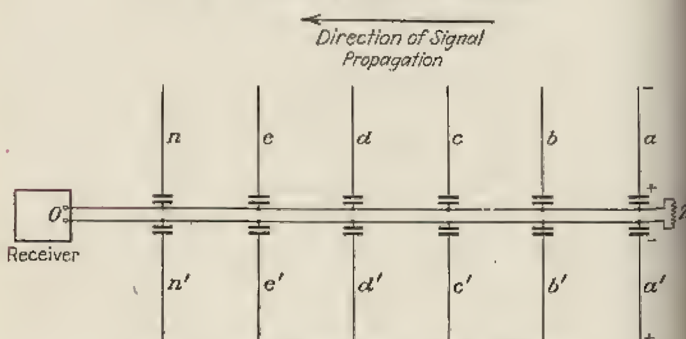


FIG. 54.—High-frequency Beverage antenna.

encounters a reactive network called a *reflection transformer*. This device reverses the phase of the wave in one of the wires and reflects the energy from the end back to the receiver, the reflected wave of energy now traveling in the two wires balanced to ground. The receiver coupling network terminates the line and absorbs all the wave energy in actuating the receiver. A wave entering the system from the reverse direction travels along the two wires in parallel against ground, producing no potential difference across the balanced termination and therefore has

no influence on the receiver. Instead, the circuit to ground is terminated in the characteristic impedance of the parallel-grounded system, and the unwanted wave is completely dissipated in a resistor.

In its very simplest form the Beverage antenna is a single straight horizontal wire a few feet above grade level, the length being anywhere from one to several wave lengths. The characteristic impedance of this wire unbalanced to ground is roughly calculable by using the image as the second conductor in a parallel wire system. The receiver is coupled in at one end of the line, and the other end is terminated in a resistance equal to the characteristic impedance. Stable ground systems are necessary at both ends.

A form of the Beverage antenna used at high frequencies (horizontally polarized exposure) is shown in Fig. 54. This is a plan view of the conductors. The side wires extract from the traveling waves energy which is coupled into the central transmission line which is balanced to ground.

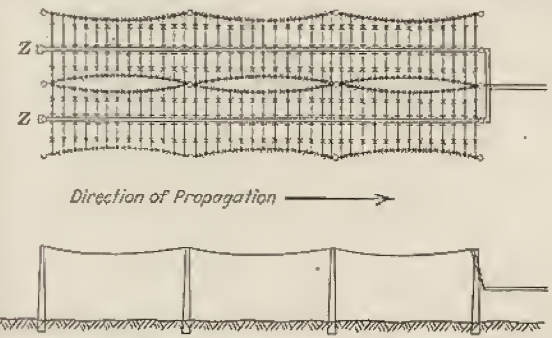


FIG. 55.—Double broadside Beverage antenna.

The side wires act as distributed loading of the transmission line, modifying (reducing) its phase velocity of propagation and its characteristic impedance. The branches with their coupling condensers have a capacitive effect on the line within the desired frequency range, and they must be close enough together to produce the effect of continuous loadings (maximum separation three-eighths wave length at the shortest wave length to be received).

A practical form of the antenna, where two are used in broadside for higher directivity, is shown in Fig. 55. This also indicates the method of rigging it, the location of insulators, etc. With antennas of this type, the signal/noise ratio is reduced from 24 to 39 db over that obtained with a single dipole, when the static directions are not in the line of maximum response.

49. Diversity Reception. The fading of h-f radio signals has always been a major problem. Antenna design, in the phases treated in this work, is at best only moderately effective in reducing it. Diversity reception has proved a long step forward in combating signal fading. In this system three separate receiving equipments are employed, the antennas for them being located at different geographical points. The distance between antennas is arbitrary, being in practice sometimes a mile or more. For obvious reasons the three antennas are not in a straight line but disposed somewhat as shown in Fig. 56. Diversi of fading with geographical separations of this sort produces an avera cumulative effect which is quite constant. To eliminate the effect phase relations when the outputs from the three systems are mix this function is achieved af detection.

Any type of receiving anten may be employed, but in th RCA Diversity Receiving Sy tem, the Beverage-Peterson an tenna shown in Fig. 56 is used the unit.

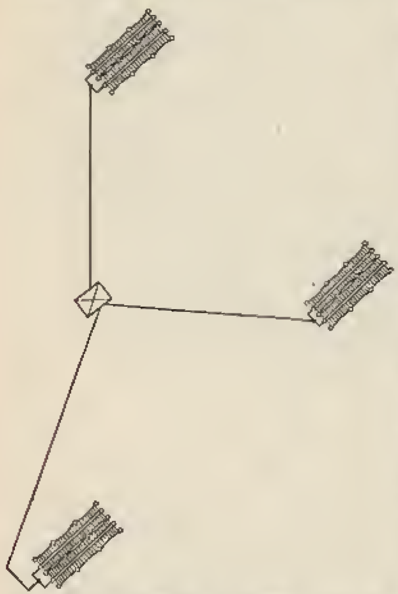


Fig. 56.—Antennas arranged for diversity reception.

BROADCAST RECEIVING ANTENNAS[1]

50. All-wave Receiving Antennas. The all-wave receiv is now more or less standardiz and contains usually three fr quency bands: the broadca band of 550 to 1,600 kc; th "police" band, from 1,600 6,000 kc; and the "short-wave band from 6 to 22 Mc. The limits sometimes are sligh changed, and one of the tw higher frequency bands is omi ted in some sets.

In general, an ordinary Ma coni open-wire antenna about to 100 ft. in length gives sat factory signal voltage, but o account of "man-made static or interference produced b electrical apparatus generati transient currents having components in the bands above mentioned, th simple open-wire antenna is not satisfactory, particularly in metropoli areas.

With the advent of television and u-h-f broadcasting it is necessar more than ever to provide special types of antennas, having very directi characteristics, and transmission lines between the antenna proper an the receiver, incapable of picking up interference.

51. Types of Antennas. All the antenna structures commonly used broadcast reception may be classified into doublet- and Marconi-typ antennas according to whether they act by virtue of phase differen within the antenna wire, or as elevated capacities with respect to th surrounding medium, called *ground,* which may be the metallic structu of a building, the piping, or even the power line. The choice of prop ground makes a lot of difference in the signal/noise ratio.

The doublets consist of two arms, usually of nearly equal length (Fi 57), and called *simple doublets;* or they may contain several pairs of arm

[1] By J. G. Aceves, Amy, Aceves & King.

interconnected at the common gap (Fig. 58). Each pair of arms is made approximately of one-half wave length for the mean frequency of the band intended to be covered, although with a transmission line that matches the impedance of the doublet, either directly or through a matching device, the resonance is not critical.

For television work it is necessary to receive signals from only one direction to avoid the formation of secondary images called *ghosts,* which originate from reflected waves arriving with a certain phase retardation. This requires the use of reflectors to make the doublet unidirectional as

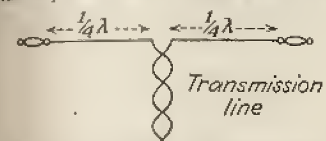


Fig. 57.—Simple doublet antenna.

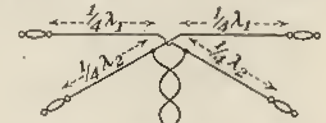


Fig. 58.—Doublet antennas for noise reduction.

far as possible. In some cases a double doublet of the same length for the two units and with a double reflector may help considerably in boosting the pickup of the structure in the desired direction. Figure 59 shows a typical television antenna.

The Marconi-type antennas act, as stated above, as capacitative generators below their natural period. This is normally the case in the standard broadcast band. However, they may be used for frequencies considerably higher, in fact sometimes so high that several wave lengths may be developed in the long-wire structure. In such cases they present alternately capacitive reaction and then resistive, inductive resistive and again capacitive reaction; and the cycle will repeat itself indefinitely, according to the number of one-quarter wave lengths. This is true in general, but, when the transmission line introduces directly or through a coupling device a large resistance reaction, the long wires begin to show less and less peaks in their voltage *versus* frequency characteristics.

The Marconi antennas can be easily changed into very directive structures, when they are several wave lengths long, by terminating them in a suitable resistance at the far end, and by proper selection of the reflected impedance at the transmission-line end.

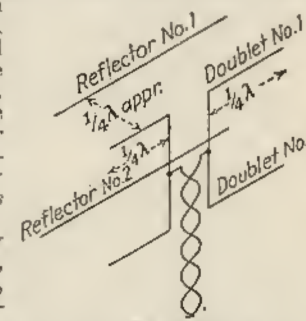


Fig. 59.—Television receiving antenna and reflection.

Combination of two such structures may become a "diamond" or rhombic antenna with very directive properties, but structures of this and other complicated types, including arrays of doublets or half doublets are seldom used for broadcast purposes, except for demonstration in stores and in localities far away from the stations (mostly for television), and where a number of receivers are to be operated from one antenna structure. Foreign reception by large commercial companies uses a number of these highly directive structures.

52. Elimination of Interference. The vast majority of radio receiver have enough sensitivity to permit reception with a very poor sign energy pickup, as can be readily seen in the typical example of automob radios. Therefore the main problem is not so much to increase th signal energy pickup by means of an antenna system very well design as it is to reduce the amount of interference which is inevitably prese even in isolated houses where other electrical apparatus containi current interrupters of one kind or another are always found. Hen it is very important in passing judgment on the merits of a given anten to examine it first of all from the advantage secured by its use in signal noise ratio. Of course, in "dead" spots, signal energy requirements m be of paramount importance.

For interference waves to assert themselves, they must contain co ponents within the band to be received. It follows that it is possible t reduce the interference by broadcasting in the region where those co ponents are a minimum or not present at all. It is well known that abo 40 Mc these components are usually very weak and "natural" static practically absent. For this reason the sound channels of televisi stations are remarkably free from noise.

An additional step in noise reduction is obtained by the use of f.m instead of a.m., thereby permitting the use of a limiter (see Sec. 1 which forms part of special receivers for frequency modulated broadca signals.

In television reception the elimination of interference is still mo necessary, and, although there are comparatively weak components the neighborhood of 50 Mc, they are sufficiently strong to make the selves obnoxious in visual reception. They originate mostly fro diathermy apparatus and internal-combustion engine ignition system

53. Noise-reduction Methods. Interference enters a radio receiv

1. Through the antenna.
2. By down-lead or transmission line pickup.
3. By direct pickup of the receiver.
4. By common coupling between the signal pickup circuit and the noi producing circuits.

The fourth mode of entry gives the greatest amount of trouble a will be treated more at length.

1. *Antenna pickup* can be reduced only by placing the antenna in field which is strong for the signals but weak for the interference.

With a low-loss transmission line it is possible to place the antenna a very great distance from the receiver. For example, this was done in a very powerful hydroelectric plant in the West where the line was near ½ mile long. From almost impossible broadcast reception it beca possible to listen to distant stations after the installation of the noi reducing system and after moving the antenna far away from the outd power network into the side of a hill.

2. *The down-lead pickup* is eliminated simply by eliminating the o down-lead and replacing it by a transmission line, which may be of t balanced type or of the concentric or shielded type. A well-balan shielded line seldom gives the expected increase in signal/noise ra over the open type, provided the terminating couplers are of the corr design.

3. *Direct Pickup.* Modern receivers are fairly well shielded, especially those for use in motor vehicles. Only the inexpensive household radio sets are likely to pick up much interference by themselves when provided with a good noise-reducing antenna system.

The exception is the loop set, which can overcome interference only by turning the loop to a minimum pickup direction with respect to the noise, whenever this is possible without also eliminating the signal. When there is more than one source of interference, this expedient fails. Loop sets have the advantage that they are considerably less affected by noise currents via the power-line connection.

The internal noise of radio receivers due to shot effect and thermal agitation has not been mentioned because the signal level is usually much higher than 30 or 40 db below the standard 1 mv input voltage, but in

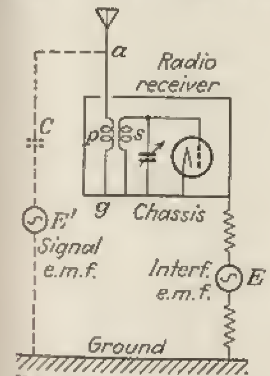


Fig. 60.—Noise circuits in radio receiver.

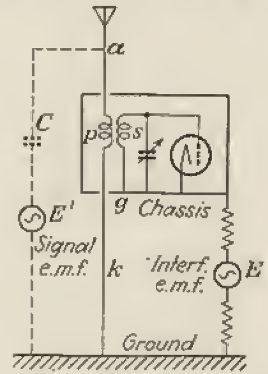


Fig. 61.—Circuit immune from power-line interference.

loop sets this does not hold as a rule, particularly if the receiver is not near a window or other free space unshadowed by metallic structures.

4. *Common Coupling between Signal and Noise Circuits.* Figure 60 illustrates schematically the principle involved in this type of interference. Let E be a source of interference in series with the lamp-cord connection of the radio set. It will force a current through the chassis of the receiver g, through the primary p of the input transformer, and through the down-lead and antenna a and its equivalent capacity C to ground. Obviously a secondary voltage will be developed and applied to the input of the tube of the receiver. Now consider Fig. 61. Everything is the same except the input-circuit connections of the receiver.

Here the signal voltage, represented by a source E', will send a current through the effective capacity of the antenna C, through the antenna and down-lead wire a, the primary p, and back to ground through a conductor k, not common to the path of the current from the source of interference E, which meets a "dead end" at the chassis of the radio receiver and therefore is incapable of delivering an e.m.f. to the input of the first tube of the receiver.

While the antenna and down-lead, in the above illustration, are op- to attack from radiated interference, the system of Fig. 61 is immune power-line interference. The capacity between t windings of the input transformer should be sm This is essential to keep in mind when design a noise-reducing antenna or radio-set transform Otherwise appreciable current will flow through t capacity and reach the antenna.

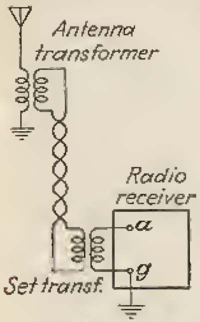


Fig. 62.—Simple system for improving signal/noise ratio.

To eliminate radiated interference, as well "conductive" due to common paths of noise a signal currents, a complete system, such as illustrated in Fig. 62, will increase the signal/n ratio by as much as 20 to 30 db. The simple s tem shown in Fig. 62 may be extended to a plurali of bands by the use of a number of transformers c ering the various selections. Figure 63 illustr a complete antenna system for the broadcast ba short-wave band, and television or f-m recept with full noise-reduction design in all the bands. this particular illustration the antenna struct acts as a Marconi antenna for the broadcast ba and as a doublet for the higher frequency bands.

54. Master Antenna Systems. The receiving system of Fig. 63 suitable for the operation of a number of radio receivers, by using

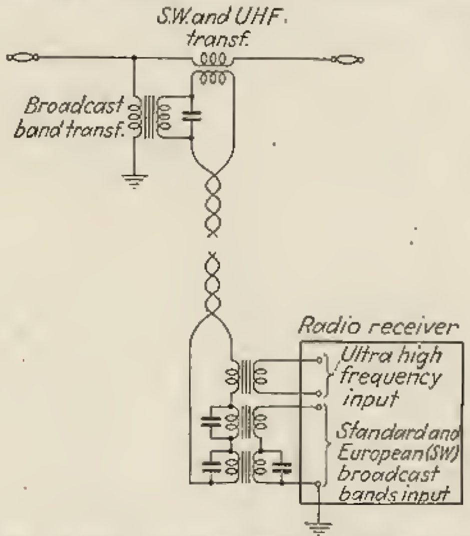


Fig. 63.—System for operating several broadcast receivers.

plurality of receiver couplers across a line terminated in its surge re ance, provided that the receiver couplers have a suitable ratio of tr

formation that will prevent not only excessive attenuation by overloading the transmission line but likewise undue reactions between the various receivers themselves.

In this case the maximum of noise reduction is to be sought and full isolation used, *i.e.*, nothing but inductive couplings and complete separation from power-line interference circuits and the signal channels. When the signal strength is too weak to operate an unduly large number of receivers, an amplifier between the antenna and the transmission-line networks is sometimes employed to increase the signal voltage available at the outlets without amplifying interference except that which the antenna itself picks up.

In places where noise is not severe and the cost of installation is a prime consideration, it is satisfactory to provide an isolation transformer at the antenna and some form of conductive receiver couplers. For example,

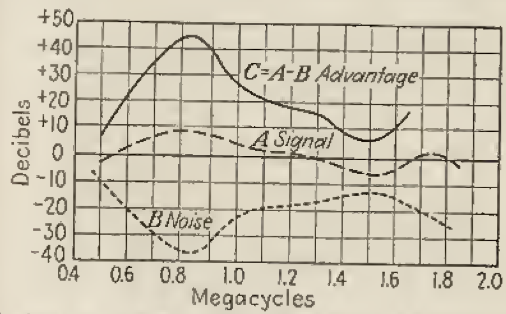


Fig. 64.—*A*, signal as received with ordinary antenna equipment; *C*, advantage secured by noise-reducing equipment.

couplers consisting of series condensers and resistors may be used or resistors or a combination thereof, mostly to minimize reactions between receivers.

References

General Properties of Radiating Systems:

Barrow, W. L.: On the Impedance of a Vertical Half-wave Antenna above an Earth of Finite Conductivity, *Proc. I.R.E.*, February, 1935.

Bashenoff, V. I., and N. A. Mjasoedoff: Effective Height of Closed Aerials, *Proc. I.R.E.*, June, 1931.

Bechmann, R.: Calculation of Electric and Magnetic Field Strengths of Any Oscillating Straight Conductors, *Proc. I.R.E.*, March, 1931.

———: On the Calculation of Radiation Resistance of Antennas and Antenna Combinations, *Proc. I.R.E.*, August, 1931.

Dellinger, J. H.: Principles of Transmission and Reception with Antenna and Coil Aerials, *Bur. Standards Sci. Paper* 354, 1919.

Hansen, W. W., and J. G. Beckerley: Concerning New Methods of Calculating Radiation Resistance with or without Ground, *Proc. I.R.E.*, December, 1936.

———, and J. R. Woodyard: New Principle in Directional Antenna Design, *Proc. I.R.E.*, March, 1938.

Hund, A.: "Phenomena in High-Frequency Systems," McGraw-Hill Book Company, Inc., New York, 1936.

Murray, F. H.: The Mutual Impedance of Two Skew Antennas, *Proc. I.R.E.*, January, 1933.

Norton, K. A.: Physical Reality of Space and Surface Waves in the Radiation Field of Radio Antennas, *Proc. I.R.E.*, September, 1937.

Pistolkors, A. A.: The Radiation Resistance of Beam Antennas, *Proc. I.R.E.*, March 1929.
Ramsey, R. R., and R. Dreisback: Radiation and Induction, *Proc. I.R.E.*, August, 19
Stratton, J. A., and H. A. Chinn: The Radiation Characteristics of a Vertical Half-wave Antenna, *Proc. I.R.E.*, December, 1932.
Uda, S.: High-angle Radiation of Short Electric Waves, *Proc. I.R.E.*, May, 1927.
Wells, N.: Aerial Resistance and Aerial Termination, *Marconi Rev.*, April, 1934.

Antenna Measurements:

Brown, G. H., and R. King: High-frequency Models in Antenna Investigation, *Proc. I.R.E.*, April, 1934.
The Problem of Auto-Radio Antennas, *Electronics*, February, 1935.
Clapp, J. K.: Antenna Measuring Equipment, *Proc. I.R.E.*, April, 1930.
Taylor, A. H., and H. F. Hastings: Determination of Power in the Antenna at High Frequencies, *Proc. I.R.E.*, August, 1931.
Radio Instruments and Measurements, *Bur. Standards Circ.* 74.

Directive Antennas for H-f Transmission and Reception:

Meissner, A.: Directional Radiation with Horizontal Antennas, *Proc. I.R.E.*, November, 1927.
Sterba, E. J.: Theoretical and Practical Aspects of Directional Transmitting Systems, *Proc. I.R.E.*, July, 1931.
———, and C. B. Feldman: Transmission Lines for Short-wave Radio Systems, *Proc. I.R.E.*, July, 1932.
Stone, J. S.: Directive Antenna Array, U.S. Patent 1643323, Sept. 27, 1927.
Yagi, H.: Beam Transmission of Ultra-short Waves, *Proc. I.R.E.*, June, 1928.

Broadcast Antennas:

Ballantine, S.: High-quality Broadcast Transmission and Reception, *Proc. I.R.E.*, May, 1934.
Chamberlain A. B., and W. B. Lodge: The Broadcast Antenna, *Proc. I.R.E.*, January, 1936.
Harbich, H., and W. Hahnemann: Wirksame Bekampfung des Nahschwundes in Rundfunksendeantennengebilde bestimmter Form, *E.N.T.*, October, 1932.
Harmon, Ralph: Some Comments on Broadcast Antennas, *Proc. I.R.E.*, January, 19
Hayes and Maclarty: The Empire Series Broadcast Station at Daventry, *Jour. I.E.E. (London)*, September, 1939.
Rickard: Graphical Method for Determining the Fundamental Wave Length of Broadcast Antenna, *Marconi Rev.*, November, 1933.
Roder, H.: Broadcast Antennas, *Broadcast News* (RCA), July, 1932.
Smith, Carl: A Critical Study of Two Broadcast Antennas, *Proc. I.R.E.*, October, 19

Aircraft Antennas:

Diamond, H., and G. L. Davies: Characteristics of Airplane Antennas for Radio Range Beacon Reception, *Proc. I.R.E.*, February, 1932.
Haller, G. L.: Constants of Fixed Antennas on Aircraft, *Proc. I.R.E.*, April, 1938.
Hyland, L. A.: Constants of Trailing Wire Antennas, *Proc. I.R.E.*, December, 1929.

SECTION 19

TELEVISION

By Donald G. Fink[1]

1. Definition. Television is the electrical transmission of transient visual images. Cathode-ray television makes use of electron beams or electron images in the camera tube (pickup device) and in the picture tube (reproducing device). A television system is considered to possess facilities for transmitting sound synchronously with visual images.

2. Elements of a Television System. The elements of a typical television system are shown in Fig. 1. The sound system consists of a

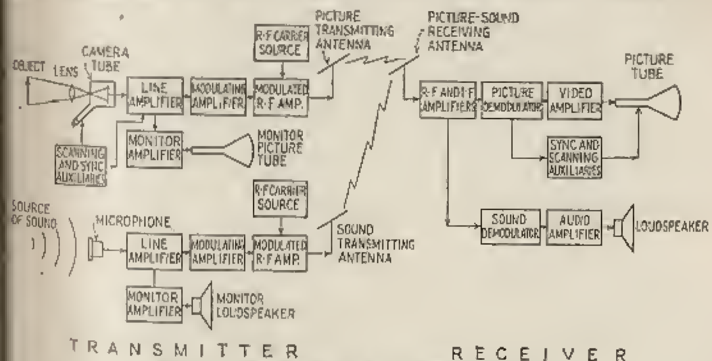


Fig. 1.—Elements of a television system.

conventional[2] transmitter and receiver operating on a carrier frequency in the u-h-f range and is separate from the picture system, except that common antennas may be employed at the transmitter and receiver and a common r-f amplifier and first detector may be used in the receiver. The picture transmitter includes the camera and synchronization circuits, (video signal generator), video amplifiers, u-h-f carrier source and r-f amplifiers, the modulator, a filter for suppressing part of one of the side-band regions in the carrier output, and the radiator. The picture receiver consists of r-f amplifier or antenna circuits, first detector and i-f amplifiers (the latter two in superheterodyne receivers), a second detector,

[1] Managing Editor, *Electronics;* author, "Principles of Television Engineering."

[2] The National Television System Committee has recommended (in 1941) the use of frequency modulation, with ±75 kc maximum deviation, for television sound transmission.

one or more video amplifiers, picture tube, synchronizing signal separat circuits, scanning generators, and power supplies.

SCANNING AND IMAGE ANALYSIS

3. Linear Scanning. The method of analyzing and synthesizing visu images employed in modern television systems is known as *linear scan ning*. As applied to the transmission of images, linear scanning involv the exploration of the image to be transmitted by an elemental spot small area, known as the *scanning agent*, which traverses the area of th image in a series of horizontal lines, moving over every point in the imag at constant speed and discovering the degree of brightness at each poi in succession. The camera tube, which includes the scanning agen generates a succession of electrical impulses which correspond with th successive values of brightness discovered by the scanning agent.

At the receiver the scanning process involves setting up an element luminous spot of small area which moves synchronously with the scannin agent in the camera tube. The brightness of this luminous spot is con trolled by the electrical impulses transmitted from the camera tube t the receiver. The values of brightness present in the original image a thereby reproduced in their proper positions. The scanning proce must be rapid enough so that all the elements of the received image a perceived simultaneously by the eye. This requirement is met if th scanning of the image is completed within the duration of persistence vision, so that the first element of brightness persists in the eye duri the production of all the succeeding elements in the image.

4. Aspect Ratio. The ratio of width (w) to the height (h) of the re tangle actively employed in reproducing the image is known as the *asp ratio*. In accordance with the standard adopted for motion pictures, the United States this ratio is given the value

$$\frac{w}{h} = \frac{4}{3} \qquad (1$$

NOTE: Those relationships marked with an asterisk (*) are recommended standa of the Radio Manufacturers Association (R.M.A.), which were used in 1940 for pub television transmissions in the United States. In 1941 the National Television Syst Committee (N.T.S.C.) recommended standards identical to those of the R.M.A. exc (1) frequency modulation for sound transmissions, (2) minor differences in the sy chronization wave form, (3) a higher modulation capability in the picture transmi (4) an increase in the number of lines from 441 to 525, and (5) the possible use of f quency modulation for synchronization.

These standards were adopted by F.C.C. for commercial televisi effective July 1, 1941.

5. Total Number of Lines per Frame. The total number of lines o which the scanning agent passes from the beginning of one compl image to the beginning of the next is known as the total number of li per frame, n.

The number of lines determines the degree of detail which may accommodated in the reproduced picture, in the vertical dimensi Hence this number sets an upper limit to the amount of pictorial det which may be accommodated in that direction. The number in mode

systems is set usually between 400 and 600 lines. According to the N.T.S.C. standards, n has the value

$$n = 525 \qquad (2)$$

The reason for the exact number 525 (see Arts. 6 and 31 of this section) is that it is an odd number composed of simple odd factors

$$(525 = 3 \times 5 \times 5 \times 7).$$

6. Interlaced Scanning. To reduce flicker in the reproduced image, a scanning technique known as *interlacing* is customarily employed, whereby the image is scanned in two or more groups of lines. The scanning motion in "two-to-one odd-line" interlaced scanning (the method now universally adopted) is shown in Fig. 2. The scanning agent traverses the area in two series of lines, alternately, passing downward (at left in Fig. 2) from point A to point B in the light solid line, following the back-and-forth motions shown by the arrows. The scanning spot then moves upward from point B to point C (at right), thence downward again from point C to point D on the heavy line, finally upward again from point D to the starting point A, where the motion repeats itself.

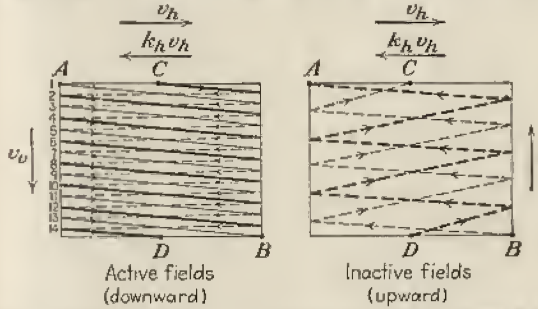


FIG. 2.—Scanning pattern for two-to-one odd-line interlaced scanning.

The scanning agent is *active* in discovering or reproducing the picture information while traveling over the lines shown solid and is inactive while traveling over the lines shown dashed. The total number of back-and-forth motions made in traversing both series of lines is n. The total number of active lines (shown solid) is n°. The inactive lines ($n - n_a$) are those made by the scanning agent in traveling from the bottom to the top of the picture area (the motions shown at the right). Practical values of the number of active lines, n_a, for a 525-line image lie between the limits 483 and 488, representing 42 and 37 inactive lines, respectively. The general expression for the number of active lines is

$$n_a = \frac{n}{1 + \frac{1}{k_v}} \qquad (3)$$

where k_v, the *vertical retrace ratio*, is the ratio between the upward scanning velocity and the downward scanning velocity, as defined in Art. 7 on the next page.

7. Scanning Velocities and Retrace Ratios. The scanning agent is caused to traverse the picture area in the interlaced pattern (Fig. 2) by imparting to it horizontal and vertical motions. The spot is displaced horizontally from left to right at a speed v_h, and simultaneously it is displaced vertically downward at a slower speed v_v. The two motions cause the spot to move slightly downward and to the right until it reaches the right-hand edge of the area. Then the spot is reversed rapidly and is moved to the left at a faster speed k_hv_h (k_h times as fast as it moves to the right), forming the retrace motion to the left-hand edge of the area. The downward velocity persists at the value v_v during the succession of back-and-forth motions until the spot reaches the bottom of the area. Thereupon the downward motion is reversed, and the spot is moved upward at a faster velocity k_vv_v (k_v times as fast as it moved downward), until it reaches the top of the pattern. During the upward motion, several back-and-forth motions are executed, since the horizontal velocities v_h and k_hv_h are maintained.

The *horizontal retrace ratio* k_h is the ratio of the backward (to the left) scanning velocity to the forward (to the right) scanning velocity. Practical values of k_h range from about 6 upward to 15. The N.T.S.C. standards set a lower limit to this ratio of 6.3. The *vertical retrace ratio* k_v between the upward and downward velocities, ranges from about 10 to 15 times. The N.T.S.C. standards set a lower limit to this ratio of 12.3.

8. Interlaced Fields. One set of the two sets of lines in the interlaced pattern is known as an *interlaced field*. Since the total number of lines in the complete frame is an odd number (525), the number of lines per field is a whole number plus one-half (262½). This accounts for the fact that at the end of the first field (Fig. 2) the spot, at point C, has formed but one-half of the horizontal motion. The half-line displacement causes the lines in the second field to be displaced vertically from those in the first field by the width of one line. Consequently the lines of one field fall directly between the lines of the preceding field. If the scanning motion is not accurately timed and if the amplitudes of the vertical and horizontal motions are not constant, this interlaced relationship is not maintained, and the lines of one field tend to overlap the lines of the preceding field. This defect is known as *pairing* of the fields. Its effect is to reduce the detail of the reproduced picture in the vertical dimension.

9. Vertical Resolution. The vertical resolution r_v of the scanning pattern is measured by the number of pictorial details or *picture elements* which may be accommodated in the vertical height of the picture area. Each active scanning line is capable of reproducing one such picture element in the vertical direction, but, since the picture elements in the image to be transmitted may not fall directly on the scanning lines, the actual number of picture elements which may be accommodated vertically is less than the number of active scanning lines. The vertical resolution r_v is accordingly the number of active scanning lines multiplied by a factor less than one, here called the utilization ratio k. The vertical resolution is accordingly

$$r_v = kn_a \text{ elements per picture height} \qquad (4)$$

Practical values of utilization ratio, depending on the method of measurement and the perfection of interlacing, range from about 0.6 to 0.9. With $n_a = 485$, r_v accordingly varies from 290 to 440 elements per picture

height. The value $r_v = 400$ is commonly reached in properly operated equipment.

10. Horizontal Resolution. The horizontal resolution r_h of the scanning pattern is measured by the number of picture elements which may be accommodated in the horizontal direction, measured in a width equal to the picture height. The picture height is used as the basis to make the horizontal resolution directly comparable with the vertical resolution. The total number of picture elements accommodated in the picture width is the horizontal resolution multiplied by the aspect ratio.

The value of r_h does not depend on the dimensions of the scanning pattern but rather on the electrical performance of the television system in reproducing rapid changes of voltage whereby the reproducing scanning agent is changed in brilliance as it moves across each line. In terms of the maximum frequency f_{max} in the video range (see Art. 16), the horizontal resolution is approximately

$$r_h = 84f_{max} \text{ elements per picture height} \qquad (5)$$

where f_{max} is expressed in megacycles. This expression assumes transmission at a rate of 30 frames per second and 525 lines. At 441 lines the expression is $r_h = 100f_{max}$.

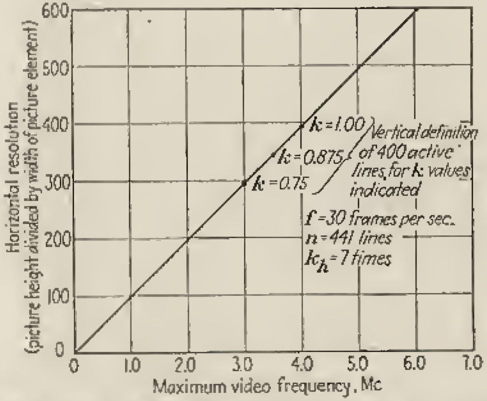


FIG. 3.—Relationship between horizontal resolution and maximum video frequency (441-line image).

11. Resolution Ratio. The ratio of the horizontal resolution to the vertical resolution is the resolution ratio m:

$$m = \frac{r_h}{r_v} = \frac{84f_{max}}{kn_a} \qquad (6)$$

Unity resolution ratio (equal resolution in vertical and horizontal directions) is not essential for good reproduction, inasmuch as the resolution in one direction may exceed that in the other by 50 per cent or more without wasting the detail in the direction of higher resolution. In present practice the resolution ratio approaches 0.95, depending on the maximum frequency in the video range. For values of $f_{max} = 4$ Mc, $n_a = 485$ lines, and $k = 0.75$, the resolution ratio is 0.925 times.

12. Total Number of Reproducible Picture Elements. A significant figure of merit of the television system is the total number N of picture elements which may be accommodated in the picture area, *i.e.*, the product of the number of elements vertically r_v, times the number horizontally $(w/h)r_h$:

$$N = \left(\frac{w}{h}\right)r_h r_v = \left(\frac{w}{h}\right)(84 f_{max})(kn_a) = \left(\frac{w}{h}\right) m k^2 n_a^2 \tag{7}$$

For values of $(w/h) = \frac{4}{3}$, $f_{max} = 4.0$ Mc, $n_a = 485$ lines, and $k = 0.75$ ($m = 0.925$), the total number is $N = 165{,}000$ picture elements. Assuming unity utilization ratio ($k = 1.0$), we obtain the maximum number available with a 4.0 Mc video range, *viz.*, $N = 220{,}000$ picture elements. Performance above 200,000 picture elements is exceptional in the present state of the art.

13. Viewing-distance Relationships. The desirable viewing distance of a television image depends on the resolution available. If we assume

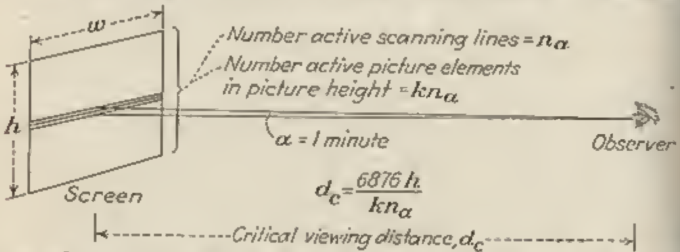


FIG. 4.—Critical viewing distance in terms of the dimensions of the scanning pattern.

a visual acuity of 1 minute of arc (typical of most normal eyes), two dark picture elements separated by a bright element (Fig. 4) may be barely resolved by the eye at a distance d_c.

$$d_c = \frac{6{,}876h}{kn_a} \tag{8}$$

and the corresponding ratio of critical viewing distance to picture height is

$$\frac{d_c}{h} = \frac{6{,}876}{kn_a} \tag{9}$$

For a vertical resolution $r_v = kn_a$ of 400 elements per picture height, the foregoing ratio is 17 times. This is the maximum viewing distance (17 times the picture height), beyond which the eye is unable to resolve the detail actually present in a stationary image.

The minimum viewing distance is determined by the tolerance of the viewer toward the structure of the picture, which becomes increasingly

evident as the viewing distance is decreased. Viewing distances shorter than 3 times the picture height are seldom considered satisfactory. A ratio of 5:1 seems to be typical of viewer habits. Figure 5 shows Eq. (9) for various numbers of scanning lines.

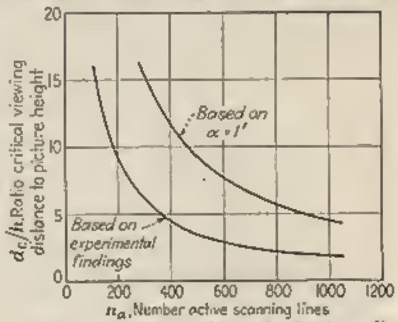


FIG. 5.—Relationship of viewing distance to number of scanning lines, in terms of the picture height. (*Experimental findings after Engstrom.*)

14. Frame-repetition Rate. The rate at which the frames are repeated (frame-repetition rate) depends (1) upon the duration of the persistence of vision of the eye and (2) upon the necessity of reproducing motion in the image in a smooth manner. In motion pictures the standard rate is 24 frames per second, with each frame projected twice, making 48 projection intervals per second. Similar values serve for television. However, since the power-supply frequency for most areas in this country is 60 cps, it is desirable to use a frame-repetition rate f which is a submultiple of the power frequency, *e.g.*, 30 per second (field repetition rate f' of 60 per second). The N.T.S.C. recommended standards for these items are accordingly

$$f = 30 \text{ frames per second} \tag{10*}$$
$$f' = 60 \text{ fields per second} \tag{11*}$$

15. Rate of Scanning Picture Elements. The maximum rate of scanning picture elements along each line depends on the number of elements in the line and the speed with which the line is scanned. These quantities in turn depend on the horizontal resolution (Art. 10) and on the number of lines per frame (Art. 5) and the rate of frame repetition (Art. 14). The general expression for the maximum rate of scanning picture elements R is

$$R = \frac{w}{h} m f k n^2 \frac{(1 + 1/k_h)}{(1 + 1/k_v)} \text{ elements per second} \tag{12}$$

where the quantities have been defined in the preceding sections. For aspect ratio $w/h = \frac{4}{3}$, resolution ratio $m = 0.925$, frame-repetition rate $f = 30$ per second, utilization ratio $k = 0.75$, number of lines per frame $n = 525$, horizontal retrace ratio $k_h = 7$ times, and vertical retrace ratio $k_v = 15$ times, the rate of scanning picture elements is approximately $R = 8{,}300{,}000$ elements per second, which is approximately the upper limit of performance of present-day equipment.

16. Maximum Frequency in Video Range. The maximum video frequency generated by the television camera is directly proportional to the rate at which the picture elements are scanned along each line. In deducing a relationship between the scanning rate R (Art. 15) and the maximum video frequency (v.f.), it is customary to assume that the picture elements are arranged as alternate black and white squares along the scanning line. An ideal scanning agent, scanning such a line, will produce a square wave, as shown in Fig. 6. The upper portion of each square wave represents a black element, the lower portion an adjacent white element. Hence there are two elements per cycle of the wave. The fundamental frequency of the square wave is accordingly one-half as great as the rate of scanning picture elements. The maximum v.f. is then derived from Eq. (12), as

$$f_{\max} = \frac{(w/h)mfkn^2}{2} \cdot \frac{(1 + 1/k_h)}{(1 + 1/k_v)} \tag{13}$$

For the conditions cited in Art. 15, $f_{\max}$ is 4.15 Mc. Table I gives other typical values. It should be noted that this frequency is the fundamental of the square wave. The reproducing equipment cannot reproduce a square wave of this frequency. Instead a sine-wave distribution of light is reproduced. This sine wave (Fig. 6) establishes the basic structure of the reproduced image.

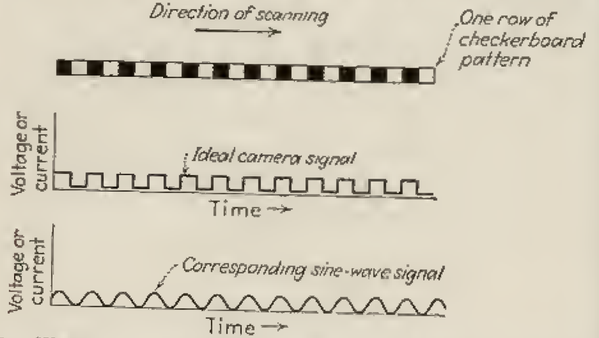


FIG. 6.—Wave forms resulting from scanning a checkerboard image. The ideal square wave becomes a sine wave when only the fundamental frequency is transmitted.

TABLE I.—MAXIMUM VIDEO FREQUENCIES FOR DIFFERENT SCANNING PATTERNS

Number of scanning lines n	Number of frames per second f	Maximum v.f. for equal vertical and horizontal resolution (m = 1.00), cps	Maximum v.f. for for horizontal resolution = 0.9 × vertical resolution (m = 0.9), cps
20	16	3,360	3,020
60	16	30,200	27,250
120	24	81,500	163,000
180	24	410,000	369,000
240	24	727,000	653,000
343 (7 × 7 × 7)	30	1,860,000	1,670,000
441 (3 × 3 × 7 × 7)	30	3,060,000	2,800,000
525 (5 × 5 × 7 × 3)	30	4,350,000	3,920,000
1029 (3 × 7 × 7 × 7)	30	16,650,000	14,800,000

NOTE: Calculation based on $w/h = 4/3$, $k_h = 7$, $k_v = 12$, $k = 0.75$.

17. Scanning Wave Forms. The deflecting forces necessary to produce the linear scanning motions shown in Fig. 2 are *saw-tooth waves*, as shown plotted against time in Fig. 7. The forward motion in the horizontal direction is produced by a deflecting force linear with time, and the retrace motion by a force which need not be linear but which must

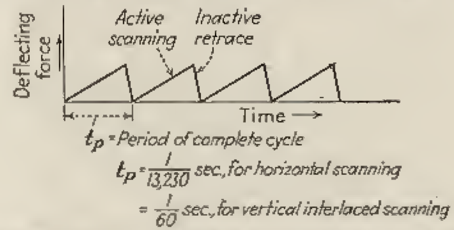


FIG. 7.—Saw-tooth waves of deflecting force used to produce the scanning pattern (441-line image).

have a rate of change high compared with that of the forward force. The ratio of the slopes is equal to the horizontal retrace ratio k_h (Art. 7). The same conditions apply to the deflecting force in the vertical direction, and the ratio of the retrace slope to the forward slope of this wave equals the vertical retrace ratio, k_v.

The scanning wave forms have fundamental frequencies determined by the number of fields per second and by the number of lines per second. In the vertical direction the scanning force must repeat itself at the field repetition rate, $f' = 60$ cps. In the horizontal direction the deflecting force must repeat itself at the line-scanning frequency (525 lines per frame, 30 frames per second), which is the product

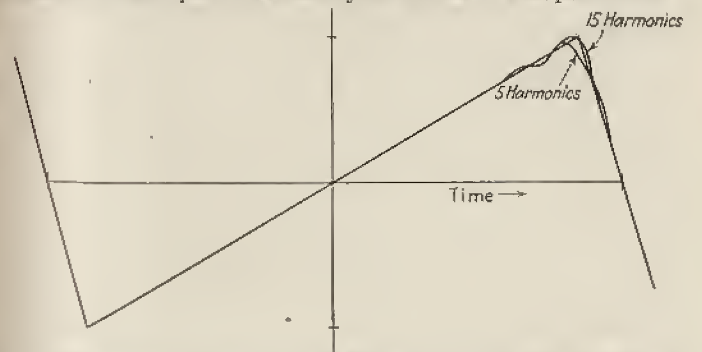


FIG. 8.—Ideal saw-tooth wave (k_h = 6.66) and approximations resulting from inclusion of 5 and 15 harmonics.

$$nf = 525 \times 30 = 15{,}750 \text{ cps}$$

These values of scanning frequency are the fundamentals of the saw-tooth wave. From 5 to 20 harmonics should be present if the wave form is to approximate the saw-tooth shape sufficiently accurately for scanni purposes. Figure 8 shows the degree of approximation for a saw-too wave having a slope ratio of 6.66 times (retrace ratio) when 5 and harmonics are included. The fifteenth harmonic extends the range to 900 cps for the vertical scanning system, and up to 250,000 cps the horizontal scanning system. Practical scanning generators ar discussed in Art. 67.

References

Scanning and Image Analysis:

BEDFORD, A. V.: Figure of Merit for Television Performance, *R.M.A. Eng.*, **2**, No. 1, November, 1937; also *RCA Rev.*, **3**, No. 1, 36, July, 1938.

ENGSTROM, E. W.: A Study of Television Image Characteristics, *Proc. I.R.E.*, Part **21**, 1631, December, 1933; Part II, **23**, 295, April, 1935.

JESTY and WINCH: Television Images—Analysis of Their Essential Qualities, *Jou Telev. Soc.*, **2**, 316, December, 1937.

KELL, BEDFORD, and TRAINER: Scanning Sequence and Repetition Rate of Televisi Images, *Proc. I.R.E.*, **24**, 559, April, 1936.

MERTZ and GRAY: Theory of Scanning and Its Relation to the Transmitted Signal Telephotography and Television, *Bell System Tech. Jour.*, **13**, 464, July, 1934.

SOMERS, F. J.: Scanning in Television Receivers, *Electronics*, **10**, No. 10, 18, Octobe 1937.

VON ARDENNE, M.: Distortion of Saw-tooth Waveforms, *Electronics*, **10**, No. 11, 2 November, 1937.

WHEELER and LOUGHREN: The Fine Structure of Television Images, *Proc. I.R.E.*, 2 540, May, 1938.

WILSON, J. C.: "Television Engineering," Chap. III, p. 46, and Chap. IV, p. 73, S Isaac Pitman & Sons, Ltd., London, 1937; extensive bibliography to periodicals an patents.

THE VIDEO-SIGNAL WAVE FORM

18. Video Signal. The video signal (or "composite video signal" is the succession of electrical impulses transmitted through the televisi system to convey the information from the scanning agent in the camer to the scanning agent in the receiver. Three direct functions are carri out through the video signal: (1) the transmission of impulses correspon ing to the brightnesses of the scanned picture elements, conveyed by th *camera signal;* (2) the blanking of the scanning agent at the receiver du ing the retrace motions, by the *blanking level* or *pedestal;* and (3) th synchronization of the scanning agents, by the *vertical* and *horizont synchronization signals.* The first item of the video signal is generated the camera, the second two in the *synchronization signal generator.* Th three items are combined in the *video mixing amplifier.*

19. Envelope of the Modulated Picture-carrier Signal. When t video signal is imposed on a carrier wave, the envelope of the modulat carrier wave constitutes the video-signal wave form. Such a modulat picture carrier and the details of the envelope are shown in Fig. 9. Th particular form of video signal shown is that recommended in the stan ards of the R.M.A. (practically identical to the N.T.S.C. proposal).

In the R.M.A. standard video signal (Fig. 9) the carrier amplitude divided by the black level (blanking level or pedestal) at a value fro 75 to 80 per cent (75 ± 2.5 per cent according to the N.T.S.C. recom mendation) of the maximum amplitude. The amplitude region abo the black level is called the *infra-black* region and is occupied by th synchronizing signals. Signal levels in this region do not produ light in the received image. The synchronizing signals are of two typ horizontal signals (Fig. 9) for initiating the motion of the scanning age along each horizontal line and vertical signals (Fig. 12) for initiating th

motion of the scanning agent vertically at the beginning of each field. The peak amplitude of the wave, the height of the synchronizing pulses, and the black-level amplitude are maintained constant throughout each broadcast at the values shown in Fig. 9.

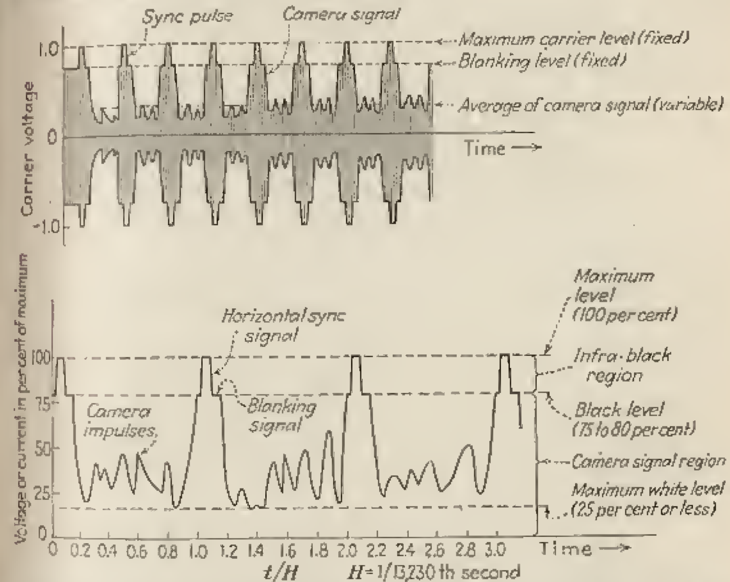


FIG. 9.—Top, modulated television carrier signal. Bottom, details of modulation envelope, according to R.M.A. recommended standard. According to the N.T.S.C. recommendations the black level is 75 ± 2.5 per cent and the maximum white level 15 per cent or less.

20. Camera Signal. The portion of the carrier envelope extending below the black level is called the *camera signal.* The polarity of transmission in the N.T.S.C. standards is negative, *i.e.*, an increase in the light on the camera plate results in a decrease in the carrier amplitude, as shown in Fig. 9. The maximum white level is 25 per cent or less (15 per cent or less according to the N.T.S.C. recommendation) of the maximum carrier amplitude. Intermediate gray tones exist between the maximum white level and the black level.

The camera signal has two components (Fig. 10): an a-c component, which describes the variations in brightness from the average brightness; and the d-c component, averaged over the frame-scanning interval ($\frac{1}{30}$ sec.), which represents the average or background brightness of the picture. The a-c and d-c components must be capable of being varied independently of each other, so that the same detail may be presented either on a dark background or on a bright background. Variation of the d-c component also permits the screen brightness to be "faded in" or "faded out" at the will of the studio operator. In order that the d-c component be independent of the a-c component, regardless of the changes in wave form, it is necessary that the black level be constant in the carrier envelope, and furthermore that the black level be maintained constant at the control grid of the picture tube (see Art. 61).

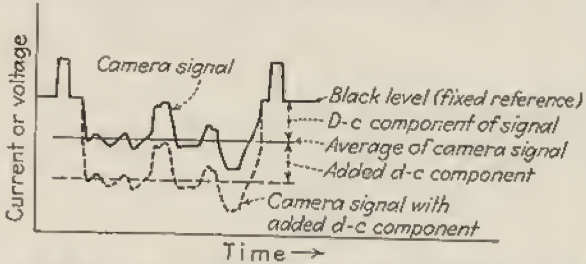


FIG. 10.—Both d-c and a-c components of the modulation envelope. The dashed line represents an increase in the background brightness without change in detail.

21. Frequency Range in the Video Signal. The maximum frequency in the video range (Art. 16) results from scanning the finest detail in the image, *i.e.*, from the scanning of adjacent picture elements. The value of f_{max} [Eq. (13)] depends on the rate at which adjacent picture elements are scanned; values up to 4 or 5 Mc are commonly employed in present equipment.

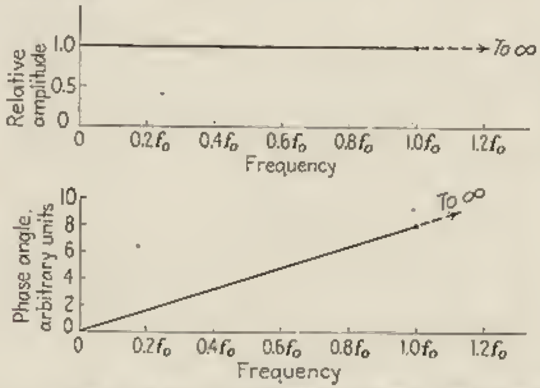


FIG. 11.—Ideal amplitude and phase transmission characteristics.

The lowest frequency in the video range, f_{min}, depends on the rate at which the background brightness of the scene changes. Brightness changes which take longer than the duration of a single frame to complete themselves are usually introduced by changes in the d-c component of the signal. Changes that take less than the duration of a single frame are accommodated by video frequencies extending downward to 30 cps (corre-

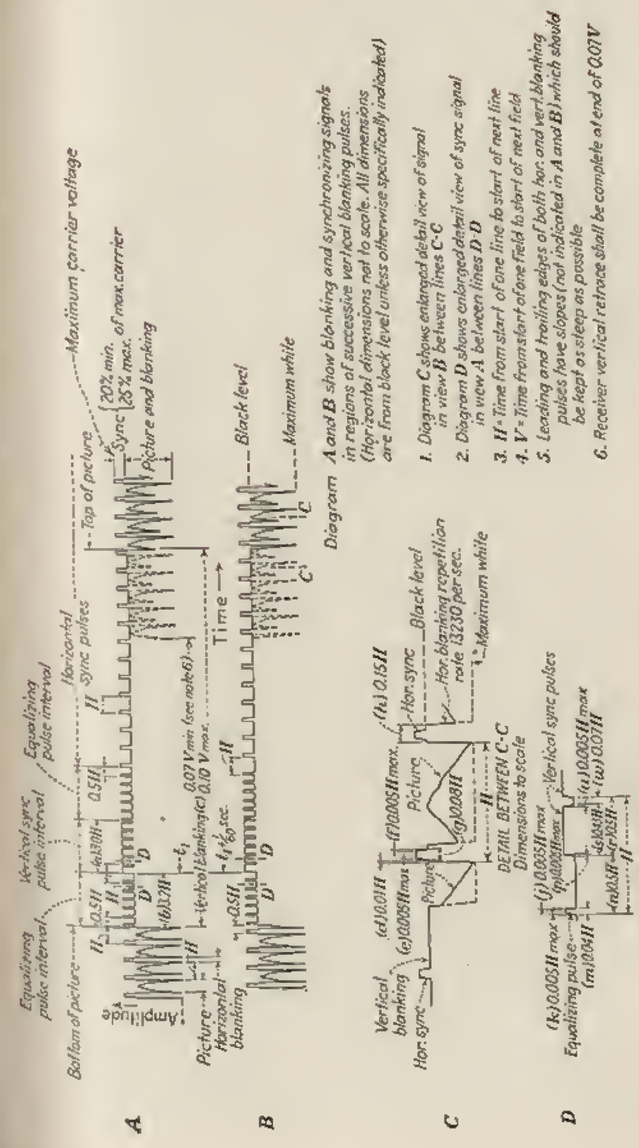


FIG. 12.—The dimensions of the R.M.A. standard video-signal wave form. The N.T.S.C. version is identical to the R.M.A. except for the blanking level (75 ± 2.5 per cent) and minor changes in the dimensions. The N.T.S.C. recommendations also impose tolerances on the rate, as well as the rate of change of recurrence, of the line sync pulses.

sponding to the frame-repetition rate of 30 per second). Consequently the significant frequency range in the video signal, based on the N.T.S.C. standards, is from 30 cps to 4 or 5 Mc.

Degrees of picture detail intermediate between the whole frame area and the area of a single picture element are reproduced by frequencies intermediate between 30 cps and 4 or 5 Mc. Since such intermediate degrees of detail may be present in any scene, the video-signal transmission system must be equally responsive to all frequencies within these limits.

22. Requirements for Transmission of the Video Wave Form. Fourier analysis of wave forms reveals that any wave form encountered in practice is composed of a number of sine-wave components having specified relative amplitudes and specified relative phases. If the wave form is to be reproduced accurately, the transmission system must be capable of transmitting all such sine-wave components, throughout the v-f range, without altering the relative amplitudes and phases of the components. This requirement is met if the amplitude-*versus*-frequency response curve of the transmission system is a horizontal straight line over the v-f range, and if the phase-*versus*-frequency curve of the transmission system is an oblique straight line passing through the zero-frequency point and extending over the v-f range. The ideal characteristics are shown in Fig. 11.

If the amplitude transmission characteristic is not ideal, the wave form is distorted symmetrically about a vertical axis. If the phase transmission characteristic is not ideal, the wave form is distorted asymmetrically about a vertical axis. Inadequate h-f response produces improper reproduction of steep vertical changes in the wave form. Inadequate l-f response produces improper reproduction of the flat top portions of the wave which extend over intervals comparable with the period of the low frequencies.

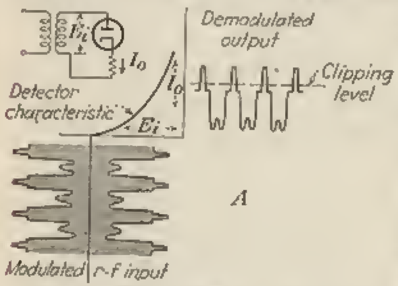


Fig. 13.—Demodulation of the modulated picture carrier by a diode detector, showing clipping level which separates the camera signal from the synchronizing pulses.

23. Dimensions of Synchronizing Pulses. The dimensions of the sync pulses in the R.M.A. recommended standard wave form are shown in Fig. 12. There are three distinct types of sync pulse. The *horizontal sync pulse* exists on the blanking pulse between the scanning of each line and occupies a duration of 8 per cent of the duration of the line-scanning interval. The *vertical sync pulse* exists on the blanking impulse between the scanning of successive fields, and extends for 3 times the duration of the line-scanning interval. *The equalizing sync pulses* exist immediately before and after the vertical sync pulse in two groups of six pulses each. The frequency of the equalizing pulses is twice that of the horizontal sync pulses.

The horizontal scanning generators at transmitter and receiver are usually synchronized by the leading edge of the horizontal pulses. Since the synchronizing action must be precise, the leading edge must be sharp.

The rise of this edge must complete itself in one-half of 1 per cent of the line-scanning interval.

The vertical scanning generators at transmitter and receiver are usually synchronized by the integrated effect of the equalizing and vertical sync pulses which are used to charge a condenser. The leading edge of the condenser charge curve (Fig. 51) acts as the synchronizing agent. This leading edge must have precisely the same shape for each vertical pulse. The equalizing pulses are inserted to ensure that this condition is met equally for fields ending on a half line and for fields ending on a whole line.

GENERATION OF THE VIDEO-SIGNAL WAVE FORM

24. Video-signal Generator. The video-signal generator consists of three essential parts: (1) the camera and its auxiliaries, which generate the camera signal component; (2) the synchronizing signal generator, which times and shapes the vertical and horizontal synchronizing signals and the blanking signals; and (3) the control amplifier which mixes the camera signals with the synchronizing signals and the blanking signals, forming the composite video signal.

25. Television Cameras. The television camera consists of a light-tight housing fitted with an adjustable camera lens which focuses the scene on the photosensitive plate of the camera tube enclosed within the housing. Also enclosed in the housing is a preamplifier which raises the level of the camera signal (usually to about 0.1 volt peak to peak) so that it can be transmitted over coaxial cable without interference. One or more scanning generators or scanning amplifiers may also be included in the camera housing. The camera is ordinarily mounted on a flexible standard so that it may be moved readily, and a universal mounting is provided so that the camera may be directed at any angle. Some form of auxiliary optical system is also provided to enable the operator to keep the image in focus.

At present, there are four important types of electronic camera tubes: (1) the iconoscope; (2) the orthiconoscope (orthicon); (3) the image iconoscope (iconotron or superemitron); and (4) the image dissector. The first three employ the storage principle whereby the effect of the incident light is stored as charge across a capacitance element. The image dissector is an instantaneous device, using only that light present on each picture element at the instant it is scanned. The storage devices display a luminous sensitivity from 10,000 to 100,000 times that of the instantaneous devices, depending on the storage and photoelectrical efficiencies and the number of reproducible picture elements (Art. 12).

26. The Iconoscope. A typical iconoscope is shown in Fig. 14 together with its optical and electrical auxiliaries. The image is focused on the mosaic plate, which is a mica sheet coated with several million globules of photosensitized silver, insulated from each other and from a graphite coating on the reverse side of the plate. The optical image releases electrons from the mosaic, thereby charging the plate positively with a charge distribution corresponding point for point with the distribution of light in the image. The insulation prevents redistribution of this charge and permits the charge image to increase in magnitude for as long as the light falls on the mosaic.

The mosaic is scanned by a beam of electrons generated in the electron gun in the side arm of the tube. The beam, impinging on the mosaic, releases secondary electrons. The number of secondary electrons released from a given point of the mosaic depends on the potential of that point, which in turn depends on the previous photoelectric emission from that point. Consequently, as the scanning agent passes over the mosaic, it generates a secondary emission current which corresponds to the successive values of brightness in the picture elements. The secondary emission is small for brightly illuminated portions of the mosaic;

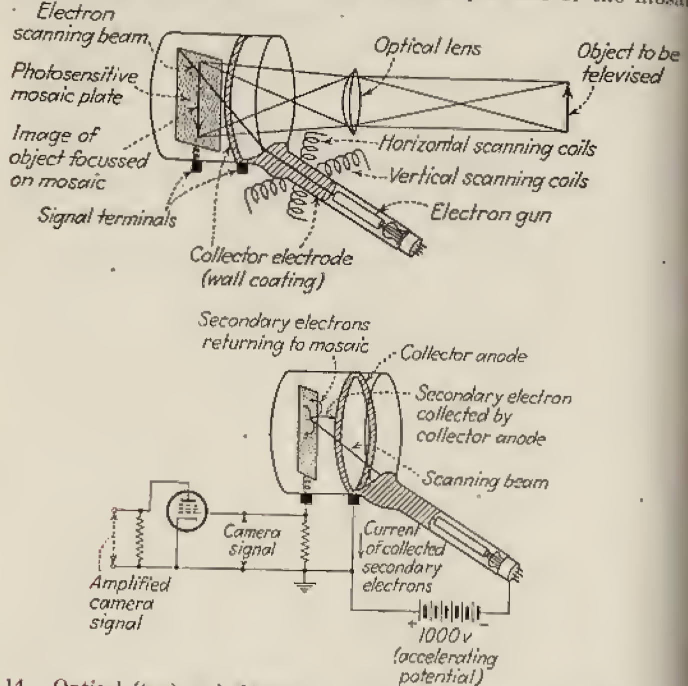


FIG. 14.—Optical (top) and electrical (bottom) arrangements used with the iconoscope camera tube.

consequently the output current is "negative" with respect to the illumination responsible for it.

The secondary emission is collected by a collector anode, and conducted through an external coupling resistor back to the graphite signal plate on the back of the mosaic support. The series circuit through which the electron current passes is accordingly composed of the ohmic resistance of the secondary emission path, the coupling resistance, and the capacitance between the signal plate and the group of globules under the scanning agent. No d.c. can flow through the capacitance; hence the output consists simply of the a-c component of the camera signal. The d-c component must be evaluated either by visual observation or by a

phototube which integrates the light on the scene. A d-c voltage derived from a manual control (or from the phototube in the second case) is inserted in series with the output of the iconoscope.

Since the mosaic is insulated, the current flowing toward or away from it must be zero, when averaged over any extended period of time. The average d-c value of the collected secondary emission must accordingly be replaced by electrons from the scanning beam.

Only a part of the secondary emission is collected from the mosaic. The remainder, falling back on the mosaic, sets up a distribution of charge which, when scanned, produces a spurious signal whose effect is to produce an unevenness in the background shading of the reproduced picture. This spurious signal ("dark-spot signal") must be compensated by a *shading-correction* signal generator (Art. 35).

The color response of the typical iconoscope mosaic (when the mosaic has been silver sensitized) is very similar to that of the usual panchromatic negative-film emulsion used in motion pictures.

The sensitivity of modern iconoscopes under optimum conditions varies from about 1 mv per millilumen per square centimeter illumination on

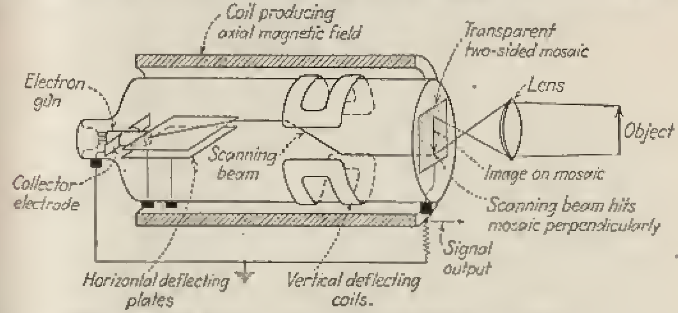


FIG. 15.—Structure of the orthiconoscope ("orthicon").

the mosaic surface (low values of illumination) to about 0.25 mv (at higher illumination). The curve between input illumination and output voltage of the typical iconoscope is characteristized by a gamma (Art. 72) of about 0.7, *i.e.*, the curve is of the saturating variety. The output voltages may be increased by increasing the current used in the scanning beam, as well as by illuminating the interior of the tube envelope by a bias light.

27. The Orthiconoscope. The orthiconoscope (Fig. 15) operates similarly to the iconoscope except that low-velocity electrons are used for scanning. Consequently no observable secondary emission effects arise, and no spurious "dark-spot" signal is generated. The scanning electrons themselves are collected and passed through the coupling resistor back to the mosaic. A two-sided mosaic is used.

The photoelectric emission from the mosaic is saturated in the orthiconoscope; consequently the relationship between input illumination and output voltage is linear (the gamma is unity). The sensitivity of current models is about 2 mv per millilumen per square centimeter on the mosaic, although theoretical sensitivities as high as 10 mv are possible.

To use low-velocity electrons for scanning without incurring defocusing of the beam, it is necessary that the scanning beam impinge perpendicularly on the mosaic at all points in the scanning pattern. This requirement is met by a rather unorthodox deflection technique which employs a combination of axial magnetic field and transverse electric field for horizontal scanning and a transverse magnetic field for vertical scanning. The high sensitivity, freedom from dark-spot signal, and convenient optical arrangement of the orthiconoscope have made it a serious competitor of the iconoscope in current broadcasting practice.

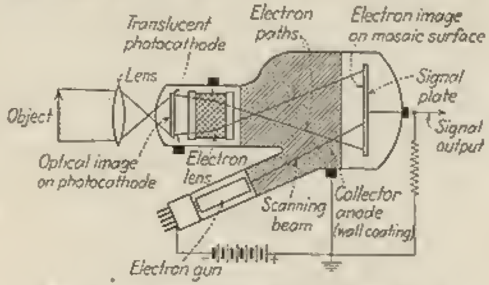


FIG. 16.—Structure of the image iconoscope.

28. Image Iconoscope. The image iconoscope has not found wide use in this country but has been used in Great Britain. An American version is shown in Fig. 16. The scene is focused on a photocathode which releases an electron image of the scene. The electron image is conveyed down the tube and brought to focus at the surface of a secondary emission surface that acts also as a storage mosaic. The high secondary emission ratio of this surface produces a correspondingly high value of charge stored on its surface, and the charge increases throughout the frame-scanning interval. A conventional scanning beam is used which induces further secondary emission in a manner exactly analagous to that in the iconoscope. The initial secondary emission, resulting from the arrival of the electron image, produces a higher value of stored charge than in the iconoscope and the sensitivity is proportionately increased. Values of sensitivity as high as 5 mv per millilumen per square centimeter have been found in typical tubes.

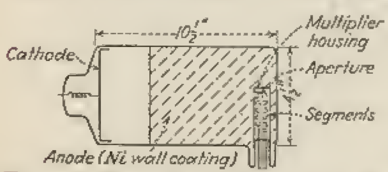


FIG. 17.—Structure of the image dissector.

29. Image Dissector. The image dissector (Fig. 17) is used principally for the televising of motion-picture film, where the light source may be brilliant and highly concentrated. For general pickup work, the low sensitivity of the device is a disadvantage when compared with storage pickup tubes.

The image dissector consists of a cylindrical envelope with an optical window at one end through which the image is admitted to the photocathode

at the opposite end. Here an electron image is generated and drawn to the opposite end of the tube where it is focused in the plane of the scanning aperture. The aperture is fixed on the end of a finger support. The image is moved past the scanning aperture by transverse magnetic fields applied from coils external to the tube. Inside the finger an electron multiplier structure is employed to increase the sensitivity of the device before the signal current (composed of the electrons entering the aperture) is applied to the coupling resistor. With this amplification, the signal/noise ratio of the output current is 10:1 when the mosaic illumination is 200 foot-candles. The sensitivity, when used with a 11-stage multiplier is about 50 μv per millilumen per square centimeter on the photocathode, at a signal/noise ratio of 5:1. No secondary emission effects are observed. The output-input curve is linear (gamma unity). Also the output current contains a d-c component which is directly proportional to the average brightness of the scene. Hence no auxiliary evaluation of the d-c component is necessary.

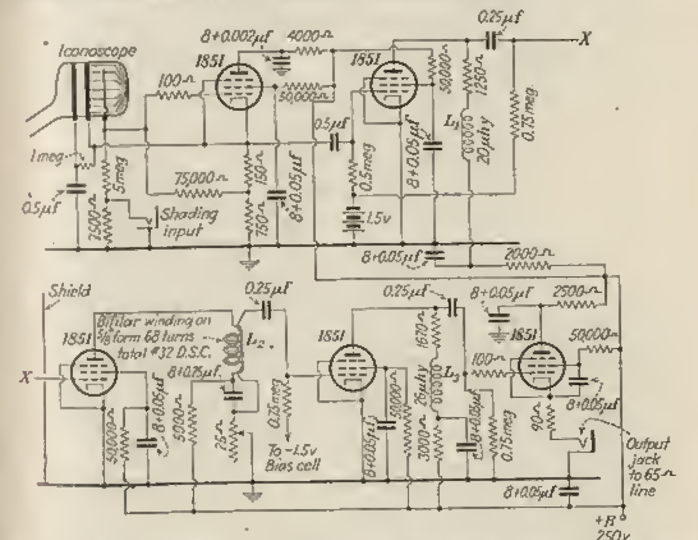


FIG. 18.—Circuit diagram of an iconoscope preamplifier. (*After Barco.*)

30. Preamplifier. A typical preamplifier for use with an iconoscope is shown in Fig. 18. To preserve a high signal/noise ratio in the first stage, an effective value of about 300,000 ohms is used as the coupling resistance, with a shunt capacitance of about 8 $\mu\mu$f. The poor h-f response incident to this combination is compensated in the third stage, which employs a bifilar winding (L_2) to remove the effect of the impedance in the power supply. The second and fourth stages are conventional video-amplifier stages (Art. 38) with flat response to 5 Mc. The output stage is a cathode-coupled stage having less than unity gain and presenting an output impedance which matches the characteristic impedance (65 ohms) of the coaxial cable. The camera signal is sent over this cable to the control amplifier for mixing with the synchronization impulses.

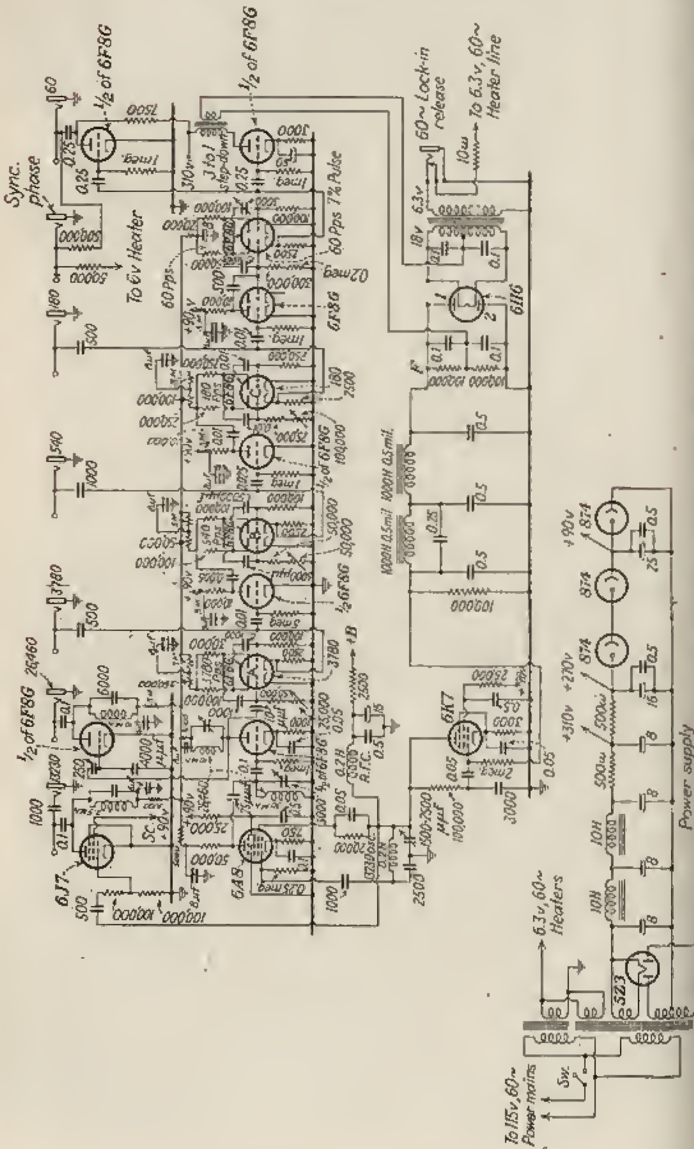


FIG. 19.—Typical synchronization signal generator (timing unit) (441-line image). (After Deal.)

The output of the shading-correction generator (Art. 35) is inserted directly in series with the signal plate of the mosaic.

31. Synchronization Signal Generator. *Timing Unit.* The sync pulses (Fig. 12, Art. 23) must be properly timed and properly shaped. The timing function is carried out in a timing unit, a typical example of which is shown in Fig. 19. The unit produces two outputs at 60 cps and at 13,320 cps (for the R.M.A. standard video signal). The 60 cps output is derived from the basic 13,230 cps oscillation by frequency multiplication and division; multiplication to 26,460 cps and division in four steps of 7 (to 3,780 cps), 7 (to 540 cps), 3 (to 180 cps), and 3 (to 60 cps). Frequency multiplication is carried out in a frequency converter tube, the divisions in multivibrators isolated by buffer stages.

The locally generated 60 cps signal is then compared with the 60 cps voltage of the power system by feeding the two sources to a discriminator diode which develops a d-c voltage proportional to the amount and direction of the phase difference between the two sources. This d-c is used in an a-f-c circuit to correct the frequency of the basic 13,230 cps oscillator from which the locally generated 60 cps is derived. In this way the 13,230 and 60 cps outputs are maintained in synchronous relationship with each other and with the frequency of the power system.

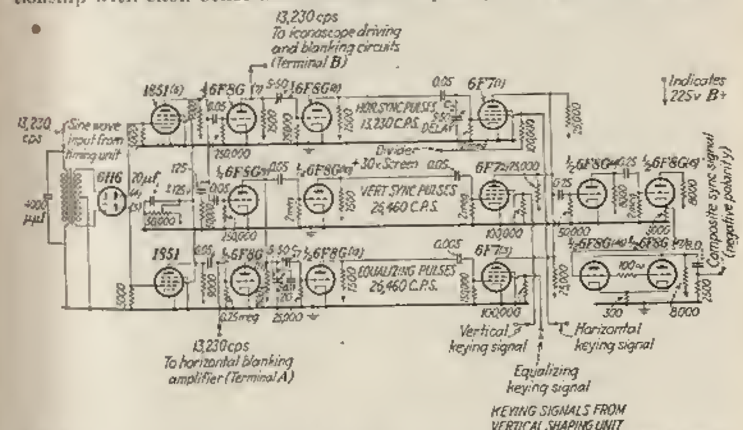


FIG. 20.—Typical horizontal shaping unit of a synchronization signal generator (441-line image). (After Deal.)

32. Synchronization Signal Generator. *Horizontal Shaping Unit.* Figure 20 shows one portion of the shaping unit of the synchronizing signal generator, the horizontal shaping unit. This unit accepts the 13,230 cps output of the timing unit and produces from it the several wave forms required for the R.M.A. standard signal (Fig. 12, Art. 23). The upper chain of tubes produces the horizontal sync pulses at 13,230 cps, the successive tubes being employed to obtain the required duration, shape, and steepness of front required for these pulses. The middle chain of tubes produces the serrated vertical sync pulses continuously at 26,460 cps. The bottom chain produces equalizing pulses at 26,460 cps. All three types of pulse are produced continuously in this portion of the generator. They are interspersed in the proper order (Fig. 12) by the action of keying signals in the 6F7 tubes at the right of each chain. The interspersed signal (composite synchronizing signal) is then amplified by the stages at the extreme right and is applied to the control amplifier for mixing with the camera signal.

33. Synchronization Signal Generator. *Vertical Shaping Unit.* The vertical shaping unit (typical example shown in Fig. 21) has the function

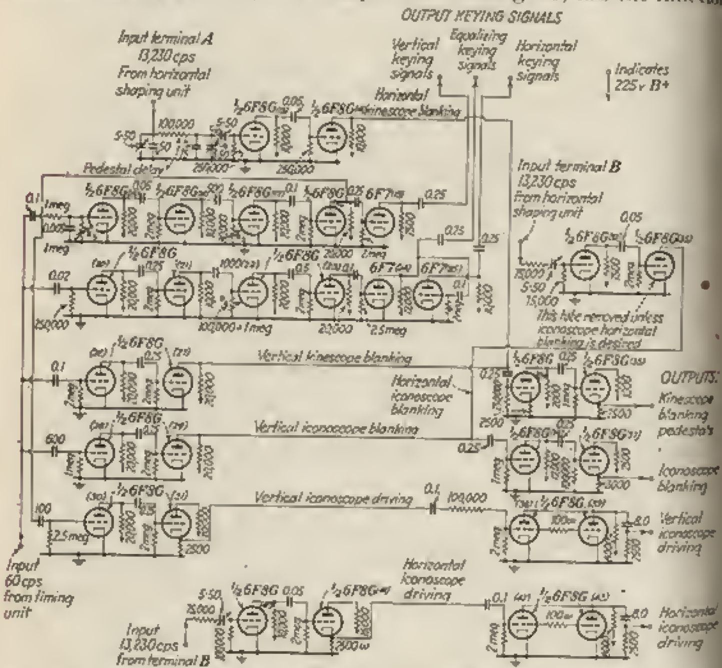


FIG. 21.—Typical vertical shaping unit of a synchronization signal generator (441-line image). (*After Deal.*)

of producing so-called *keying signals* at a frequency of 60 cps. These keying signals are applied to the screen grids of the keying tubes in the horizontal shaping unit. The action of the keying signals is to allow to pass, or to block, the synchronizing signals passing through the keying tubes. Thus a negative keying signal is required to block the passage of the horizontal synchronizing signals during the vertical blanking period (Fig. 12); a positive keying signal is required to allow the serrated vertical sync pulse to pass at the proper time during the field blanking interval; and a two-part positive keying signal is needed to allow the equalizing pulses to pass immediately before and after the vertical sync pulses. The shape of the keying signals and the synthesis of the composite synchroniz-

ing signal are shown in Fig. 22. The vertical shaping unit accepts the 60 cps output of the timing unit and forms the required keying signals by several chains of shaping tubes which introduce the necessary wave shaping and delaying actions.

The vertical shaping unit also provides blanking signals which are applied to the control amplifier to introduce the black level during the retrace periods. Two sets of blanking signals are generated: one for the composite video signal and another, of somewhat shorter duration, for the control of the scanning beam in the camera tube. Each group of blanking signals consists of horizontal and vertical square waves recurring at 13,230 and 60 cps, respectively.

The camera-tube scanning generators are controlled by vertical and horizontal driving impulses, which are somewhat narrower and sharper than the corresponding sync pulses in the composite video signal. These

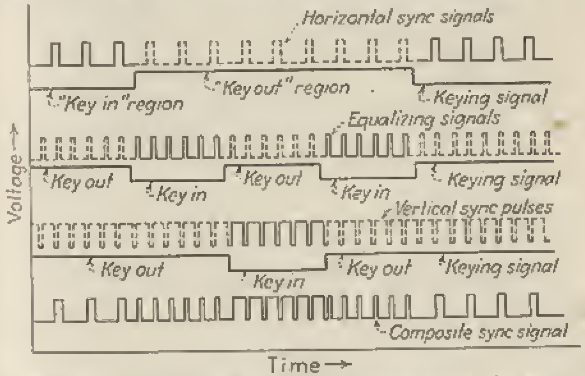


FIG. 22.—Function of the keying signals in interspersing the components of the composite synchronization signal.

driving pulses are formed from 13,230 and 60 cps signals derived from the horizontal and vertical shaping units, respectively.

34. Mixing Amplifier. The mixing amplifier (Fig. 23) has three input terminals which accept the camera signal from the camera preamplifier, the blanking signals from the synchronizing generator, and the composite synchronizing signals from the synchronizing generator. The camera signal and blanking signals are first combined by means of two amplifier tubes feeding a common load resistor, across which the "semicomposite" signal appears. The d-c component of the video signal is controlled by varying the bias on the blanking signal amplifier tube, thus controlling the amplitude relationship between the average of the camera-signal component and the blanking level.

The composite synchronizing signal is similarly added to the camera and blanking components in two amplifier stages feeding a common load resistor, across which the composite video signal appears. Bias controls across these tubes control the relative amplitude of the camera and synchronizing signal amplitudes, thus allow the establishment of the 75-25 per cent relationship demanded by the standard signal.

35. Shading-correction Generator. The shading-correction generator is a device for producing wave shapes of saw-tooth, sine, and parabolic

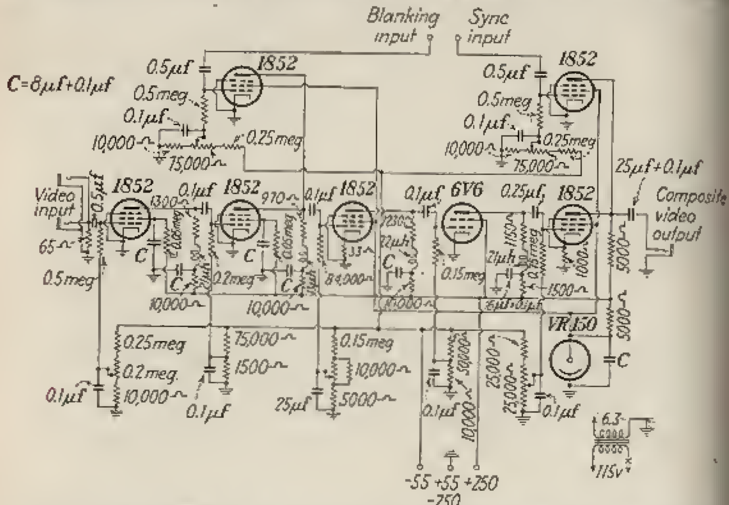


Fig. 23.—Typical mixing video amplifier for combining camera signal, blanking signals, and composite synchronization signals. (*After Barco.*)

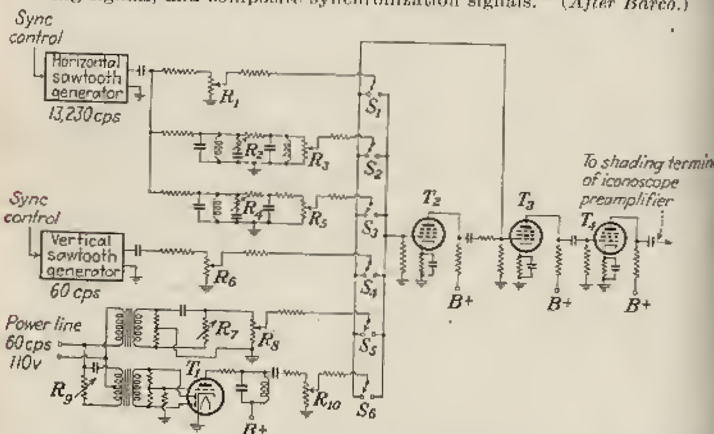


Fig. 24.—Typical shading-correction signal generator. (*After Bedford.*)

shape at vertical scanning and horizontal scanning rates (60 and 13,230 cps, respectively) in synchronism with the scanning motion. These wave shapes, controlled as to amplitude, phase, and polarity, are introduced in the preamplifier (Fig. 18) to compensate the spurious shading signal generated in the iconoscope. A form of shading-correction generator is shown in Fig. 24. The horizontal saw-tooth generator used to deflect the beam in the iconoscope tube is used directly to produce saw tooths of controlled amplitude and polarity, as well as 13,230 and 26,460 cps sine waves of controllable amplitude, phase, and polarity. Similarly the output of the vertical saw-tooth generators is used to produce similar signals. Reversal of polarity is provided by an amplifier stage. The switches control the polarity, resistors R_1, R_3, R_5, R_6, R_8, and R_{10} control amplitude, and resistors R_2, R_4, R_7, and R_9 control the phase. Methods of producing saw-tooth waves of controllable phase are also available, using "clipped-off" portions of the basic saw-tooth waves. The shading-correction generator controls are manipulated manually to correct for the observed defects of shading in the image as viewed on the monitor picture tube.

References

The Video Signal and Its Generation, Including Camera Tubes:

Barco, A. A.: Iconoscope Preamplifier Report LB-448 of the RCA License Laboratory. Information made available by special permission. See also: *RCA Rev.*, **4,** No. 1, 89, July, 1939.

Burnett, C. E.: The Monoscope, *RCA, Rev.*, **2,** No. 4, 414, April, 1938.

Farnsworth, P. T.: Television by Electron Image Scanning, *Jour. Franklin Inst.*, **218,** 411, October, 1934.

———: Image Amplifier Pick-up Tubes (delivered before the Rochester section, I.R.E., Nov. 14, 1938). Described briefly in *Electronics*, **11,** No. 12, 8–9, December, 1938.

Fiske, H. A.: A Television Pick-up Tube, *Proc. I.R.E.*, **27,** 144, February, 1939.

Iams, Janes, and Hickok: The Brightness of Outdoor Scenes and Its Relation to Television Transmission, *Proc. I.R.E.*, **25,** 1034, August, 1937.

Iams, Morton, and Zworykin: The Image Iconoscope (presented before the Annual I.R.E. Convention, June 17, 1938). Described briefly in *Electronics*, **11,** No. 12, July, 1938.

Iams, Morton, and Zworykin: The Image Iconoscope, *Proc. I.R.E.*, **27,** 541, September, 1939.

Iams and Rose: Television Pick-up Tubes Employing Cathode-ray Beam Scanning, *Proc. I.R.E.*, **25,** 1048, August, 1937.

——— and ———: A New Television Pickup Tube (presented before the New York Section, I.R.E. June 7, 1939). Described in "The Orthicon," *Electronics*, **12,** No. 7, 11, July, 1939.

——— and ———: Television Pickup Tubes Using Low-velocity Electron Beam Scanning, *Proc. I.R.E.*, **27,** 547, September, 1939.

Janes and Hickok: Recent Improvements in the Design and Characteristics of the Iconoscope, *Proc. I.R.E.*, **27,** 535, September, 1939.

Larson and Gardner: The Image-dissector, *Electronics*, **12,** No. 10, 24, October, 1939.

Lewis, H. M.: Standards in Television, *Electronics*, **10,** No. 7, 10, July, 1937.

McIlwain, Knox: Survey of Television Pickup Devices, *J. Applied Phys.*, **10,** 432, July, 1939.

Maloff, I. G.: Gamma and Range in Television, *RCA Rev.*, **3,** No. 4, 409, April, 1939.

Murray, A. F.: R.M.A. Television Standards, *R.M.A. Eng.*, **1,** No. 2, November, 1936.

———: R.M.A. Completes Television Standards, *Electronics*, **11,** No. 7, 28, July, 1938.

Rose and Iams: The Orthicon, Television Pickup Tube, *RCA Rev.*, **4,** No. 2, 186, October, 1939.

Zworykin, V. K.: The Iconoscope, a New Version of the Electric Eye, *Proc. I.R.E.*, **22,** 16, January, 1934.

———: Iconoscopes and Kinescopes in Television, *RCA Rev.*, **1,** No. 1, 60, July, 1936.

———, Morton, and Flory: Theory and Performance of the Iconoscope, *Proc. I.R.E.*, **25,** 1071, August, 1937.

VIDEO AMPLIFICATION

36. Requirements for Video Amplification. The transmission system must transmit all sine-wave components within the video range (*e.g.*, 30 cps to 4 Mc) without amplitude discrimination and without phase discrimination. The gain G of a pentode amplifier stage (plate resistance large compared with the load resistance) is

$$G = g_m Z_0 \tag{14}$$

where g_m is the grid-plate transconductance of the tube, and Z_0 is the output impedance of the coupling connection between the stage and the following transducer. Over the video range g_m is independent of frequency; hence the amplitude and phase responses of the amplifier are determined solely by Z_0.

In video amplifiers, Z_0 consists of R, L, and C components so proportioned as to display a constant magnitude of impedance and a phase angle proportional to frequency over the video range. The lower frequency limit over which these conditions may be met is determined by the series impedance of the coupling capacitor, whereas the h-f limit is determined by the shunt impedance of the capacitance existing in shunt across the coupling connection. The value of gain within these limits depends on g_m and on the value of the load resistor, since this is the principal component of Z_0 within the video range.

37. High-frequency Compensation. To minimize the effect of the shunt capacitance, it is usual in video amplifiers to insert a small inductance either in series with the load resistor (shunt peaking), in series with the coupling connection (series peaking), or a combination of the two (shunt-series peaking). The inductance is used to form a resonant circuit with the shunt capacitance at a frequency above the upper limit of the required v-f range, and the rising resonance characteristic is used, to counteract the falling off of the Z value at the upper frequency limit. The load resistor must similarly be chosen in terms of the total shunt capacitance, so that the gain in the mid-frequency range (where reactive effects are not prominent) will be the same as at the upper limit (where reactive effects are predominant).

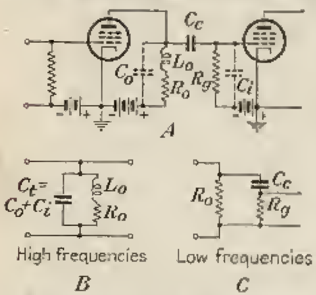


FIG. 25.—High-frequency compensation by the shunt-peaking method, with equivalent circuits for high and low frequencies.

In all cases of h-f compensation the basic factor is the total shunt capacitance C_t associated with the coupling connection

$$C_t = C_{pk} + C_{gk} + C_{gp}(G + 1) + C_{stray} \tag{15}$$

where C_{pk} = output tube capacitance
C_{gk} = input capacitance of the following tube
C_{gp} = grid-plate capacitance of the following tube
G = stage gain of the following stage
C_{stray} = total shunt capacitance due to wiring, tube sockets, terminals, etc.

In pentode amplifiers C_{gp} may ordinarily be neglected.

38. Shunt-peaking Compensation. The most widely used h-f compensation scheme (Fig. 25) is known as shunt peaking, because the resonating

(peaking) inductance L_0 is in shunt across the shunt capacitance C_t. The design values of L_0 and R_0 (the load resistor) are based on the shunt capacitance C_t, on the maximum required frequency in the video range f_{max} and on two design constants k_L and k_R which relate the impedance of L_0 and R_0, respectively, to the impedance of C_t at the maximum frequency f_{max}.

$$k_R = \frac{R_0}{1/(2\pi f_{max} C_t)} \tag{16}$$

$$k_L = \frac{2\pi f_{max} L_0}{1/(2\pi f_{max} C_t)} \tag{17}$$

The values of k_R range from 0.8 to 1.0; most designs are based on $k_R = 1$, *i.e.*, the load resistance is made equal to the impedance of C_t at the maximum v.f. The values of k_L range from 0.3 to 0.7, but most designs are based on $k_L = 0.5$, *i.e.*, impedance of the inductance L_0 is made one-half as great as the impedance

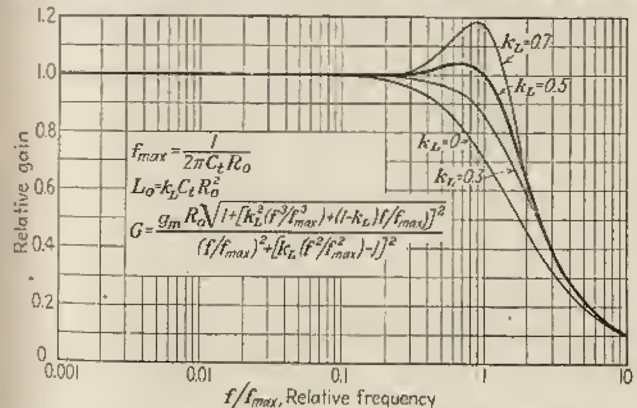


FIG. 26.—Gain of a shunt-compensated video amplifier.

of C_t, at f_{max}. This is equivalent to making the resonant frequency between L_0 and C_t equal to 1.41 times f_{max}.

On the assumption that $k_R = 1.0$, the expression for the gain of the shunt-compensated video amplifier is

$$G = \frac{g_m R_0 [1 - j\{k_L^2 (f/f_{max})^3 + (1 - k_L)(f/f_{max})\}]}{(f/f_{max})^2 + [k_L (f/f_{max})^2 - 1]^2} \tag{18}$$

where G is the gain at frequency f, and the other quantities have been defined. The absolute magnitude of this equation is plotted in Fig. 26, and its phase angle in Fig. 27, for several values of k_L.

The simplified design equations for shunt peaking ($k_R = 1$ and $k_L = 0.5$) are as follows:

$$R_0 = \frac{1}{2\pi f_{max} C_t} \tag{19}$$

$$L_0 = 0.5 C_t R_0^2 \tag{20}$$

Typical values of R_0 are 2,000 to 4,000 ohms and of L_0 are 50 to 100 μh.

39. Series-peaking Compensation. The compensation in Fig. 28 has an advantage over the shunt-peaking system in that the inductance L_c isolates the effects of the output and input capacitances C_0 and C_i; whereas in the shunt-peaking systems, C_0 and C_i are directly additive. Since C_0 is less than C_i, for a given h-f limit R_0 may be made correspondingly larger; hence the

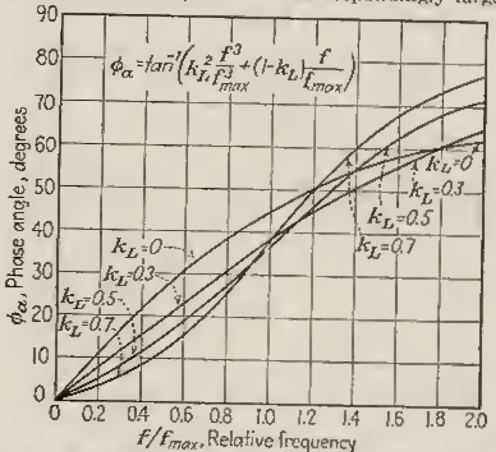


FIG. 27.—Phase angle introduced by a shunt-compensated video amplifier.

gain of the stage is increased. On the assumption that $C_i/C_0 = 2$ (usually assumed condition), the design equations for R_0 and L_c are as follows:

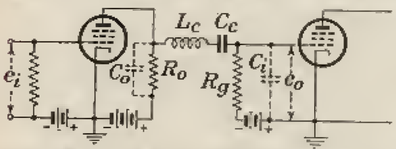


FIG. 28.—The series-peaking system of high-frequency compensation.

$$R_0 = \frac{1.5}{2\pi f_{max}(C_0 + C_i)} \tag{21}$$

$$L_c = 0.67 C_i R_0^2 \tag{22}$$

With these values the gain is uniform up to f_{max}, and its value is 50 per cent greater than the gain of the shunt-compensated stage with the same values of C_0, C_i, and C_i, provided $C_i/C_0 = 2$.

40. Shunt-series-peaking Compensation. The combination of shunt and series peaking (shown in Fig. 29) allows still higher gain by combining the

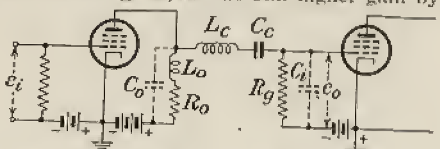


FIG. 29.—Combined shunt- and series-peaking system of high-frequency compensation.

virtues of both connections. Assuming $C_i/C_0 = 2$, the design equations are

$$R_0 = \frac{1.8}{2\pi f_{max} C_i} \tag{23}$$

$$L_0 = 0.12 C_i R_0^2 \tag{24}$$

$$L_c = 0.52 C_i R_0^2 \tag{25}$$

The stage displays up to f_{max} uniform gain, which is 80 per cent greater than that of the simple shunt-peaking stage. The relative merits and design factors of the three methods of h-f compensation are shown in Table II.

TABLE II.—HIGH-FREQUENCY COMPENSATION SYSTEMS

Type	R_0	L_0	L_c	Relative gain at f_{max}	Variation in time delay, seconds up to f_{max} cps
Uncompensated	$1/(2\pi f_{max} C_i)$	....	....	0.707	$0.035/f_{max}$
Shunt	$1/(2\pi f_{max} C_i)$	$0.5 C_i R_0^2$	....	1.0	$0.023/f_{max}$
Series ($C_i/C_0 = 2$)	$1.5/(2\pi f_{max} C_i)$	....	$0.67 C_i R_p^2$	1.5	$0.0113/f_{max}$
Shunt-series ($C_i/C_0 = 2$)	$1.8/(2\pi f_{max} C_i)$	$0.12 C_i R_0^2$	$0.52 C_i R_0^2$	1.8	$0.015/f_{max}$

41. Low-frequency Compensation. The amplitude response of conventional resistance-capacitance-coupled amplifier stages at low fre-

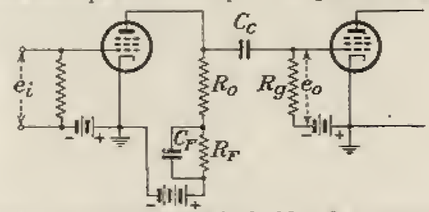


FIG. 30.—Resistance-capacitance method of low-frequency compensation.

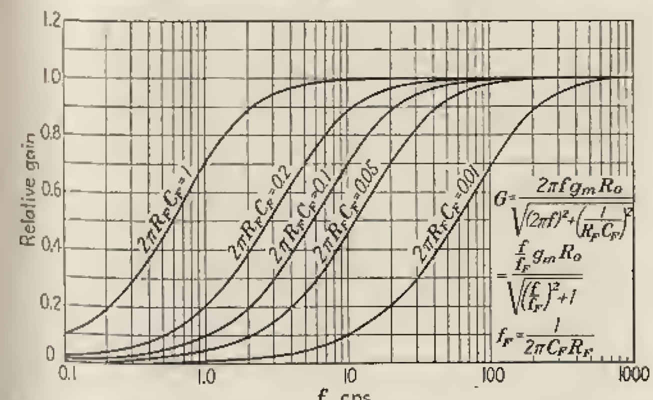


FIG. 31.—Amplitude response of low-frequency compensation system.

quencies is usually satisfactory, but the phase response at the low frequencies is troublesome.

The phase angle introduced by the coupling connection C_c and the grid resistor R_g of the following stage is sufficient to prevent proper reproduction of square waves of 30 or 60 cps fundamental frequency, unless very large values of C_c and R_g are employed. Large values of C_c introduce shunt capacitance to ground, and large values of R_g introduce grid-current difficulties in the following stage. Large values of C_cR_g may induce relaxation oscillations. Accord-

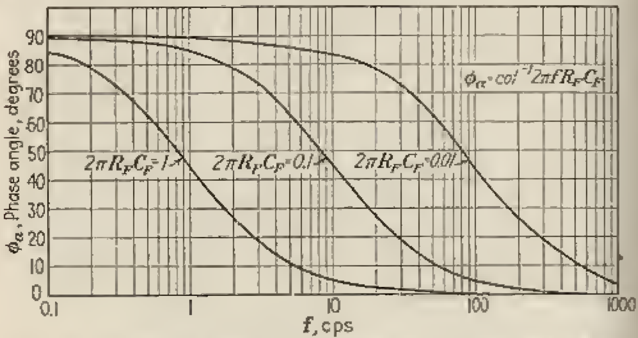


FIG. 32.—Phase response of low-frequency compensation system.

ingly it is usual to compensate the effect of the time constant C_cR_g by the introduction of a filter R_FC_F shown in Fig. 30.

The design equation is

$$\frac{C_F R_0 R_F}{R_0 + R_F} = C_c R_g \qquad (2$$

When this condition is met, the gain at low frequencies is

$$G = \frac{(f/f_F)g_m R_0}{(f/f_F - j)} \qquad (27$$

where G = the gain at frequency f
$f_F = 1/(2\pi C_F R_F)$
$j = \sqrt{-1}$.

The amplitude and phase of Eq. (27) are shown in Figs. 31 and 32. Values of R_FC_F from 0.15 to 0.5 should be used to keep the point of zero-phase shift below 30 cps, as indicated in Fig. 32.

42. Cathode-coupled Stage. For many purposes a video-amplifier stage displaying low output impedance is necessary (to match the impedance of coaxial cables and to permit the stage to feed many high impedance sources at once). The cathode-coupled stage (Fig. 33) is commonly used for this purpose. The gain of this stage is less than unity, and its output impedance can be designed readily for values as low as 50 ohms. The amplifier, being degenerative, has lower values of input capacitance, is freer from amplitude distortion, and is less affected by changes in supply voltages than is the conventional amplifier stage.

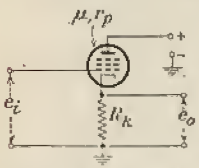


FIG. 33.—Fundamental cathode-coupled stage.

The gain of the cathode-coupled stage is

$$G = \frac{\mu R_k}{r_p + R_k(\mu + 1)} \qquad (2$$

where μ = amplification factor of the tube
r_p = its internal plate resistance
R_k = value of the cathode resistor.

The effective output impedance Z_0 is

$$Z_0 = \frac{R_k r_p/(\mu + 1)}{R_k + r_p/(\mu + 1)} \qquad (29)$$

An important practical advantage of the cathode-coupled stage is that it may be coupled to the following transducer without the intervention of a coupling capacitor, so that the d-c as well as a-c components of the video signal are transmitted. No pains need be taken to preserve the h-f response, since the low value of impedance makes the shunting effect of the output capacitance negligibly small.

43. Transient Response of Video Amplifiers. The response of a video amplifier to the Heaviside unit pulse of voltage is a general criterion of

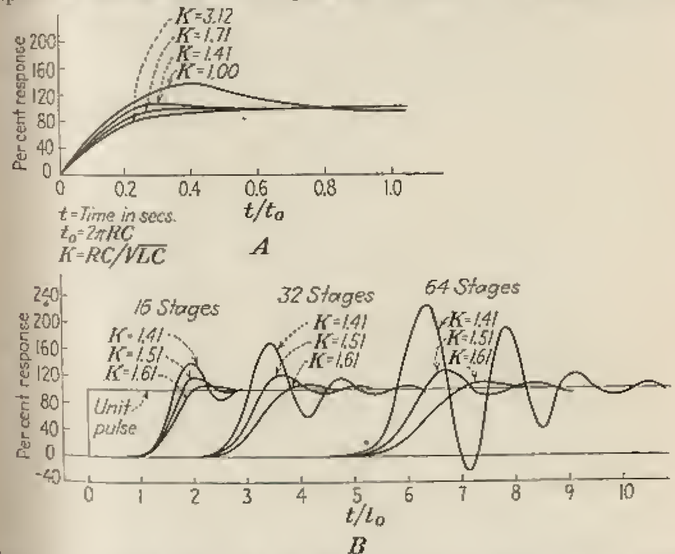


FIG. 34.—Transient response of single and multistage compensated video amplifiers. (*After Bedford and Fredendall.*)

video-amplifier response. The response to a single unit pulse is difficult to measure experimentally, but a square wave may be used as the exciting voltage, provided that the period of the wave is long compared with the duration of the transient response. Responses calculated on this assumption are shown in Fig. 34, for a single stage and for several multistage amplifiers. Simple shunt peaking is assumed, for various values of the parameter $K = RC/\sqrt{LC}$, *i.e.*, the ratio of the load resistor R_0 to the impedance of the shunt capacitance C_t at the frequency at which L_0 and C_t are resonant. The case for $K = 1.41$ is equivalent to the cases of $k_R = 1$ and $k_L = 0.5$ (Art. 38).

44. Noise Limitations to Video Amplification. One of the principal limitations to proper video amplification is inadequate signal/noise

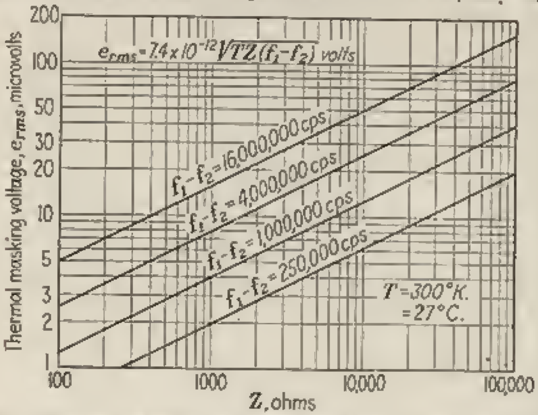


FIG. 35.—Thermal agitation voltage generated in wide-band circuits.

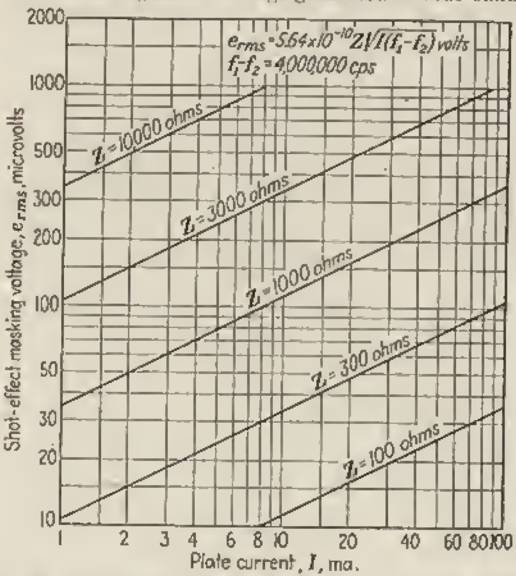


FIG. 36.—Shot-effect noise generated in wide-band circuits.

ratio. The two sources of circuit noise, thermal agitation and shot effect, are evaluated in Figs. 35 and 36 for a transmission system responsive to the video range. Values of 50 to 100 μv are common. For a signal/noise ratio of 10:1, commonly assumed as the minimum acceptable for entertainment purposes, the desired signal must accordingly have an r-m-s amplitude of from 0.5 to 1.0 mv.

References

Video Amplification:

BARCO, A. A.: Measurement of Phase Shift in Television Amplifiers, *RCA Rev.*, **3**, No. 4, 441, April, 1939.

BARBER, A. W.: Video Amplifier Design, *Communications*, **18**, No. 6, 13, June, 1938.

BEDFORD and FREDENDALL: Transient Response of Multistage Video Amplifiers, *Proc. I.R.E.*, **27**, 277, April, 1939.

BUILDER, G.: The Amplification of Transients, *Wireless Eng., Exp. Wireless*, 246, May, 1935.

CARNAHAN, C. W.: The Steady-state Response of a Network to a Periodic Driving Force of Arbitrary Shape, and Its Applications to Television Circuits, *Proc. I.R.E.*, **23**, 1393, November, 1935.

EVEREST, E. A.: Wideband Television Amplifiers, *Electronics*, **11**, No. 1, 16, January, 1938; **11**, No. 5, 24, May, 1938.

FREEMAN and SCHANTZ: Video Amplifier Design, *Electronics*, **10**, No. 8, 22, August, 1937.

HEROLD, E. W.: High-frequency Correction in Resistance-coupled Amplifiers, *Communications*, **18**, No. 8, 11, August, 1938.

KAUZMANN, A. P.: New Television Amplifier Receiving Tubes, *RCA Rev.*, **3**, 3, 271, January, 1939.

KEALL, O. E.: Correction Circuits for Amplifiers, *Marconi Rev.*, **54**, 15, May, 1935.

LANE, H. M.: Resistance-capacitance Amplifier in Television, *Proc. I.R.E.*, **20**, 722, April, 1932.

MCLACHLAN, H. W.: Reproduction of Transients by Television Amplifiers, *Wireless Eng.*, **13**, 519, October, 1936.

NAGY, P.: The Design of Vision-frequency Amplifiers, *Television*, **10**, No. 160, 220, 279, March, April, May, 1937.

OAKEY, C. W.: Distortionless Amplification of Electrical Transients, *Wireless Eng. Exp. Wireless*, 245, May, 1931.

POLLACK, DALE: Choice of Tubes for Wide-band Amplifiers, *Electronics*, **12**, 3, 38, April, 1939.

PREISMAN, A.: Some Notes on Video Amplifier Design, *RCA Rev.*, **2**, No. 4, 421, April, 1938.

PUCKLE, O. S.: Transient Aspect of Wideband Amplifiers, *Wireless Eng.*, **12**, 251, May, 1935.

ROBINSON, G. D.: Theoretical Notes of Certain Features of Television Receiving Circuits, *Proc. I.R.E.*, **21**, 833, June, 1933.

SEELEY and KIMBALL: Analysis and Design of Video Amplifiers, *RCA Rev.*, **2**, No. 2, 171, October, 1937; **3**, No. 3, 290, January, 1939.

SHIFFENBAUER, R. G.: Phase Distortion in Television, *Wireless Eng.*, **13**, 21, January, 1936.

SWIFT, G.: Amplifier Testing by Means of Square Waves, *Communications*, **14**, No. 2, 22, February, 1939.

WHEELER, H. A.: The Interpretation of Amplitude and Phase Distortion in Terms of Paired Echoes, *Proc. I.R.E.*, **27**, 359, June, 1939.

———: Wideband Amplifiers for Television, *Proc. I.R.E.*, **27**, 429, July, 1939.

WILSON, J. C.: "Television Engineering," Chap. VI, Sir Isaac Pitman & Sons, Ltd., London, 1937.

MODULATION, R-F AND I-F AMPLIFICATION, DETECTION

45. Video Modulation. Video modulation is based on the same considerations as audio modulation, with certain specialized requirements. One of the limitations is the small amount of video-signal voltage which may be generated in currently available tubes and circuits. The high capacitance to ground of large water-cooled tubes requires the use of very low values of load resistance to maintain response over the v-f range. The voltage which can be developed across the load resistance depends on the emission current. At present it is difficult to generate more than the 1,500 or 2,000 volts, peak-to-peak over the video range from 30 cps to 4.5 Mc. When high-level modulation is used, therefo is usually considered expedient to use grid-circuit modulation, ra than plate-circuit modulation, since the voltage requirements for modulation are less by the amplification factor of the modulated st Low-level modulation is not similarly restricted but has not pro popular because of the very low efficiency of the modulated r-f ampli which follow the modulator and also because of the difficulty of maint ing the characteristics of vestigial side-band transmission (Art. 46) un the r-f amplifiers are highly linear.

The second unusual requirement in video modulation is the neces for maintaining two levels in the modulation envelope at const amplitudes. These levels are (1) the tips of the sync pulses, *i.e.*, maximum amplitude of the envelope; and (2) the blanking level pedestal. Since these levels must remain constant regardless of changes in the wave form of the camera-signal component, it is necess to couple the modulating amplifier conductively to the modulated am fier. This makes necessary a separate power supply for each stage.

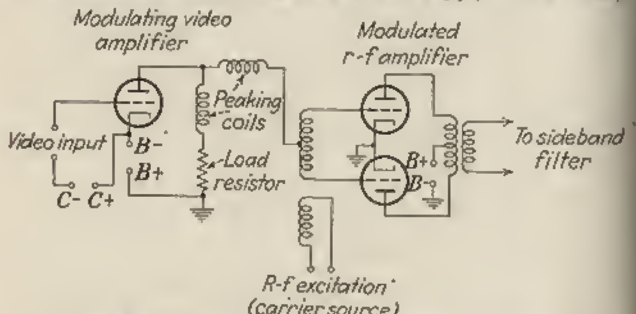


FIG. 37.—Fundamental circuit of video modulator.

typical arrangement is shown in Fig. 37. Here the modulating vi amplified is coupled conductively to the grids of the r-f amplifier. T B supply for the modulating amplifier is in series with the cathode.

At the grid of the modulating amplifier, it is necessary that the blank level and sync-pulse tip level be constant. The latter levels are cau to assume fixed values by passing the video wave form through a di rectifier whose cathode is connected to the modulating video-ampl grid. The load circuit values are chosen so that the rectified d-c pot tial across the diode assumes a level at the tips of the sync pulses just below the tips (the difference being required to supply the di current). The voltage across the diode forms a part of the fixed b of the modulating amplifier. The composite wave form, extend more positively than the tips of the sync pulses, causes the modulat amplifier output voltage to extend more negatively than the sync pul This output voltage, applied to control the amplitude of the modul r-f amplifier, causes the sync-pulse tips to assume the peak position in t envelope, while the blanking level and camera-signal components ext to lower levels in the envelope. The sync pulses and blanking le maintain constant amplitudes, whereas the average on the camera-sig

component changes with the background illumination of the scene (see Fig. 10, Art. 19).

46. Vestigial Side-band Transmission. The side bands of the modulated r-f signal, assuming a maximum video-modulating frequency of 4.5 Mc, extend over a total region of 9 Mc. To conserve space in the ether and at the same time to secure greater efficiency from r-f and i-f amplifiers, the N.T.S.C. recommended standards specify vestigial side-band transmission (sesqui-side-band, selective side-band, or "single" side-band transmission). In this system a part of the lower frequency side band is completely attenuated. By this means the upper side band can be transmitted completely with 4.0 to 4.5 Mc width, within the 6-Mc channel assigned by the FCC. A portion of the lower side band, within 1.25 Mc of the carrier frequency, is also transmitted.

The channel composition for vestigial side-band transmission is shown in Fig. 38, at the top. The lower figure shows the corresponding characteristic

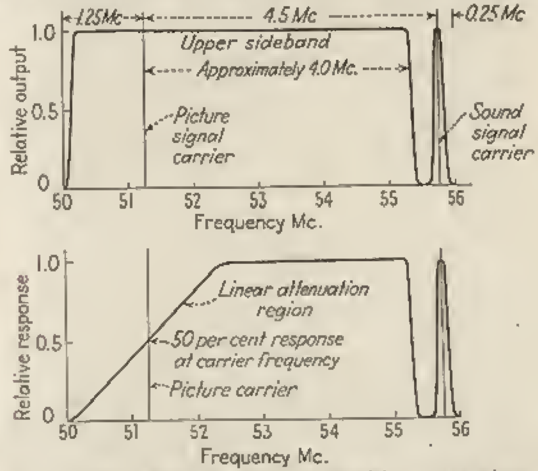


FIG. 38.—Top, output characteristic of television transmitter. Bottom, corresponding input response characteristic of receiver.

of the receiver. In the receiver characteristic the picture signal-carrier voltage is attenuated to 50 per cent of its original level, and the curve of attenuation is linear over a range of 2.5 Mc. This arrangement develops 50 per cent modulation in those portions of the carrier which receive double-side-band treatment (within 1.25 Mc of the carrier). The modulation of the components further removed from the carrier in the upper side band are inherently 50 per cent modulated, so all portions of the signal, when presented to the detector, produce an equal amplitude in the detector output.

To form a transmitted signal of the character shown at the top in Fig. 38, a filter having sharp cutoff characteristics is required. In Fig. 39 the desired upper side band is passed through a capacitor to the antenna, whereas the undesired lower side band is passed through an inductance to an absorbing resistor. Filter structures for this purpose, when employed for high-level modulation, are customarily formed from sections of coaxial transmission lines. In addition to the filter shown, a sharply tuned "notching filter" is usually used to provide additional attenuation at the sound-carrier frequency of the adjacent channel.

47. Allocation of Television Channels. Figure 40 shows the allocation of 18 six-megacycle channels for television in the region between 44 and 300 Mc allocated by the FCC. The five channels lowest in frequency are currently considered most useful for public service, because of the greater transmitter and receiver efficiencies in the lower frequencies. Frequencies from 150 to 500 Mc have been employed for relaying television signals between stations.

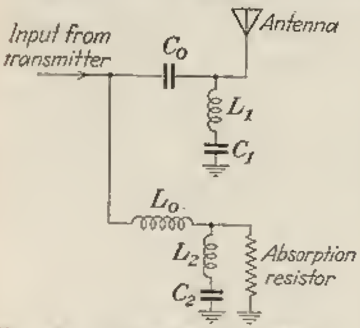


FIG. 39.—Elements of a vestigial side-band filter.

48. Wide-band R-f Amplification. The gain of a wide-band r-f pentode amplifier is equal to the product of the tube transconductance g_m by the load impedance Z_0. For maximum gain per stage both g_m and the absolute value of Z_0 must be as large as possible. To maintain the proper band-pass characteristics, Z_0 must be designed to have as nearly constant amplitude and as nearly linear phase as possible over the desired operating range and to have nearly zero impedance outside these limits. The design of the optimum Z_0 to meet these conditions is best attacked from the standpoint of band-pass filter theory, as indicated in Sec. 6 of this handbook.

Some general considerations are revealed simply in the analysis of the single loaded tuned circuit, shown in Fig. 41. At the resonant frequency f_r the impedance of the circuit is equal to the value of the shunting resistor. At

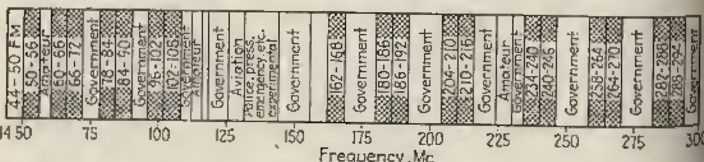


FIG. 40.—Allocation of frequencies from 30 to 300 Mc, according to regulations of the Federal Communications Commission.

frequencies removed from resonance, the impedance is less by the amount shown, and the degree of attenuation depends on the ratio of the resistance to the impedance of the inductance at the resonant frequency. The general relationships between the impedance and resistance are shown in Fig. 41. The corresponding phase relationships are shown in Fig. 42.

The design of the single loaded tuned circuit is based on the necessity of (1) obtaining resonance at or near the carrier frequency and (2) loading the circuit to present nearly uniform response over the side-band regions. In particular, if it is desired that the circuit display an impedance at the edge of the side-band regions equal to 0.707 times the impedance at resonance, the value of the resistance required is

$$R = \frac{f_r}{\Delta f}\sqrt{\frac{L}{C}} \tag{30}$$

where f_r = resonant frequency
Δf = total frequency width of the region within which the response is desired within unity and 0.707
R = maximum value of impedance which the tuned circuit impedance Z_0 may have.

For maximum gain the L/C ratio should be as high as possible. It is usual to employ as C only the stray and distributed capacitance present in the circuit and to bring this capacitance to resonance by employing a variable value of L. The loading is determined from Eq. (30).

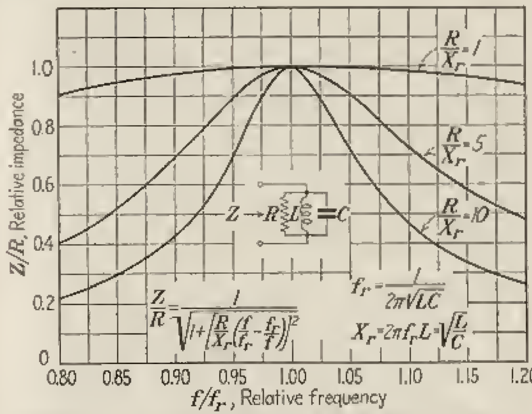


FIG. 41.—Impedance characteristics (absolute magnitude) of the single loaded tuned circuit (inset).

Since the single-tuned circuit cannot display a uniform impedance over an extended band width, it is usually desirable to employ coupled circuits (coupled capacitively, self-inductively, or mutually inductively) to obtain a flat-top response curve. In coupled circuits the impedance is a complicated function, but in general its value is independent of the carrier frequency, inversely proportional to the band width, and directly proportional to the L/C ratio of the tuned circuits.

The phase response of a wide-band r-f amplifier should be as linear as possible over the band-pass region. Such linearity is associated with symmetry in the amplitude characteristic; hence it is unwise to allow an amplifier to be unsymmetrical in one stage and to compensate with an antisymmetrical characteristic in the next stage, since poor phase response will usually result.

49. Picture I-f Amplification. The design of i-f amplifier circuits for television i-f signals is similar to the design of wide-band r-f circuits, except that lower carrier frequencies are used. Also, since most of the gain in a television receiver resides in the i-f amplifier stages, the problem of selectivity against interference from adjacent channels must be contended with.

The band-pass characteristic of an ideal picture i-f amplifier is shown in Fig. 43. The values of carrier frequencies are those recommended by the

R.M.A. Committee on Television, *i.e.*, 8.25 Mc for the sound carrier and 12.75 Mc for the picture. The adjacent carrier frequencies are 14.25 Mc for the adjacent sound channel, and 6.75 Mc for the picture carrier of the oppositely adjacent channel. Ordinarily the picture i-f circuit need be designed to display selectivity against the two sound channels only, at 8.25 Mc against the associated sound channel and at 14.25 Mc against the adjacent-channel sound carrier. The usual values of attenuation are 40 db at 8.25 Mc and 60 db at 14.25 Mc.

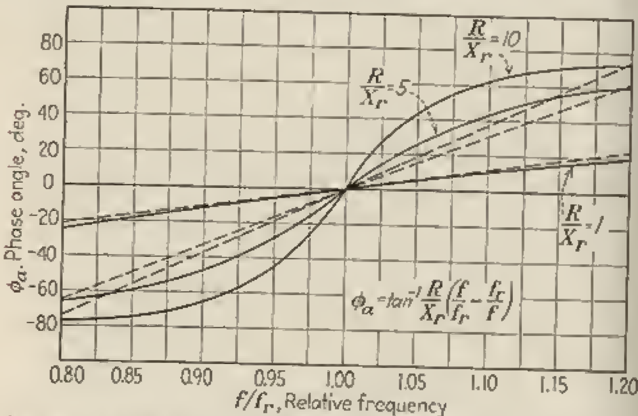


FIG. 42.—Phase angle of impedance of single loaded tuned circuit.

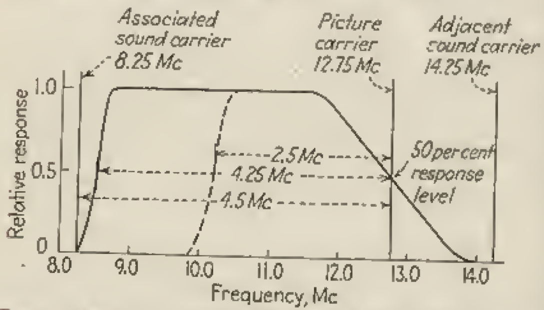


FIG. 43.—Response curve of typical television intermediate frequency amplifier, according to R.M.A. recommended practice.

To produce the recommended standard i.f. of 8.25 and 12.75 Mc, the frequency of the local oscillator must be 8 Mc higher in frequency than the upper frequency limit of the channel under consideration, *i.e.*, 64 Mc for the 50- to 56-Mc channel.

The gain per stage in picture i-f amplifiers depends directly on the band width passed. Stage gains of 10 are possible when accepting the full band width of 4 Mc shown in Fig. 43. For a band width of 2.5 Mc, typical in receivers using a 5-in. cathode-ray tube, the gain per stage may rise to 15 per

stage. A total i-f gain of 10,000 is usually considered sufficient. The effective tube and circuit noise at the input to the first i-f stage is usually 100 µv

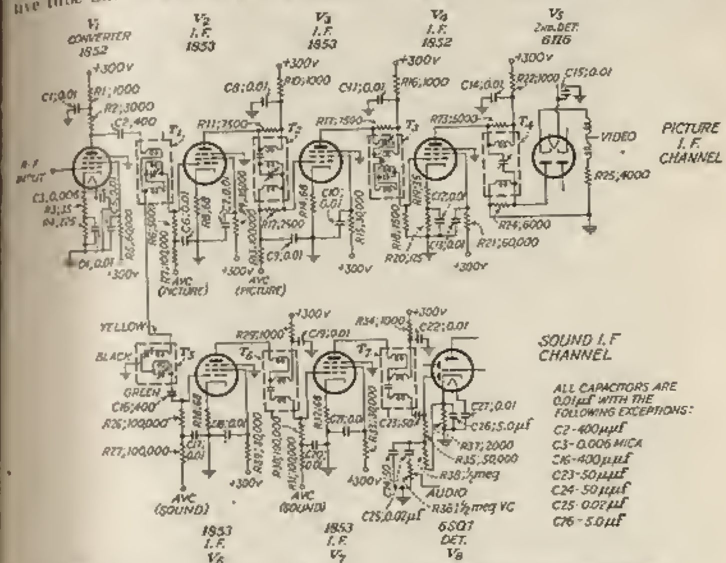


FIG. 44.—Connection diagram of typical picture i-f amplifier.

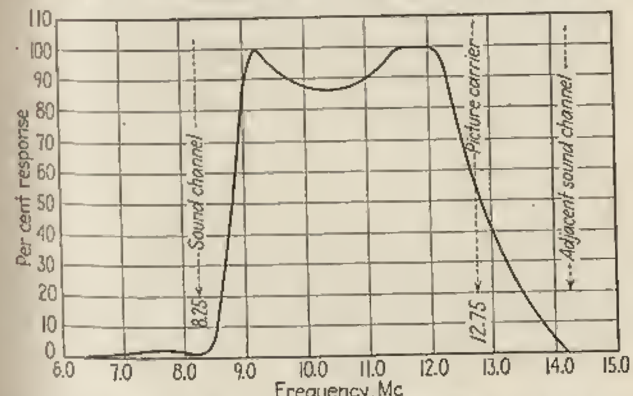


FIG. 45.—Response curve of picture i-f amplifier shown in Fig. 44.

or more. With a gain of 10,000 the noise voltage applied to the detector would be 1 volt, which is sufficient to make it plainly visible in the cathode tube. Sensitivity greater than this is clearly not necessary. Total i-f gain as low as 2,000 may be used in low-priced receivers, intended for use w input r-f signals of 1,000 μv or more.

50. Video Detection. The diode detector is used almost univers for video demodulation in current receivers. The important conside tions are (1) the amplitude and phase responses of the load circuit the detector over the video range, (2) the discrimination of this circ against components of carrier frequency, (3) the loading exerted by th circuit on the i-f coupling circuit which feeds the detector, and (4) t polarity of the detected voltage output.

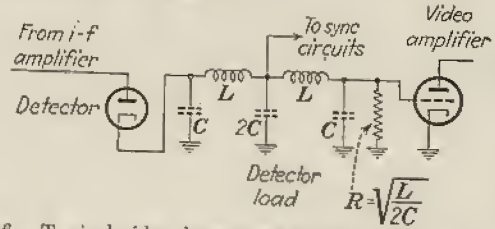


FIG. 46.—Typical video detector circuit with filter load circuit.

In designing the detector load circuit, the important factors are the capac tance to ground of the detector output and the input capacitance of the follow ing video amplifier. The circuit usually used is very similar to the serie peaking circuit (Art. 39), and the expressions for R_0 and L_e [Eqs. (21) and (22 can be used, under the assumption that $C_i/C_0 = 2$. The values of R_0, determined, usually range from 2,000 to 5,000 ohms.

The simple series-peaking circuit possesses sufficient discrimination again carrier-frequency components when the detection occurs at radio frequenci

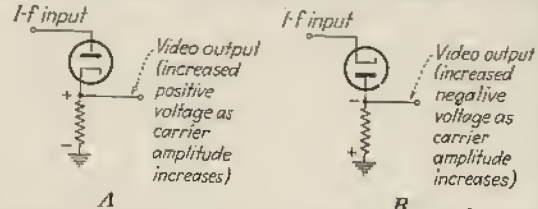


FIG. 47.—Detector polarities: A, cathode-above-ground connection; B anode-above-ground connection.

(above 40 Mc). But when i-f detection is considered (carrier frequenci from 8.5 to 13.0 Mc), it is preferable to design the detector load circuit in the form of a low-pass filter having a sharp cutoff above the v-f limit (5 Mc). typical constant-k filter section of this type is shown in Fig. 46.

The loading of the detector load circuit on the preceding i-f circuit calculated from

$$R_{\text{eff}} = \frac{R}{2\nu} \quad (31$$

where R_{eff} = effective load resistance on the i-f amplifier
R = actual value of the detector load resistor
ν = detection efficiency (very close to unity in most practical case

The polarity of the detected voltage output is important because it determines the number of video-amplifier stages required between the detector and the picture tube control grid to produce a picture having positive tone values. The two possible detector polarities are shown in Fig. 47. The cathode-above-ground connection produces an increased voltage output as the initial light in the studio decreases (assuming negative modulation, see Art. 20). Consequently one phase reversal is necessary between the detector and the picture tube. Any odd number of video-amplifier stages suffices (usually one stage is used). In the anode-above-ground connection the reverse polarity exists and an even number (usually two) of stages is required between the detector and picture tube. The same polarity considerations govern the number of amplifier stages required between the detector and the synchronizing input terminals of the scanning generators. With scanning generators synchronized by positive pulses (usual type) the cathode-above-ground connection shown at A requires an even number of intervening stages, whereas the anode-above-ground connection B requires an odd number of stages.

It is usual to operate video detectors with a maximum peak-to-peak input i-f voltage of 10 volts. Assuming full modulation, the peak-to-peak output voltage (with detector internal resistance and load resistance values equal) will be 5 volts, three-quarters of which constitutes the camera signal. A single video stage having a gain of 12 is consequently capable of delivering $5.0 \times 0.75 \times 12 = 45$ volts, peak to peak, in the picture tube grid. This value is sufficient to operate the usual picture tube over its entire control range.

References

Modulation, R-f and I-f Amplification, Detection:

BENHAM, W. E.: Aerial Coupling System for Television, *Wireless Eng.*, **15**, 555, October, 1938.

———: Asymetric Sideband Phase Distortion, *Wirelss Eng.*, **15**, 616, November, 1938.

CARTER, P. S.: Simple Television Antennas, *RCA Rev.*, **4**, No. 2, 168, October, 1939.

COCKING, W. T.: Television I-f Amplifiers, *Wireless Eng.*, **15**, 358, July, 1938.

CORK and PAWSEY: Aerial Feeders for Television, *Television*, **12**, 282, May, 1939.

ENGSTROM and BURRILL: Frequency Assignments for Television, *RCA Rev.*, **1**, No. 3, 88, January, 1937.

HOLLYWOOD, J. M.: Single Side-band Filter Theory with Television Applications, *Proc. I.R.E.*, **27**, 457, July, 1939.

LINDENBLAD, N.: Television Transmitting Antenna for Empire State Building, *RCA Rev.*, **3**, No. 4, 387, April, 1939.

LYMAN, H. T.: Television Radio Frequency Input Circuits, *R.M.A. Eng.*, **3**, No. 1, 3, November, 1938.

MOUNTJOY, GARRARD: Television Signal-frequency Circuit Considerations, *RCA Rev.*, **4**, No. 2, 204, October, 1939.

PARKER, W. N.: A Unique Method of Modulation for High-fidelity Television Transmitters, *Proc. I.R.E.*, **26**, 946, August, 1938.

POCH and EPSTEIN: Partial Suppression of One Sideband in Television Reception, *RCA Rev.*, **1**, No. 3, 19, July, 1937.

——— and ———: Partial Suppression of One Sideband in Television Reception, *RCA Rev.*, **1**, No. 3, 19, January, 1937.

RODER, HANS: Analysis of Load-impedance Modulation, *Proc. I.R.E.*, **27**, 386, June, 1939.

SEELEY: Effect of Receiving Antenna on Television Reception Fidelity, *RCA Rev.*, **2**, No. 4, 433, April, 1938.

STRUTT, M. J. O.: High Frequency Mixing and Detector Stages in Television Receivers, *Wireless Eng.*, **16**, 174, April, 1939.

SEPARATION OF THE SYNCHRONIZING SIGNALS

51. Amplitude Separation. The separation of the composite synchronizing signal from the camera signal is performed after the composite video signal has been developed by the second detector. The composite video signal (Fig. 48) is applied to a "clipper" tube, which is a tube that cuts off all current beyond a certain negative amplitude limit. A triode clipper tube and characteristic are shown in Fig. 48. In Fig. 49 a di clipper arrangement is shown in conjunction with the second detector. is necessary, of course, that the clipping level be maintained continuo

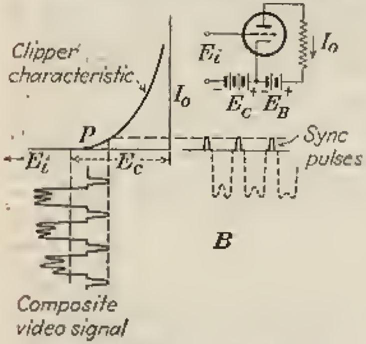


FIG. 48.—Clipper circuit and characteristic, used to separate composite sync signals from camera signals.

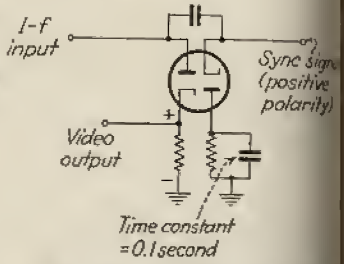


FIG. 49.—Combined second tector and sync amplitude separ (clipper).

at the blanking level to ensure that the camera signal does not aff synchronization, on the one hand, and to ensure that the maxim amplitude of sync pulses is developed, on the other.

52. Wave-form Separation. After the composite synchronizing sig has been separated from the video signal, it necessary to develop the horizontal sync pu independently of the vertical sync pulses. T latter separation is carried out by a meth known as wave-form separation, since the t sets of pulses cannot be distinguished by am tude means. Essentially wave-form separat depends on circuits which respond to the re tive frequency content of the two sets of pul The horizontal sync pulses that are of sh duration occur 13,230 times per second a have a predominance of h-f components, wh as the vertical pulses that are of long durat and occur 60 times per second have a pred inance of l-f components. The ratio of t frequencies of the two sets of pulses 13,230 /60 220½ is the index of the degree of frequ difference on which the separator circuits m operate.

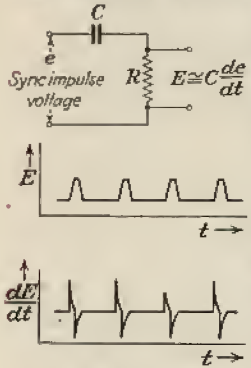


FIG. 50.—Differentiator circuit (top) and action on sync pulses (bottom).

53. Differentiator Circuit for Horizontal Sy Pulses. The differentiator circuit shown Fig. 50 is used to develop the h-f component the composite synchronizing signal, *i.e.*, the horizontal sync pulses. series capacitance passes the high frequencies associated with the lead edge of the sync pulse, while retarding all lower frequency componen

The RC product (time constant) of the combination is made short compared with the frame-repetition interval (1/60 sec.) and long compared with the line scanning interval (1/13,230 sec.). The leading edge of the differentiated wave forms is applied, in the proper polarity, to the synchronizing terminal of the horizontal scanning generator.

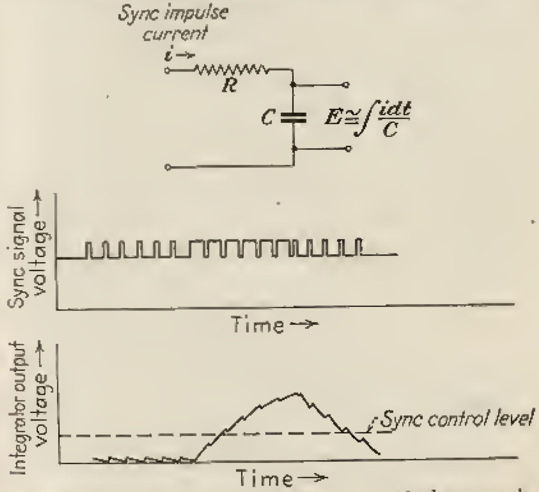


FIG. 51.—Integrator circuit (top) and action on vertical sync pulse (bottom).

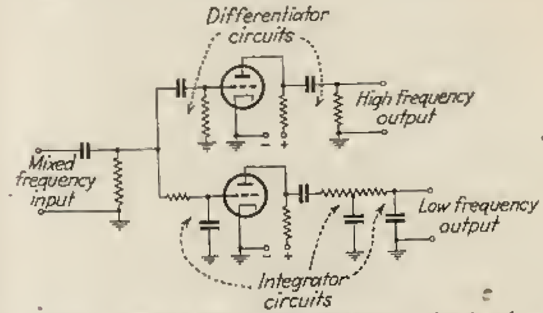


FIG. 52.—Combined differentiator and integrator circuits for wave-form separation.

54. Integrator Circuit for Vertical Sync Pulses. The integrator circuit shown in Fig. 51 develops a sync pulse from the serrated vertical pulse and equalizing pulses. The wave forms of input and output are shown. It will be noted that the initial portion of the integrated output pulse is not so sharply rising as that of the differentiated horizontal pulse, and consequently the intersection with the synchronizing control level is not so precisely marked. This fact makes it necessary to have the wave shape of each successive integrated pulse precisely the same. The function of the equalizing in this respect has been pointed out in Art. 23. In addition it is necessary that all traces of the horizontal sync pulses be completely removed from the integrating circuit.

Several differentiating and integrating circuits may be used in cascade to improve the degree of separation. The cascaded circuits may be connected directly together (usually done with integrator circuits), or they may occur in the grid and plate circuits of a sync separator amplifier tube. A typical synchronizing amplifier circuit is shown in Fig. 52.

PICTURE TUBES AND ASSOCIATED CIRCUITS

55. Picture Tubes. The conventional cathode-ray picture tube is a funnel-shaped evacuated structure containing an electron gun which forms an electron beam, and a fluorescent screen on which the beam

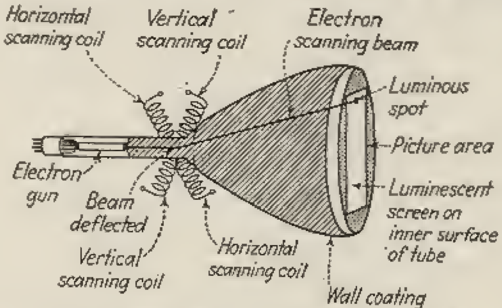


Fig. 53.—Structure of a typical picture tube.

impinges. The beam is deflected by the application of transverse electric or magnetic fields which cause the end of the beam to trace out the interlaced scanning pattern over the fluorescent screen. The current in the beam is capable of variation from zero (cutoff) to a maximum of several hundred microamperes, under the control of the signal potential applied between the cathode and the control electrode of the electron gun.

The beam is deflected synchronously with the scanning agent in the camera tube, and the beam current is controlled by the camera signal. The variations in the beam current produce corresponding variations in the brightness of the fluorescent spot, and the picture is thereby reproduced.

The operating characteristics of picture tubes depend on the design of the electron gun and on the physical and chemical properties of the fluorescent screen. The electron gun requires a power supply to form the electron beam. Finally the deflection fields must be provided by scanning generators, and these generators must operate under the control of the synchronizing signals of the video signal.

Picture tubes are classified according to (1) the type of focusing employed (electrostatic or magnetostatic) in the electron gun, (2) the type of deflection (electric or magnetic), (3) the type of phosphor (sul-

phide and non-sulphide), and (4) the color of the light produced (usually white).

56. Characteristics of Electron Guns. Electrostatically focused electron guns are characterized (1) by the ratio of the voltages applied to the second anode and the first anode. In present designs this ratio varies from 4 to 6. In addition the guns are characterized (2) by the control electrode characteristic which specifies the relation between control electrode voltage and beam current (second-anode current) for different values of second-anode voltage. A typical control characteristic of an electrostatically focused gun employed in the 12AP4 tube is shown in Fig. 54. Curves of this shape are typical of all types of electron guns, whether electrostatically or magnetostatically focused.

An important characteristic of electron gun is the degree of fineness of focus, *i.e.*, the size of the fluorescent spot formed on the screen. Guns of good design are capable of forming a fluorescent spot about 0.005 in. in diameter, but production tubes usually have spots from 0.01 to 0.015 in. in diameter. The latter spot size permits a picture resolution of 350 lines when the picture height is 6 in. or more (picture width 8 in. or more). For smaller tubes the spot size sets the upper limit of picture resolution at a figure lower than 350 lines. In 5-in. tubes, for example, resolution of 200 to 250 lines is typical performance of current tubes.

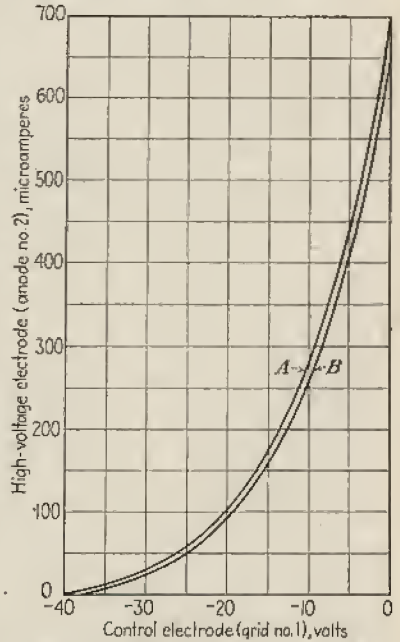


Fig. 54.—Electron-gun control characteristic. *A*, 7,000-volt; *B*, 6,000-volt second anode voltage.

57. Characteristics of Phosphors. The important operating characteristic of the phosphors (fluorescent materials) employed in picture tubes is the relationship between the light produced, the beam current (second-anode current), and the second-anode potential. Figure 55 shows a typical family of such curves, taken for the "P4" white-light phosphor employed in the 5AP4, 5BP4, 7AP4, 9AP4, and 12AP4 tubes.

58. Transfer Characteristic of Picture Tube. The transfer characteristic of a transducer in a television system is the relationship between the significant variational input quantity and the significant variational output quantity. In picture tubes the significant input is the control-electrode voltage, and the significant output is the corresponding light produced on the screen. The transfer characteristic of the 12AP4 tube is shown in Fig. 56. Note that the relationship is not linear but has the "antisaturation" shape, corresponding to a gamma greater than unity. This characteristic tends to enhance the apparent contrast of the picture (see Art. 72).

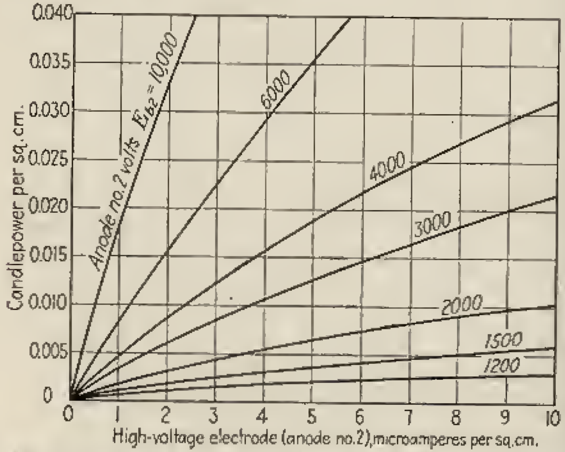


FIG. 55.—Phosphor light-output characteristic (type P4 phosphor).

59. Contrast in Picture Tubes. The ratio of the brightness of the brightest portion of an image to the brightness of the darkest portion is called the *brightness-contrast ratio.* Owing to the effects of light spreading (halation) within the glass envelope of the tube the maximum contrast ratio of current picture tubes is about 50:1. Between closely adjacent portions of the image, halation reduces the maximum obtainable contrast to about 10:1.

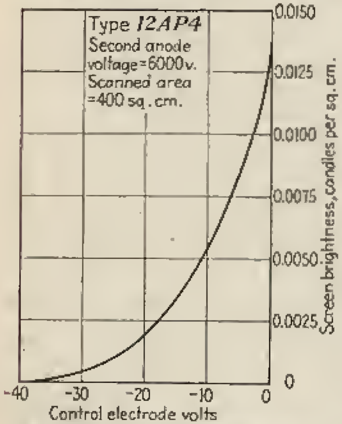


FIG. 56.—Typical transfer characteristic of a picture tube, derived from Figs. 54 and 55.

60. Dynamic Action of Picture-tube Control Circuit. The dynamic action of the picture-tube circuit is represented by applying the video signal wave form to the transfer characteristic (Fig. 57). The video wave form is applied so that the blanking level corresponds to the zero light (cutoff) point on the transfer characteristic as shown. This bias level must remain fixed at all times. Then the camera signal extending to the right of the blanking level produces light on the screen in accordance with the camera signal, whereas the synchronizing signals to the left of the blanking level are in the infra-

black region (beyond cutoff) and do not produce light. The total excursion of the camera signal should be limited so that the control-electrode voltage never becomes positive; usually the control electrode does not go beyond the −5 or −10 volts mark. The average of the picture signal component, taken over the frame interval, establishes the background brightness of the scene, provided the blanking level remains fixed at the light cutoff point.

61. Direct-current Restoration Circuits. Two typical circuits used to maintain the blanking level constant at the picture-tube control electrode

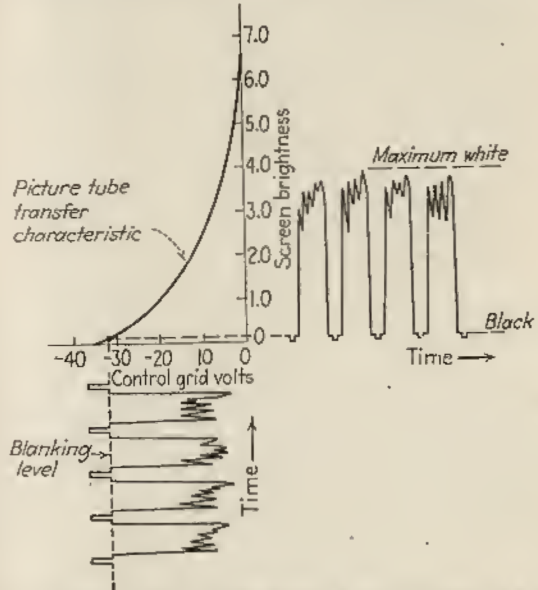


FIG. 57.—Dynamic action of video signal in picture-tube control circuit.

are shown in Fig. 58. The proportions of C_c and R_g are chosen to have a time constant long compared with the line-scanning interval but short compared to the duration of the changes in background light. In the upper diagram the grid and cathode of the video-amplifier tube act as a diode, whereas in the lower case a separate diode is employed. The diode, acting in conjunction with C_cR_g, develops a bias equal to the peak of the video signal. This peak value remains fixed (assuming no change in signal strength), consequently the remainder of the video signal (the camera-signal components) act in series with a fixed bias. This fixed bias is coupled conductively (either through the amplifier tube, at the top, or directly, at the bottom) and forms a part of the control-electrode bias. By this means the blanking level remains fixed, and, if the total control-electrode bias is fixed so that the blanking level coincides with the light cutoff point, the background brightness of the scene depen[ds] only on the average of the camera-signal component, as is required.

62. Picture-tube Power Supplies. The picture-tube power supp[ly] consists of (1) a source of high voltage for the first and second anod[es] which act to draw the electrons from the gun and (in the case of electr[o]statically focused tubes) bring the beam to focus; (2) a source of heat[er]

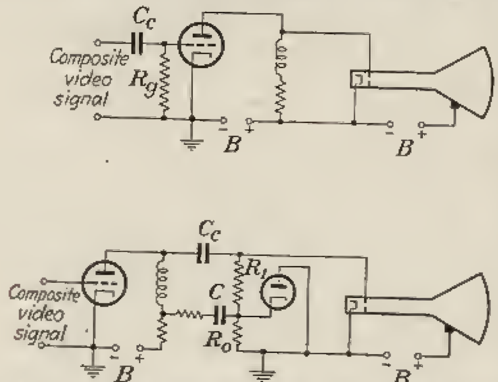


FIG. 58.—Direct-current restoration circuits: top, using the video amplif[ier] grid current for rectification; bottom, using a separate diode.

current for the cathode of the electron gun; and (3) a source of focusin[g] coil current (in the case of magnetostatically focused tubes).

A typical high-voltage power supply is shown in Fig. 59. It consists of [a] single-winding transformer of r-m-s output voltage equal to approximatel[y] $V_{dc}/1.4$, where V_{dc} is the desired output d-c voltage; two capacitors of roughl[y] 0.03 to 0.05 μf; a series filter resistor of roughly 100,000 to 500,000 ohms; an[d] a tapped bleeder resistor of about 5 megohms. A resistor is also connected i[n] series with the second-anode output tap to limit the total output current to a

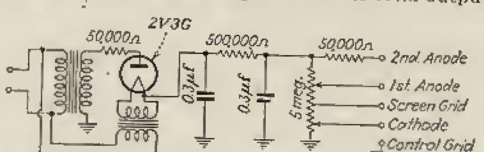


FIG. 59.—Typical anode voltage supply for a picture tube.

safe value in case of accidental contact by the operator. The taps require[d] for the various electrodes of an electrostatically focused electron gun ar[e] shown.

The current required for the focusing coil of a magnetostatically focus[ed] gun depends on the focus-coil design. A typical value is 100 ma at 25 volt[s] which may be obtained from the current drain of the receiver proper at th[e] sacrifice of 25 volts in the low-voltage power supply for the receiver.

The heater currents have usually one of two r-m-s a-c values: 2.5 volts at 2.1 amp. or 6.3 volts at 0.6 amp.

DEFLECTION OF ELECTRON BEAMS

63. Electron-beam Velocity. The amount of deflection suffered by an electron scanning beam depends on the velocity with which the electrons in the beam move. This velocity v is expressed by

$$v = 3 \times 10^{10}\sqrt{1 - \left(\frac{1}{2 \times 10^{-6}E + 1}\right)^2} \text{ cm per second} \tag{32}$$

where E is the accelerating voltage in volts (approximately equal to the second-anode voltage). This expression takes into account the change in electron mass with velocity. The values of v vary from 2.66×10^9 cm per second at 2,000 volts to 4.93×10^9 cm per seconds at 7,000 volts.

64. Electric Deflection. The deflection d, in centimeters, of the scanning beam across the screen of a picture tube caused by passage between parallel deflecting plates is given by

$$d = \frac{1.77 \times 10^{15} E_d l(D + \frac{1}{2}l)}{sv^2} \text{ cm} \tag{33}$$

where E_d = voltage in volts applied to the deflection plates
v = electron beam velocity in centimeters per second
l = length of the deflection plates in centimeters
s = separation between them in centimeters
D = distance from the screen end of the deflection plates to the center of the screen measured along the axis of the tube in centimeters.

Typical electrically deflected tubes have deflection sensitivities of from 0.15 to 0.35 mm deflection per volt applied to the deflecting plates, when operated at maximum rated second-anode voltage.

65. Magnetic Deflection. The deflection d, in centimeters, across the screen of a picture tube, caused by passage through a uniform magnetic field is given by

$$d = \frac{1.77 \times 10^7 BlD}{v} \text{ cm} \tag{34}$$

where B = flux density of the field in gauss
l = its length in centimeters
D = field-to-screen distance in centimeters
v = electron-beam velocity in centimeters per second.

66. Ion Spot. Negative ions liberated from the cathode of the electron gun are focused and deflected in much the same manner as the electrons. In electric deflection the deflection is independent of the charge/mass ratio of the particles; hence the ions and electrons are equally deflected. In magnetic deflection, however, the deflection depends on the square root of the change/mass ratio. Since the ions have masses several thousand times that of the electron, they suffer correspondingly small deflection. The lack of deflection subjects the center of the scanned area to continual bombardment by the ions, and this eventually results in the formation of a black or yellowish spot. The ion spot is characteristic of the combination of electrostatic focusing and magnetic deflection. Specialized electrode structures have been devised, however, which intercept the ions before they reach the screen.

When magnetostatic focusing is employed, the heavy ions are not brought to focus by the same value of magnetic field as are the electrons; consequently the bombardment by ions is spread over a larger area of the surface. Accordingly the combination of magnetostatic focusing and magnetic deflection is comparatively free from the ion-spot difficulty.

SCANNING AND SYNCHRONIZATION

67. Saw-tooth Generators. The saw-tooth wave form (Fig. 7) is generated for scanning purposes by the periodic charging and discharging of a capacitor. The charge-time curve is used to produce the active scanning motion, and the discharge curve forms the retrace. To main-

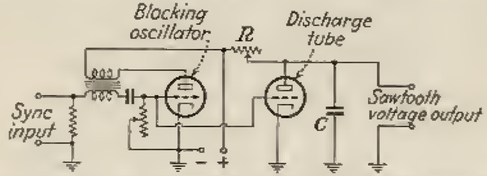


FIG. 60.—Blocking-oscillator type of impulse generator.

tain a linear charge curve, it is customary to restrict the charge time to about 0.4 time the *RC* product of the circuit, or less, and also to make use of the non-linear dynamic characteristic of the following amplifier to introduce a compensating non-linearity. Certain forms of multivibrator circuits may be used to produce saw-tooth waves directly.

Usually a separate *discharge tube* is used to discharge the capacitor. The discharge current is passed through a high-vacuum triode whose grid controls the timing of the discharge. The impulses applied to the grid of the discharge tube are usually derived from an impulse generator, although they may consist of the synchronizing signal itself properly amplified.

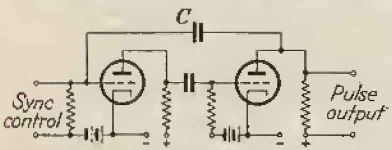


FIG. 61.—Multivibrator type of impulse generator.

Impulse generators used to control the discharge tube in scanning generators take one of two forms, the multivibrator or the blocking oscillator. The blocking oscillator (Fig. 60) consists of a grid-plate coupled oscillator whose grid is driven negative by the passage of grid current, thus blocking the oscillations suddenly. As the charge leaks off the grid through the grid resistor, the oscillations recommence, to be followed by the sudden blocking of the grid circuit. The sharp impulses appearing between the grid and ground are used to control the discharge tube as shown.

A multivibrator type of saw-tooth generator is shown in Fig. 61. This circuit operates by virtue of the connection between the plate circuit of the output tube and the grid circuit of the input tube. The alternate charge and discharge of the coupling capacitor can be used to produce either impulses or saw-tooth waves, depending on the circuit constants.

68. Production of Current Saw-tooth Waves. Saw-tooth waves of voltage produced by saw-tooth generators suffice to deflect the beam of an electrically deflected tube, which is a voltage-operated device, provided only that the peak-to-peak value of the saw-tooth wave is great enough to produce full deflection. In magnetically deflected tubes the deflection is proportional to the current in the deflection coils; hence saw-tooth waves of current are required. The voltage wave form required to produce saw-tooth waves of current depends on the inductance and resistance present in the scanning-coil windings. An "impulse" voltage wave is required for coils exhibiting a large inductance-resistance ratio. For lower L/R ratios the voltage wave form is a combination of impulse and saw-tooth waves. The several voltage and current wave forms for these cases are shown in Fig. 62.

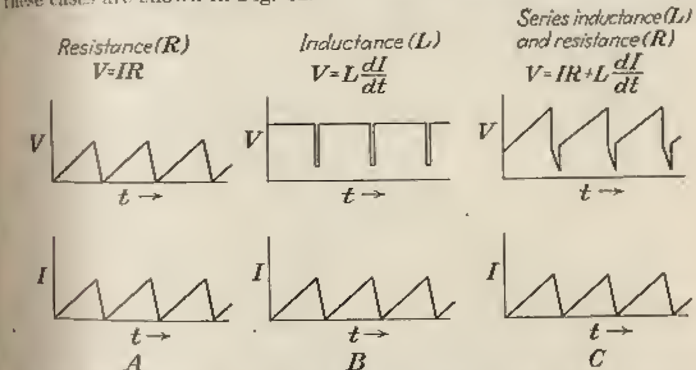


FIG. 62.—Saw-tooth waves of current and corresponding voltage waveforms in circuits of *R*, *L*, and *L-R*.

The part-impulse part-saw-tooth wave form may be produced simply by applying a saw-tooth wave to a series *RC* combination. The saw-tooth component develops across the resistance, while the impulse portion develops across the capacitance.

The presence of distributed capacitance in the scanning-coil windings gives rise to resonance oscillations when the impulse voltage wave form is applied. These residual oscillations may be damped out by connecting a rectifier tube and a shunt *RC* circuit in series across the scanning-coil terminals.

69. Amplification of Scanning Wave Forms. The preservation of the scanning wave form in the amplifier subsequent to the scanning generator is based on the considerations for video amplifiers. Usually it is desirable to pass the fundamental and 20 harmonics, which makes the range 60 to 1,200 cps for the vertical scanning amplifier and 15,750 to 315,000 cps for the horizontal amplifier. The phase and amplitude characteristics must be linear over these ranges.

For electric deflection it is essential that the scanning-generator output be disposed symmetrically with respect to the deflection plates, and this is carried out by employing push-pull amplification. The center point of the push-pull output is connected through a high resistance to the second-anode terminal of the picture tube. Care must be taken to all the necessary peak-to-peak voltage to develop across the amplifier outp without breakdown of insulation and excessive stress in the tube stru tures. The necessity for high scanning voltages has limited applicati of electric deflection to tubes operating below 3,000 or 4,000 volts, secon anode voltage.

In magnetic deflection, heavy current rather than high voltage required to secure full deflection. To secure the current, it is customar to employ a voltage step-down transformer in the output of the scanni amplifier. This transformer must meet the amplitude- and phase-fr quency characteristics of the amplifier itself. High voltage develo across the primary of this transformer as a result of the rapid changes current in the secondary. The amplifier tubes and other componen must be capable of withstanding these voltage peaks, which often atta several thousand volts amplitude.

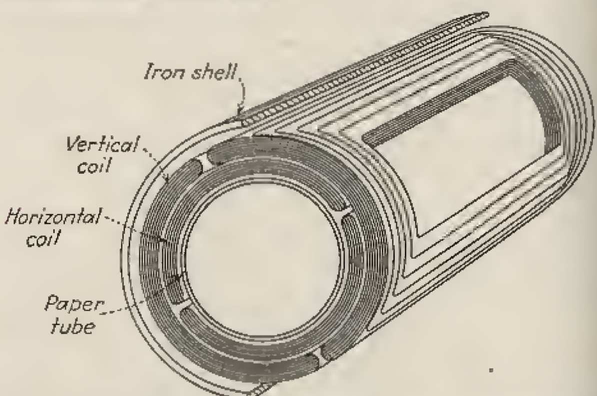


FIG. 63.—Construction of a typical scanning yoke (magnetic-deflection co system).

70. Scanning Yokes. The set of deflection coils required for magnet deflection is called a *scanning yoke*. It consists of two sets of coils. One arranged about a vertical axis transverse to the tube axis, produces th horizontal deflection, and another set of coils, arranged on a horizonta axis transversely to the tube axis, produces the vertical deflection.

Among the factors on which the yoke design depends are (1) the angl of deflection required (which determines the required number of ampere turns as well as the allowable physical length of the yoke); (2) the nece sity of providing a uniform field, to avoid defocusing the spot and distorting the orthogonal shape of the scanning pattern; and (3) the proportioning of the L/R ratio to secure linear deflection with a give deflection amplifier and output transformer.

CONTRAST AND GRADATION OF TELEVISION IMAGES

71. Over-all Brightness Transfer Characteristic. The ability of the television system to reproduce brightness contrasts and tonal gradation

is expressed by the over-all brightness transfer characteristic (Fig. 64). The ordinates give the range of brightness in the reproduced image (image brightness) corresponding to the range of brightness in the original object (object brightness) plotted in the abscissas.

The actual shape of this curve depends on the transfer characteristic (input-output relationship) of each item of equipment in the transmission system. In general the actual characteristics cannot be expressed in simple analytic form. However, if an idealization is made, the curves may be expressed in the following form:

$$B_i = k_0 B_0^{\gamma_0} \tag{35}$$

where B_i is the image brightness corresponding to the object brightness B_0, k_0 is the proportionality factor relating the image brightness scale to the object brightness scale, and the exponent γ_0 (gamma) determines the extent and direction of the curvature of the characteristics. For unity gamma

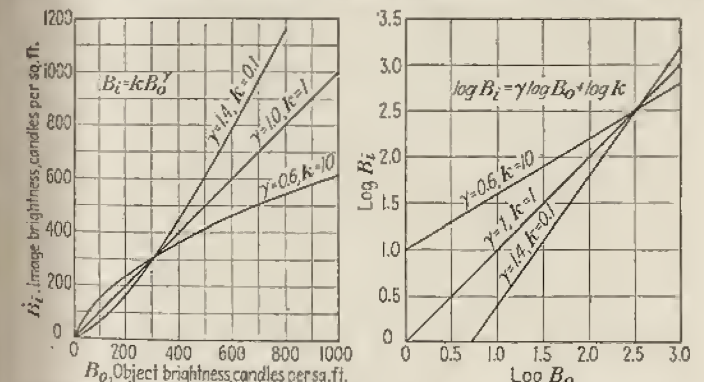


FIG. 64.—Brightness transfer characteristics of a television system.

($\gamma_0 = 1$) the relationship between B_i and B_0 is linear. For gamma greater than unity the curve has an "antisaturation" shape; for gamma values below unity the curve has a "saturation" shape.

The value of gamma determines the subjective contrast of the image as viewed by the observer, since the sensation of light in the mind is approximately proportional to the logarithm of the brightness. When Eq. (35) is expressed in logarithmic form

$$\log B_i = \log k_0 + \gamma_0 \log B_0 \tag{36}$$

all the relationships between $\log B_i$ and $\log B_0$ become linear and the slope of the lines is directly proportional to the gamma value. In consequence high contrast is produced by correspondingly high values of gamma.

72. Subsidiary Transfer Characteristics. The input-output characteristics of each subsidiary item of equipment in the system can be expressed by a similar relationship

$$(\text{Output}) = k\ (\text{input})^{\gamma}$$

where k relates the scales of the input and output quantities and γ is the gamma exponent describing the curvature of the characteristics. By combining each curve in the transmission system successively, equating the output of one device to the input of the succeeding device, it can be shown th the over-all gamma of the system is equal to the product of all the subsid gammas. In consequence, the effect of one item of equipment whose gam is lower than unity may be compensated by that of another whose gam is the inverse of the first. The gamma of iconoscope tubes, for examp lies at about 0.7, whereas that of picture tubes is about 2.0. Assuming t the subsidiary amplifiers, modulators, and demodulators are linear (gam unity), the over-all gamma is then $0.7 \times 2.0 = 1.4$, *i.e.*, the gamma is so what above unity. The orthicon camera, on the other hand, has a gamma unity, and the over-all gamma in this case would be 2.0, producing a consid ably more contrasty reproduced image. The desirable value of over gamma, following motion-picture practice, is between 1.2 and 1.7. The hi

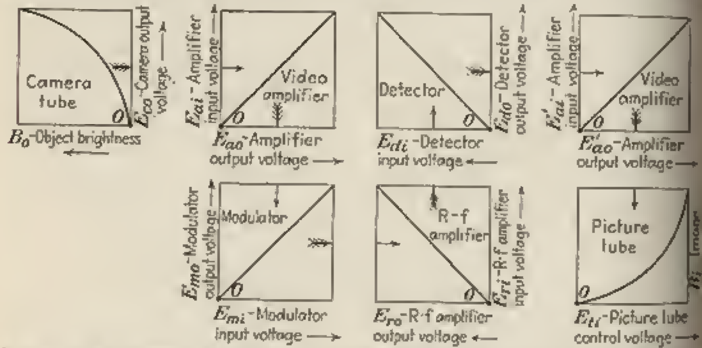


FIG. 65.—Transfer of object brightness to image brightness through su sidiary transfer characteristics of the elements of a television system.

value of gamma aids in restoring color contrasts lost through the monochr matic nature of the reproduction. It should be noted that high contrasts a limited by picture tube performance.

The values of the subsidary gammas also bear on the signal/no ratio of the system. If a transmitter gamma less than unity is employe most of the picture information consists of signal excursions havi amplitudes high on the dynamic characteristic, above the noise. compensating higher value of gamma in the receiver may be used produce an over-all value within the desirable range of 1.2 to 1.7.

References

Picture Tubes and Image Reproduction:

BACHMAN and CARNAHAN: Negative Ion Components in the Cathode Ray Beam, *Pro I.R.E.*, **26**, 529, May, 1938.

BURNAP, R. S.: Television Cathode-ray Tubes for the Amateur, *RCA Rev.*, **2**, No. 3, 29 January, 1938.

BURNETT, C. E.: A Circuit for Studying Kinescope Resolution, *Proc. I.R.E.*, **25**, 99 August, 1937.

Characteristics of Phosphors for Cathode-ray Tubes, *Electronics*, **11**, No. 12, 31, Dec ber, 1938.

EPSTEIN: Electron Optical System of Two Cylinders as Applied to Cathode-ray Tube *Proc. I.R.E.*, **24**, 1095, August, 1936.

IAMS, H.: A Fixed Focus Electron Gun for Cathode-ray Tubes, *Proc. I.R.E.*, **27**, 1 February, 1939.

JOHNSON, J. B.: The Cathode-ray Oscillograph, *Bell System Tech. Jour.*, **11**, 1, January 1932.

LANGMUIR, D. B.: Theoretical Limitations of Cathode-ray Tubes, *Proc. I.R.E.*, **25**, 977, August, 1937.

LAW, R. R.: High Current Electron Gun for Projection Kinescopes, *Proc. I.R.E.*, **25**, 954, August, 1937.

———: Contrast in Kinescopes, *Proc. I.R.E.*, **27**, 496, July, 1939.

LEVERENZ, H. W.: Problems concerning the Production of Cathode-ray Screens, *Jour. Opt. Soc. Am.*, **27**, 25, January, 1937.

——— and SEITZ: Luminescent Materials, *J. Applied Phys.*, **10**, 479, July, 1939.

LEVY and WEST: Fluorescent Screens for Cathode-ray Tubes for Television and Other Purposes, *Jour. I.E.E.*, (*London*), **79**, 11, July, 1936.

——— and ———: Luminescence and Its Application to Television, *Jour. Telev. Soc.*, **2**, 337, March, 1938.

McGEE and LUBSZYNSKI: E. M. I., Cathode-ray Television Tubes, *Television*, **12**, 78, February, 1939.

MALOFF, I. G.: The Cathode-ray Tube in Television Reception, "Television," Vol. I, p. 337, RCA Institutes Technical Press, New York, 1936.

———: Direct-viewing Type Cathode-ray Tube for Large Television Images, *RCA Rev.*, **2**, No. 3, 289, January, 1938.

———: Gamma and Range in Television, *RCA Rev.*, **3**, No. 4, 409, April, 1939.

——— and EPSTEIN: Theory of the Electron Gun, *Proc. I.R.E.*, **22**, 1386, December, 1934.

——— and ———: Luminescent Screens for Cathode-ray Tubes, *Electronics*, **10**, No. 11, 31, November, 1937.

——— and ———: "Electron Optics in Television," Chap. XII, McGraw-Hill Book Company, Inc., New York, 1938.

NOTTINGHAM, W. B.: Electrical and Luminescent Properties of Willemite under Electron Bombardment, *J. Applied Phys.*, **8**, 762, November, 1937.

SCHMIDLING, G. T.: Fluorescent Materials for Television Tubes, *Communications*, **14**, No. 4, 30, April, 1939.

WALLER, L. C.: Kinescopes for Television Receivers, *Communications*, **14**, No. 4, 20, April, 1939.

WILSON, J. C.: "Television Engineering," Chap. VIII, Sir Isaac Pitman & Sons, Ltd., London, 1937.

ZWORYKIN, V. K.: Iconoscopes and Kinescopes in Television, *RCA Rev.*, **1**, No. 1, 60, July, 1936.

——— and PAINTER: Development of the Projection Kinescope, *Proc. I.R.E.*, **25**, 937, August, 1937.

General Bibliography

General—Books:

FINK, D. G.: "Principles of Television Engineering," McGraw-Hill Book Company, Inc., New York, 1940.

LOHR, L. R.: "Television Broadcasting," McGraw-Hill Book Company, Inc., New York, 1940.

MALOFF and EPSTEIN: "Electron Optics in Television," McGraw-Hill Book Company, Inc., New York, 1938.

WILSON, J. C.: "Television Engineering," Sir Isaac Pitman & Son, Ltd., London, 1937.

ZWORYKIN and MORTON: "Television," John Wiley & Sons, Inc., New York, 1940.

Periodicals:

ASHBRIDGE, N.: Television in Great Britain, *Proc. I.R.E.*, **25**, 708, June, 1937.

BEAL, R. R.: Equipment Used in the Current RCA Television Field Tests, *RCA Rev.*, **1**, No. 3, 36, January, 1937.

BEERS, ENGSTROM, and MALOFF: Some Television Problems from the Motion Picture Viewpoint, *Jour. Soc. Mot. Pict. Eng.*, **32**, 18, January, 1939.

BROLLY, A. H.: Television by Electronic Methods, *Trans. A.I.E.E.*, **53**, 1153, August, 1934.

CONKLIN and GIHRING: Television Transmitters Operating at High Powers and Ultra-high Frequencies, *RCA Rev.*, **2**, No. 1, 30, July, 1937.

EDDY, W. C.: Television Studio Considerations, *Communication and Broadcast Eng.*, **4**, No. 4, 12, April; **5**, 14, May; **6**, 20, June; **7**, 17, July, 1937.

———: Television Lighting, *Jour. Soc. Mot. Pict. Eng.*, **33**, 41, July, 1939.

ENGSTROM, BEERS, and BEDFORD: Application of Motion Picture Film to Television, *Jour. Soc. Mot. Pict. Eng.*, **33**, 3, July, 1939; see also *RCA Rev.*, **4**, No. 1, 48, July, 1939.

ENGSTROM and HOLMES: Television Receivers, a series of six articles: Part I, Antenna and R-f Circuits, *Electronics*, **11**, No. 4, 28, April, 1938; Part II, Television I-f Amplifiers, *ibid.*, **11**, No. 6, 20 (June, 1938); Part III, Television V-f Circuits, *ibid.*, **11**, No. 8, 18, August, 1938; Part IV, Television Synchronization, *ibid.*, **11**, No. 11,

18, November, 1938; Part V, Television Deflection Circuits, *ibid.*, **12**, No. 1, 1 January, 1939; Part VI, Power for Television Receivers, *ibid.*, **12**, No. 4, 22, April, 1939.
———, Kell, Bedford, Trainer, Holmes, Carlson, and Tolson: An Experimental Television System, *Proc. I.R.E.*, **22**, 1241, 1246, 1266, November, 1934.
Fink, D. G.: A Laboratory Television Receiver (in six parts) *Electronics*, Part I, *ibid.* **11**, No. 7, 16 July, 1938; Part II, *ibid.*, **11**, No. 8, 26, August, 1938; Part III, *ibid.* **11**, No. 9, 22, September, 1938; Part IV, *ibid.*, **11**, No. 10, 16, October, 1938; Part V, *ibid.*, **11**, No. 11, 26, November, 1938; Part VI, *ibid.*, **11**, No. 12, 16, December, 1938.
———: A Television Receiver for the Home, *Electronics*, **12**, No. 9, 16, September, 1939.
Gannett and Green: Wire Transmission System for Television, *Bell System Tech. Jour.*, **6**, 616, October, 1927.
Goldmark, P. C.: A Continuous Type Television Film Scanner, *Jour. Soc. Mot. Pict. Eng.*, **33**, 18, July, 1939.
———: Problems of Television Transmission, *J. Applied Phys.*, **10**, 447, July, 1939.
Goldsmith, A. N.: Television Economics, *Communications*, **14**, No. 2, 18, February; No. 3, 17, March; No. 4, 26, April, 1939.
Gray, Horton, and Mathes: The Production and Utilization of Television Signals, *Bell System Tech. Jour.*, **6**, 560, October, 1927.
Hanson, O. B.: Experimental Studio Facilities for Television, *RCA Rev.*, **1**, No. 4, 3, April, 1937.
Ives, H. E.: Television, *Bell System Tech. Jour.*, **6**, 551, October, 1927.
———: Transmission of Motion Pictures over a Coaxial Cable, *Jour. Soc. Mot. Pict. Eng.*, **31**, 256, September, 1938.
Kaar, I. J.: The Road Ahead for Television, *Jour. Soc. Mot. Pict. Eng.*, **32**, 18, January, 1939.
Lubcke, H. R.: An Introduction to Television Production, *Jour. Soc. Mot. Pict. Eng.*, **33**, 54, July, 1939.
Morris and Shelby: Television Studio Design, *RCA Rev.*, **2**, No. 1, 14, July, 1937.
Protzman, A. W.: Television Studio Technic, *Jour. Soc. Mot. Pict. Eng.*, **33**, 26, July, 1939.
Seeley and Kimball: Transmission Lines as Coupling Elements in Television Equipment, *RCA Rev.*, **3**, No. 4, 418, April, 1939.
Strieby, M. E.: Coaxial Cable System for Television Transmission, *Bell System Tech. Jour.*, **17**, 438, July, 1938.
Television Transmitters, *Electronics*, **12**, No. 3, 26 March, 1939.
Wilson, J. C.: Trichromatic Reproduction in Television, *Jour. Roy. Soc. Arts*, **82**, 841, July, 1934.
Zworykin, V. K.: The Iconoscope, a Modern Version of the Electric Eye, *Proc. I.R.E.*, **22**, 16, January, 1934.
———: Television, *Jour. Franklin Inst.*, **217**, 1, January, 1934.
Kallman, H. E.: The Gradation of Television Pictures, *Proc. I.R.E.*, **28**, 170, April, 1940.
Rosenthal, A. H.: A System of Large-Screen Television Reception Based on Certain Electron Phenomena in Crystals, *Proc. I.R.E.*, **28**, 203, May, 1940.
Goldmark and Dyer: Quality in Television Pictures, *Proc. I.R.E.*, **28**, 343, August, 1940.
Baldwin, M. W.: The Subjective Sharpness of Simulated Television Images, *Proc. I.R.E.*, **28**, 458, October, 1940.

SECTION 20

FACSIMILE

By R. E. Mathes,[1] B.S.

1. General Considerations. The term *facsimile* has been applied to that branch of the science of graphic electrical communication which endeavors to convey the physical form, and even the light shadings of the original subject matter. Such information cannot be instantaneously or simultaneously transmitted, and it is thus necessary to do so bit by bit sequentially. The manner of doing this is to divide effectively the original into a large number of elemental areas and to transmit signals to indicate the relative light shades of these areas. Such shades are then reproduced more or less accurately at the receiver. The elemental areas are recorded in the same sequence, thus building up the record similar to the building of a brick wall.

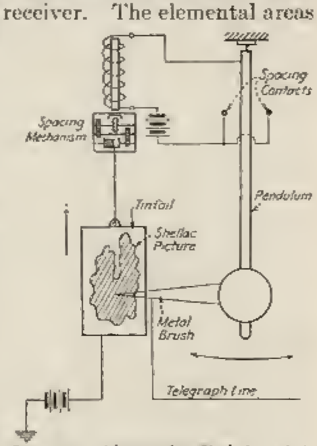


Fig. 1.—Alexander Bain's original picture apparatus.

The accuracy of reproduction depends upon the number of these elemental areas in the picture. It makes no difference as to the size of the finished record; the resolving power is entirely a matter of the number of elemental areas satisfactorily transmitted. It takes just as many tiny areas to represent a face well on a postage stamp as it does to represent a face well in larger areas on a 10-ft. enlargement.

To transmit sequentially and to record these areas with necessary fidelity requires highly accurate mechanisms, synchronizing means, communication circuits, amplifier circuits, and scanning and recording devices. Such means have been the subject of intensive development efforts for many years.

Because of the close similarity of requirements and equipment for both landline and radio facsimile transmissions, it has seemed best to treat the subject of facsimile in all its phases.

2. Historical. In 1842 Alexander Bain proposed a facsimile system which had in it all the pertinent functions included in the most modern derivations. Figure 1, a copy of his original system, shows these clearly. Synchronous action was afforded by the pendulums at transmitter and receiver, line advance by moving the message plates upward a short

[1] RCA Communications, Inc.

distance at the end of each swing of the pendulum, and elemental ar[ea] scanning by the contact of the metal brushes. Caselli produced a[n] improved system in 1865, Korn another in 1902, Belin in 1920. Th[e] American Telephone and Telegraph Company opened a public service [in] the United States in 1925, and the Radio Corporation of America inaugu[]rated public service with London in 1926. An excellent bibliograph[y] covering this growth is given by J. L. Callahan.[1]

3. Transmission. The functions necessary to transmit a facsimil[e] record are as follows:

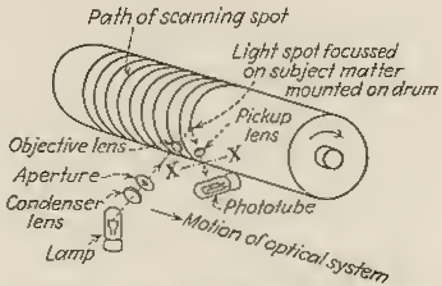


FIG. 2.—Depicting necessary elements for scanning.

1. A scanning system to explore the elemental areas of the subject [and] identify their individual light shadings in terms of an electrical current.
2. A modulation circuit to provide this fluctuating current in a form sui[t]able for transmission over the communication system available.
3. A mechanism to provide a[n] orderly exploration of these are[as] by the scanning system.
4. An electrical drive syste[m] to run the mechanism at a un[i]form, predetermined rate, with[in] close tolerances.

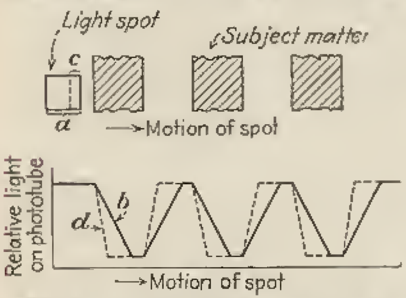


FIG. 3.—Distorting effect of finite width of the scanning light spot or aperture. (a) and (b) spot width = three-fourths width of narrowest line to be scanned. (c) and (d) spot width = one-fourth width of narrowest line to be scanned.

4. Scanning. Modern met[h]ods invariably use an element[al] area of intense light, eith[er] transmitted through or re[]flected from the original subje[ct] and picked up by a phototub[e]. The diaphragm is sometim[es] placed at point X-X after th[e] pickup lens rather than at th[e] condenser lens. The effectiv[e] light of the scanning system [is] proportional to the product o[f] the intrinsic intensity of th[e] light source and the solid angle subtended by the objective lens at th[e] surface of the subject.

[1] CALLAHAN, J. L., A Narrative Bibliography of Radio Facsimile,"Radio Facsimile," RCA Institutes Technical Press, 1938.

Ideally, the elemental area will be of infinitesimal width. This cannot be realized practically, and therefore all scanners have an effective light spot of finite width. This gives rise to a distortion known as the *aperture distortion* which modifies the electrical signals so they are not a true representation of the instantaneous changes of the shadings or tonal values of the subject.

In Fig. 3 a light spot of a width nearly as great as that of the finest vertical bars is shown at *a* and the resultant electrical response at *b*. Such relatively great finite width of the spot produces a trapezoidal wave form which becomes triangular as the spot width becomes just equal to that of the vertical lines of the subject. On the other hand, a narrower spot, such as *c*, will produce a wave form *d*, which, although still trapezoidal, approaches a true rectangular or "square" wave shape as the spot width approaches zero.

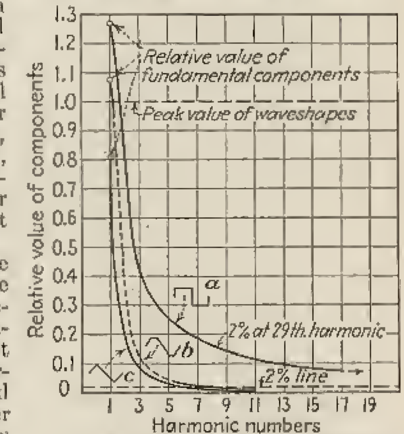


FIG. 4.—Relative value of harmonic content of signal produced by various width apertures: (*a*) for square wave shape from aperture of infinitesimal width; (*b*) for trapezoidal wave shape from aperture two-thirds width of narrowest line; (*c*) for triangular wave shape from aperture width equal to narrowest line. Also note relative value of fundamental component for each.

The narrow spot will permit the interpretation of more detail of the subject but will result in an electrical wave form which has pertinent and necessary component frequencies (harmonics) considerably higher than those produced by the wider spot. In other words, the aperture distortion has an action approximately equivalent to that of a low-pass filter having a perfect phase characteristic. Figure 4 shows the manner in which the amplitude of the harmonic components decreases for three different apertures.[1] The curve for the rectangle corresponds to an aperture of infinitesimal width; that for the trapezoid corresponds to an aperture whose width is two-thirds that of the vertical line to be scanned (Fig. 3); that for the triangle corresponds to an aperture whose width is equal to that of the vertical line. It will be seen that both the triangle and trapezoid components drop to a negligible value very quickly, whereas the value of rectangular components does not drop to 2 per cent until about the twenty-ninth harmonic. The peak amplitude of these shapes were all taken as 1.0.

The practical determination of the detail required depends upon the use to be made of the record. However, commercial experience to date teaches that, for an ultimate enlargement of the recording by not more than 4 to 1 over the original, a texture of 100 to 120 scanning lines per inch is ample. One system in extensive operation uses a texture of 200

[1] JOLLEY, L. B. W., "Alternating Current Rectification," John Wiley & Sons, Inc., New York, 1928.

lines per inch and this, of course, permits of still greater enlargement of the recording. If detail is set as equal in both dimensions, a minimum width of vertical lines is indicated as about 0.008 in. It is therefore necessary merely to produce a sufficiently accurate representation of a line of this width on the record. Practically, considerable inaccuracy may be permitted because the minimum detail the average eye can differentiate is that area which subtends an angle at the eye of 0.000 to 0.00070 radian (approximately 2 minutes). It is found that a rectangular aperture 0.006 in. wide is a practical compromise, which reduces the pertinent frequency components as much as is permissible while preserving a sufficient detail in the record.

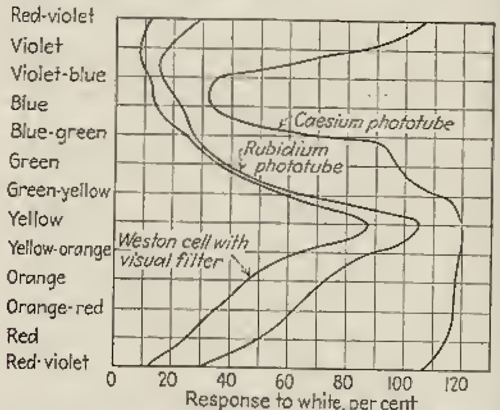


Fig. 5.—Showing spectral characteristic of standard caesium oxide and rubidium phototubes. Curve of Weston cell with visual filter—shown for comparison—is very close to the stimulus of various colors on the normal eye.

Another phase of scanning is the spectral distribution of the light reflected from the surface of the subject, as referred to the distribution of the phototube sensitivity. Figure 5 shows curves of such distribution. It will be seen that the curve for the rubidium tube fairly closely follows that of the eye, *viz.*, it is about equally panchromatic. Use of this type tube will give a black and white recording in which the various colors, as well as the tonal values, are given a weighting closely approximating that assigned by the eye and results in a more effective reproduction, even though the original was in color. This is of importance for facsimile service because the system should be capable of handling any type of subject matter that may be submitted. Unfortunately, the sensitivity of rubidium is low, and this type tube can be used only where there is an excess of light available to the phototube. Therefore the caesium-surfaced phototube is the type more generally used.

The scanning system usually includes the phototube and its immediately associated amplifiers. The light intensity available at the phototube is very low—on the order of 0.001 to 0.005 lumen—and the voltage output of this tube is likewise low. It can be increased by increasing the load resistance, and this has sometimes been made as high as 10 to

25 megohms. However, the interelectrode capacitance of the tube becomes serious when shunted across such high value, even though these capacitances are of the order of 1 to 2.5$\mu\mu$f. The effect is that of a low-pass filter to limit the higher frequency components and thus limit the possible scanning speed of the facsimile equipment.

Use of lower phototube load resistance—½ to 1 megohm—to permit of sufficiently high frequency response for higher speed operation makes it difficult to amplify the weak output. The simplest way would be to use direct resistance-coupled amplifiers. However, the variations in voltages, emission, and contact potential are of the same order of magnitude as the desired signal and in such amplifiers are superposed on the signal. One

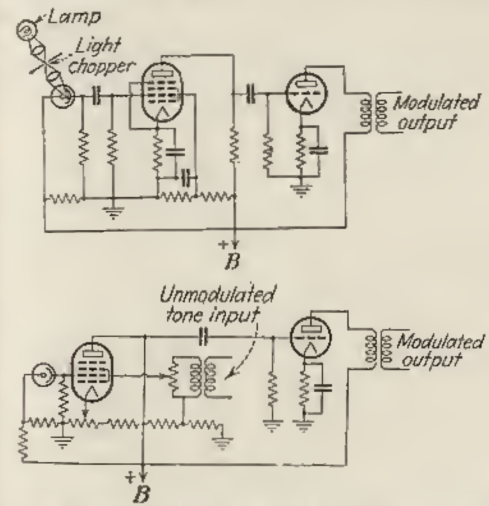


Fig. 6.—Typical scanner amplifier circuits: (above) a-c amplifier for use with light chopper in the optical system; (below) modulator amplifier for use with d-c output of phototube.

solution for this difficulty is to modulate the light beam at an a.f. and use conventional a-c amplification. The modulation can be applied directly to the lamp in the case of a gaseous light source, such as neon, helium, or mercury vapor, or it may be accomplished by cutting the beam by a mechanical chopper, such as a string galvanometer, vibrating reed, chopper disk, etc.

Another solution is to apply the phototube output to a sensitive modulator, such as a balanced bridge circuit, etc., and to amplify the resulting modulated tone. Figure 6 (above) shows a typical arrangement of the first type and Fig. 6 (below) that of the second type.

In either of these types it is essential that the audio tone which acts as the carrier for the facsimile modulation be of a high enough frequency that the shortest signal to be sent (*e.g.*, a line 0.008 in. wide) be composed of enough tone cycles to form a sufficiently accurate envelope.

5. Modulation. The signals produced by the scanner may be tra mitted over different types of communication systems, and there numerous ways in which the signals may be applied as modulation these systems. Those of present commercial importance will be outlin briefly.

1. *For Radio Circuits:*
   a. "CFVD" type of time modulation.
   b. Subcarrier f.m.
   c. Amplitude modulation on u-h-f circuits.
2. *For Landlines:*
   a. Double side-band a.m.
   b. Single side-band a.m.
3. *For Oceanic Cables:*
   a. Direct-current transmission.

For radio circuits the CFVD (constant frequency variable dot) syst has been extensively used since 1931 on long-distance short-wave circu and at centers such as New York, San Francisco, Buenos Aires, Lond

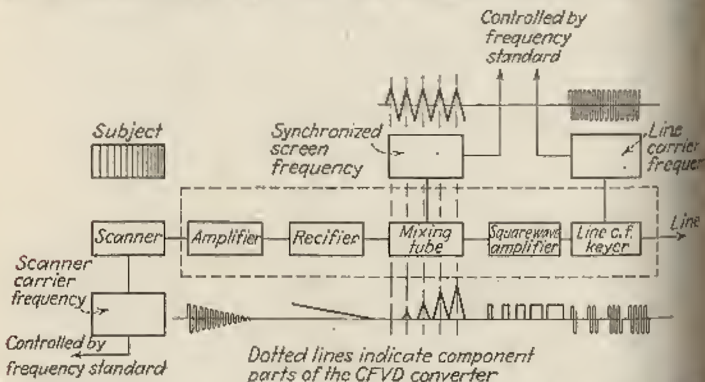


FIG. 7.—Block diagram of CFVD system for radiotelegraphic transmis of half tones.

Berlin, Melbourne, etc. The system comprises synthesizing the t values of the subject by means of dots sent at a constant frequency (ab 100 cycles) but of varying length. These are transmitted telegraphic by means of full 100 per cent keying of the r-f carrier and are record directly as dots. The recorded copy then has very much the appear of a screened half-tone picture in a newspaper. To obtain the screen is necessary to choose the screen frequency such that

$$\frac{f_s}{N} = \frac{n}{2}$$

where f_s = screen frequency in cycles per second
N = number of complete lines scanned per second
n = any odd integer.

Such limitation of the screen frequency may be avoided by provid a cam and contacts at the scanner to reverse the phase of the screen inp

to the modulator at the end of each line scanned. However, it is also best to choose the frequency such that the number of dots per inch of scanning line equals one-half the number of scanning lines per inch; measured perpendicularly to the direction of scanning. Figure 7 is a chart showing the functioning of this scheme, and Fig. 8 is the fundamental diagram of the CFVD converter.

Suggestions for improvement over this system have been elicited by the streaking in the recordings caused by fading and multipath variations inherent on long-distance radio circuits. One thought advanced is to transmit the mark and space intervals on different radio frequencies, using the start of each new signal to trigger off the recorder. Another thought is to transmit very short pulses (½ to 2 millisec.) for the start of each mark interval and to utilize interspersed locally generated pulses at the recording office to start the spacing intervals.

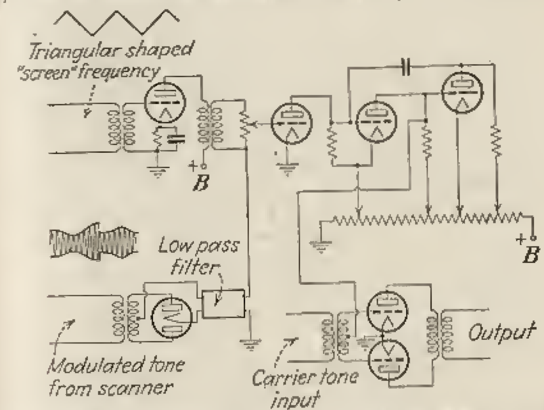


FIG. 8.—Schematic diagram of transmitting converter for CFVD.

Recently some of the long-distance commercial circuits have adopted *subcarrier frequency modulation*[1] in lieu of the CFVD system. The input from the scanner is rectified and applied to a push-pull triode, which in turn acts as a variable resistance in series with balanced trimmer condensers connected across the tank circuit of an i-f oscillator. The variation in resistance of the triode varies the effectiveness of the trimmer condenser on the natural period of the tank circuit, and thus provides f.m. in accordance with the light variations of the scanned subject matter. Another i-f oscillator of fixed frequency is beat against the first one and the difference in frequency taken to the output as in the conventional beat-frequency oscillator. The result is an audio output which is frequency-modulated over a relatively large percentage of the audio mid-frequency. This output is applied to a radio telephonic transmitter as a.m.

The principle of f.m. can, of course, be applied directly to the r-f carrier instead of through the medium of a subcarrier, but this requires special-

[1] MATHES, R. E., and J. N. WHITAKER, Radio Facsimile by Subcarrier Frequency Modulation, *RCA Rev.*, October, 1939, pp. 131–158.

ized equipment at both the radio transmitter and radio receiver, wh is not required in the above scheme. Furthermore, it does not overco any variations in the audio equipment or on the control lines as does t subcarrier method. Phase modulation can also be used if the rad propagation conditions permit.

For landlines the standard procedure is to transmit the amplitu modulated subcarrier tone, *viz.*, the signal output of the scanner. Nec sary amplifiers are used to provide the desired level, and impedan matches to the telephone lines which are used as the transmission circui Most systems now couple to the lines directly through standard rep coils.

Both double and single side-band transmission is extensively use In both cases care is taken not to utilize the frequencies below abo 1,000 cycles because of the inherent poor phase characteristic of w

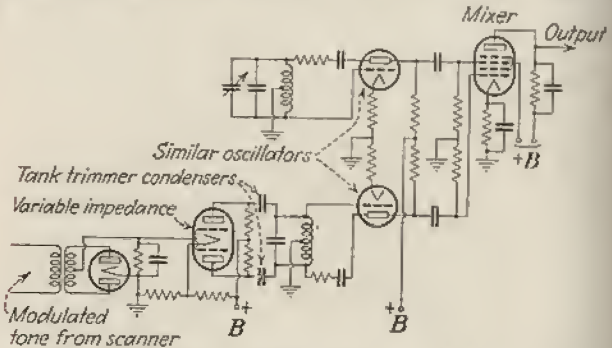


FIG. 9.—Schematic diagram of transmitting converter for subcarrier f-+ method.

lines at the low frequencies. In the case of single side band it is usuall the upper side band that is suppressed. The frequency of the carri tone used varies in different systems from about 1,800 to 5,000 cycle The exact value chosen is usually dependent upon the h-f characteristic of the line or channel to be used. It is made as high as is possibly con sistent therewith.

For Oceanic Cable. A new method for the transmission of pictures on the transoceanic cables was put into service in 1939. This system is th only practical system which transmits the picture signals directly withou any intermediary modulation or carrier. The phototube current i built up by d-c amplifiers and applied to the cable in that form.

6. Mechanisms. In order that the elemental areas be scanned sequen tially, it is necessary that mechanical means be provided to move the scanning light spot relative to the subject matter in such a manner that the entire area to be transmitted is covered in a predetermined order Many different sequences have been proposed in the past, *e.g.*, scanning alternate lines, diagonal and crisscross patterns, etc. However, the simple uniform scanning across the width of the subject, with line-by-line advance along the length of the subject is the easiest and the most readily

adaptable to various mechanisms. The relative motion may be obtained by moving either the light spot or the subject, or both. The latter is the more usual, in which the light spot is moved along one dimension of the subject and the subject itself is moved along its other dimension.

Figure 10 shows the most popular scheme, in which the subject is wrapped around a drum which revolves relatively rapidly in front of an optical and phototube pickup assembly. This assembly is carried on a track and caused by a lead screw to move along parallel to the axis of the drum at a relatively slow rate. This rate is so chosen by proper gear ratios that it will travel axially exactly the width of one scanning line for each revolution of the drum. In some designs the optical and pickup system is mounted in a fixed position, and the drum is slowly advanced along its axis as well as being revolved rapidly.

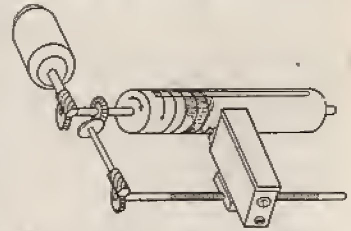


FIG. 10.—Drum type of machine.

Figure 11 indicates a different arrangement in which the optical system is comprised of two, three, or four identical lens systems for projecting the light spot on to the subject matter and for picking up the reflected light. These are mounted concentrically like spokes of a wheel and are revolved rapidly. The subject matter is applied face down to a one-half, one-

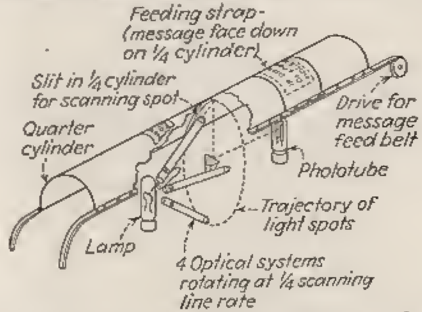


FIG. 11.—Quarter-cylinder type of continuous feed machine.

third, or one-fourth section of cylindrical tubing which is mounted concentric with the optical system. The subject is moved slowly along the cylinder by a belt or other device. This scheme has the great advantage that it can handle any size or thickness of subject matter so long as but one dimension does not exceed the peripheral length of the cylindrical section. Also successive subjects can be fed to the machine without need for stopping the machine or, alternatively, providing removable drums and rather complicated clutch mechanisms to prevent loss of synchronism. However, it has the severe handicap—yet to be fully overcome—of requiring an ultrafine degree of mechanical precision because all the optical systems must be exactly equal in optical efficiency as well as track perfectly.

Figure 12 illustrates a scheme of reciprocating motion in which a si optical system is swept across the inner face of a cylindrical section a then swept back in the reverse direction. Again high precision is requ so that the alternate sweeps will line up exactly at each point. T has been overcome in some systems for home recording by using o one direction of sweep, with the o being idle. This, of course, is q inefficient and can be considered for certain limited applications, where simplicity and cost are pred inating factors.

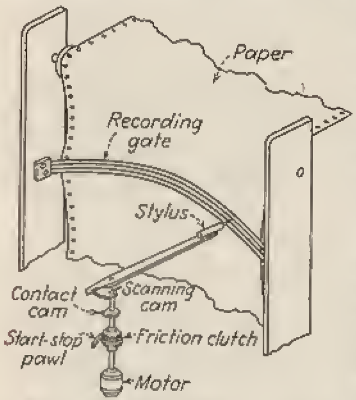


FIG. 12.—One type of start-stop recorder used for home reception.

The precision of the mechan must be of high order in any p facsimile system, but the ex requirement is practically depen upon the definition of which th cording system is capable. The p tographic method is the best f this standpoint, and in this cas instantaneous hunting, as betw successive lines, of more than deg. of drum rotation is a serious er Likewise, inaccuracies in the l screw, such as spacing variat between successive lines of 0.001 are quite noticeable in the recor particularly if the variations are a periodic nature occurring every 10 or so lines. Gear ripples, dynamic unbalance of the drum or mo minute eccentricities, voltage fluctuations, all must be carefully guar against.

Often cam and contact devices are applied and lined up with interval on the drum between the trailing and leading edges of the subj

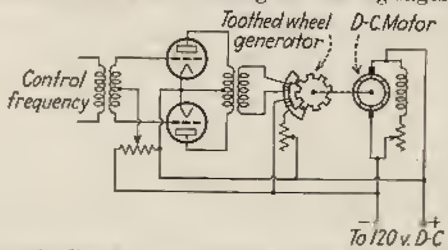


FIG. 13.—Schematic diagram of "Hammond brake" speed control for motors.

or "phasing line." These are used for special functions such as sendi synchronizing or level-control signal or for reversing polarity as in CFVD system. They are usually carried on either the drum or dri shaft.

Many commercial or military applications require a portable tr mitter. Particularly is this true for news picture work where the sou

of the subject matter may be anywhere at any time. Several designs of portable equipment have been produced for mounting small, comparatively light units in carrying cases. However, it is most difficult to get and retain the same precision as in equipment permanently installed.

7. Electrical-drive Systems. The mechanism may be variously driven by controlled d-c motors, 60-cycle synchronous motors, low-speed phonic wheels, and, lately, by the Alnico type of toothed wheel which operates directly at some 75 r.p.m. from 60 cycles, 120 volts a.c., or they may be designed to operate at a speed such as 100 r.p.m. from a drive frequency of 160 cycles. These may be driven from ordinary vacuum-tube power amplifiers. Some use one large motor to drive both the transmitting mechanism and a local monitor recorder, and others use two small motors, one for the drum and the other for the lead screw. One scheme of controlling a d-c motor which has been used for years is the Hammond brake

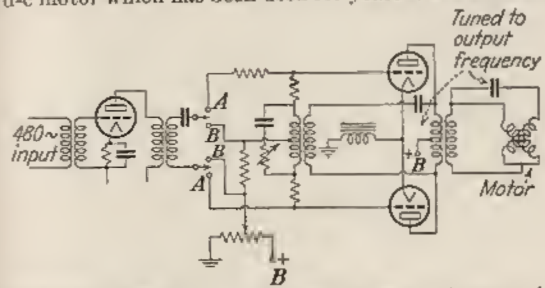


FIG. 14.—Schematic diagram of controlled thyratron inverter for driving synchronous motors.

scheme, shown in Fig. 13, and a scheme for driving a standard fractional-horsepower synchronous motor is a push-pull thyratron inverter circuit, such as shown in Fig. 14.

RECEPTION

For purposes of reception there must be the following:

1. A recording system which will translate the signals into visual markings.
2. A sensitive surface to receive such markings.
3. A mechanism to provide an orderly relative motion of the record surface.
4. Necessary filters, amplifiers, rectifiers, etc.
5. Synchronizing means to maintain the mechanism in instantaneous phase with the transmitter.

8. Photographic Recorders. The system most used records photographically on film. It has the advantages of giving the best detail and quality of definition and also readily provides a negative from which further processing, such as newspaper reproduction, can best be accomplished. It is usual to employ a thin base film that can be easily handled on a drum. The sensitive emulsion should be chosen for linearity of light response rather than extreme contrast. It should develop rapidly, and special films have been made which can be developed, fixed, and dried quickly so the picture can be used as soon after reception as possible.

Recorders take many forms. Glow tubes of neon, argon, helium, or a mixture of gases are formed so the glow takes place in a crater, thus confining the glow and giving essentially a point source of light of h intrinsic brilliance. In this tube the light intensity follows direc the modulation of the signal.

Other schemes use a local light source of fixed amplitude, such filament or arc lamps of various types. The light is then passed throu a modulating device before it strikes the film. One method sometim used in Europe is the Kerr cell coupled with Nicol prisms to polarize t light.

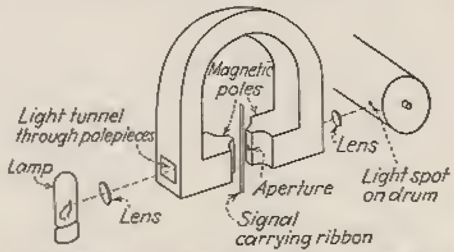


FIG. 15.—Schematic of optical system using the light valve.

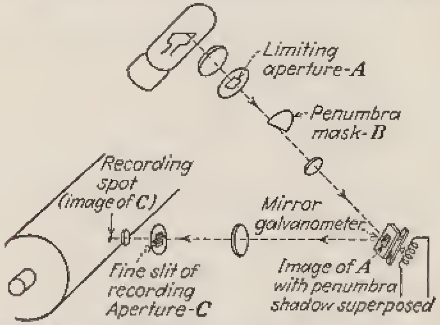


FIG. 16.—Elements of optical system with mirror galvanometer and line penumbra mask.

Another scheme is the light valve depicted in Fig. 15, in which a co ducting ribbon is suspended in a strong magnetic field. Signal pas through it causes the ribbon to move at right angles to the field. T ribbon is placed in the path of the light, and its movement acts a shutter to widen or narrow the slit of light passed through. If t ribbon is mounted parallel to the axis of the drum, this results in "va able-density" recording. If mounted perpendicular to the axis, it res in "variable-width" recording. Both methods are similar to those in recording sound on film in the motion-picture field. For facsim the former is by far the better.

A fourth scheme utilizes a D'Arsonval galvanometer or some modifi tion. In this case a small mirror is mounted on the signal-carrying wir or coil set in the magnetic field. The signal causes an angular moveme of the mirror and so changes the orientation of the reflected light on

TABLE I.—TYPICAL DIMENSIONS OF RECORDERS FOR FACSIMILE BROADCAST RECEPTION

	System *A*	System *B*	System *C*
1. Type of scanning...	Reciprocating	Rotating drum	Reciprocating
2. Scanning lines per minute..........	60	75	100
3. Total length of scanning line..........	$4\frac{3}{16}$ in.	$8\frac{3}{4}$ in.	6 in.
4. Scanning lines per inch..........	100	125	100
5. Synchronizing method.	Transmission of a separate subcarrier	Connection to common power system	Connection to common power system
6. Tone-carrier frequency..........	2,000 cycles	3,000 cycles	3,200 cycles

fixed aperture. This is diagramed in Fig. 16, which also shows a linear penumbra which permits of variable-density recording without critical adjustments of the galvanometer or signal level.

9. "Direct" Recording. Various schemes have been worked out for methods of "direct" recording in which the record appears almost instantaneously and need not be developed and fixed as with a photographic film.

Hot air has been used to discolor a presensitized paper, or to evaporate an opaque coating to permit the color of the paper base to show through. The air stream was keyed or triggered by the signal. An alcohol ink has been vaporized and blown on a glossy paper on the drum. It was keyed by a shutter on a signal relay. This has been used considerably for monitoring CFVD transmissions.

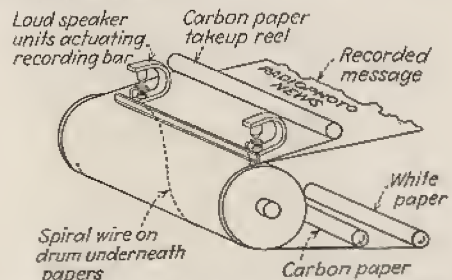


FIG. 17.—Elements of bar-and-spiral type recorder with carbon recording.

One group of commercial equipment now uses a stylus of small contact area, bearing on a coated dry paper. The signal current passes through the paper from the stylus to the drum, and the heat generated vaporizes the coating, permitting the black paper base to show through. A modification which may be developed in the future is the use of a spark discharge at a stylus or point. Heat, ultraviolet radiation, or direct electrolytic action might accomplish the actual recording.

Another important field of direct recording is that of electrochemical action. Certain groups of molybdenum and other salts will change color when current passes, as will also the organic azo dyes. The process is sensitive and very rapid, but difficulty is experienced in obtaining a stable reaction that will hold its color for a matter of years. Satisfactory recordings are being made at a speed in excess of 1,500 ft. per min linear spot travel. The definition at present is much less than with photographic recording but is ample for certain services such as message handling and home recording.

Much work has also been done to utilize carbon recording. In this process ordinary cheap carbon paper is laid with its face against plain white paper and the combination advanced slowly between a rapidly rotating drum carrying a raised spiral of wire and an axially parallel bar. The bar is driven by loud-speaker-type magnets to vary the pressure between it and the spiral, in accordance with the signal. The mechanism is indicated in Fig. 17. This method is best adapted to black and white recording but can be used to record linearly half-tone pictures if the amplifiers are carefully compensated for the non-linear pressure-density characteristic of the carbon paper.

10. Recording Sheets. Some slight use is made of photographic recording of a positive directly on paper for delivery, but the majority of the work is negative recording on film stock which is itself delivered, or from which a delivery print is made. The emulsion is chosen to match as nearly as possible the spectral distribution of the recording lamp. With this requirement it is also desirable that the emulsion have a good linear region, that the gamma be fairly high to require a lesser intensity rating of the recording spot, and that it be relatively color blind, if possible, to permit working in the darkroom under safelights. For black and white recording, as with the CFVD system, a highly contrasty and sensitive film is desired and a commercial process film is used. For linear recording a less contrasty and more sensitive film is best, and special emulsions have been made available which lie approximately between an orthochromatic and a panchromatic as regards color response.

The paper used for carbon recording is a cheap grade that merely has a fine enough grain not to cause undue loss of the definition inherent in the method. The paper for opaque coatings is a jet black of rather coarse grain and of only nominal mechanical strength; whereas that for impregnating baths must have proper absorption properties, as well as mechanical strength. This is because most of these systems use the paper in a damp condition or else pass it through a liquid bath just prior to recording. In either event it must be tough enough to withstand the feeding and rolling stresses while wet.

11. Recording Mechanisms. The recording system utilizes mechanisms of the same general type as for transmission, some of which have already been indicated, *e.g.*, in Figs. 10, 12, and 17. In this case, however, the mechanisms must be held to even closer tolerances than in the transmitters in order that deviations be not apparent in the recording. Because of its greater inherent definition the photographic process requires the most precise mechanisms. Specific tolerances are indicated in Table II.

In particular, the lead screw must be very accurate in order that irregular feed between adjacent lines not result in light or dark streaks in the recording. Also the drum or equivalent must travel smoothly and without hunt within the stroke or from stroke to stroke. This infers complete lack of gear ripple and a drive motor which has a uniform

TABLE II.—OPERATING STANDARDS OF TYPICAL FACSIMILE SYSTEMS USED ON WIRE LINES IN THE UNITED STATES

	System *A*	System *B*	System *C*	System *D*	System *E*
Drum speed, r.p.m.	180	100	90	100	90, 45
Scanning lines per inch	100	200	96	100	96
Maximum picture size	7¾ × 11½ in.	5½ × 8½ in.	8 × 10 in.	7 × 9 in.	7 × 9 in.
Frequency of standard fork	60	160—fixed stations 2,400—portable	60	60	1,800
Phasing method	Automatic—from synchronizing dash sent at start of each picture				
Carrier frequency on line	5,000 cycles on trunks 2,500 cycles on local loops	2,400	1,800	1,800	1,800
Modulation transmitted	Double side band	Lower side band only	Double side band	Double side band	Double side band
Equalization	Amplitude equalizers along line and shaping discriminators at terminals	Permanent network equalized by phone company	No equalizer—use 1,000-cycle high-pass filter	No equalizer—use single *K* section 1,800-cycle high-pass filter	Nothing
Maximum picture modulation	1,250 cycles	1,200 cycles	800 cycles	800 cycles	Approx. 400 cycles
Permissible skew	⅛ in.	1/16 in.	¼ in.	1/16 in.	
Accuracy of synchronizing frequency	1 in 100,000	1 in 1,000,000 or better	1 in 500,000	1 in 100,000	
Required stability of line level	Uses pilot channel for automatic gain control	0.1 db	....	3 db	¼ db
Type line system	Own lines	Network leased	Long-distance toll calls		

instantaneous rate of rotation. These may usually be minimized reducing the mass and inertia of the drum system or by use of flexi coupling and flywheel action to filter out these quick-speed variatio

Many receiving machines are fitted with cams for the purpose of quasi- or full-automatic synchronizing or phasing. This is particular necessary in home recorders or in a large network receiving a pict simultaneously in several offices. Most of these utilize the inter between the end of one scanning line and the beginning of the next such a phasing signal. One variation is in the start-stop type of recei cating machine, in which the travel of the stylus is arrested at the end each complete stroke cycle and is then released for the next cycle receipt of the phasing signal.

Many appliances have been developed, such as sheet folders cutters, recordings on both sides of the record sheet, automatic dru loaders and discharge schemes, automatic hold circuits to permit rer if needed, etc. Some, notably in the international field, are equipp with gear shifts to permit meeting various conditions and standards other countries.

Machines have often been designed so they may readily be used either transmitters or recorders. Others have been designed for mouit service in conjunction with transmitters.

12. Receiving Circuits. Most of the facsimile systems, either w line or radio, record the signals in essentially the same general form

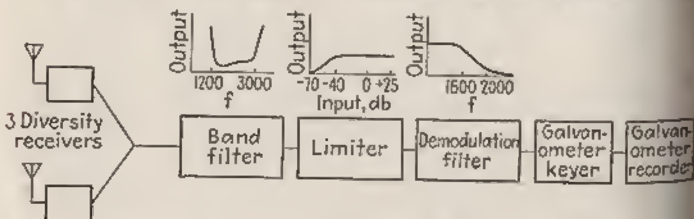


FIG. 18.—Block diagram of receiving system for subcarrier f-m method.

they are received. For these the received signal is amplified, rectifie and applied to the recorder either directly or through some type vacuum-tube keyer. Invariably the frequency band is limited to th just necessary to pass the pertinent signal components. This is usual accomplished by the use of filter networks; the process is carried to more or less high degree in the variously designed systems. Its purpo is primarily to increase the effective signal/noise ratio, but on long wi lines or extensive networks it also eliminates those lower frequency sign components whose phase would be sufficiently distorted to have a dele terious effect on the recorded copy.

Some few systems use special circuits at the recorder. One such the carbon recorder method when used for half tones. In order that th recorded densities of the copy have a satisfactorily linear relation the densities of the original subject, it is necessary that the amplitud characteristic of the amplifier be predistorted in a curve conjugate that of the recorder characteristic. Of course, this can, and is, often don at the transmitting end instead of at the receiver.

A second system which requires special "signal shaping" is the method for transmission by submarine cable. This is more fully treated in Art. 17.

A third such system is the subcarrier f-m method. The heart of this method lies in the special treatment accorded the signal at the recording location. In this case the signal is amplified to a usable value, say zero level, and then applied to a very rigid amplitude limiter (*e.g.*, one which shows no change in output level for a variation of the input on the order of 70 db). The output is then applied to a frequency-discriminating network (*e.g.*, cutoff slope of a low-pass filter). This reinserts an amplitude variation which is proportional only to the variation of frequency of the received audio subcarrier. Thereafter the signal is handled as an ordinary a.m.

SYNCHRONIZING

In order that the recording be properly built up to appear to be the equivalent of the original, it is necessary that the elemental areas be recorded in the same geometric relation with the others as are the elemental areas scanned at the transmitter. This demands that the transmitting and recording mechanisms operate in practically perfect synchronism and phase (or frame) with each other. It is now customary to drive the mechanisms by synchronous motors or by accurately speed-controlled d-c motors. The inherent accuracy of the frequency at transmitter and receiver is usually depended upon for equal driving speeds. The phasing is accomplished by manually or automatically causing the recorder to start a scanning line simultaneously with the start of a scanning line at the transmitter. This is done either at the start of a subject or periodically throughout the reception.

13. Drive Motors. For home recordings ordinary 60-cycle synchronous motors are often used. These are satisfactory if the transmitter and recorder are connected to the same power network or to networks that are interconnected or synchronized with each other. However, this is a considerable limitation, and the trend appears to be to transmit a control frequency to the recording station and to use that to control or drive the motor.

The schemes for driving or controlling the speed of the motors on the recorders are identical to those used on the transmitters, as mentioned in Art. 7. All of them require a source of standard frequency for controlling the driving circuits. The great majority of present systems use independent tuning forks at the transmitter and recorders, or else they connect directly to a-c power lines which are interlocked as to frequency. However, several designs have in the past—and there are signs of revival of the idea—actually transmitted a control tone derived from the frequency standard at the sending station.

14. Speed Controls. All commercial systems obtain their speed control from tuning-fork oscillators. The frequency may lie between 60 and 1,500 cycles.

Some take elaborate precautions to enclose the forks in heated and heat-insulated chambers, with the temperature variations held to 0.01°C. or less. The drive circuits are carefully engineered to minimize or eliminate variations in the drive due to supply voltage, tube emission, load, or similar variations. Recently it has been realized that changes in atmospheric pressure have a considerable effect on the frequency stability of forks.

Other systems utilize a fork which is compensated for temperature a therefore does not require careful temperature control. Such forks a made of a bimetallic layer structure in which the two metals have oppo temperature coefficients. By proportioning the two metals properly, t resultant for the fork can be made to within 0.05 p.p.m. per deg centigrade. However, its frequency is susceptible to variations in dr and atmospheric pressure.

15. Phasing Methods. In start-stop systems the recording gear is chosen so the scanning-line cycle is traversed faster than at the tra mitter, and the motion is arrested at the end of each cycle. The tra mitter then sends a release or start signal at the commencement of ea new scanning-line cycle, which releases the drive at the recorder. Th the recorder is in step with the transmitter at the commencement of ea line.

In other systems a phasing signal is sent at the start of each transm sion schedule, and this cooperates with a cam on the recorder to slow dow or speed up the drive until the two machines are accurately phase This circuit is then disabled, and the relative equality of the fork f quencies at the two stations is depended upon to retain the phasing. I still others a special signal is sent which releases a clutch at the record and thus starts it in phase with the transmitter. This is essentially th same procedure as is used in the start-stop system except that it is do but once rather than at the start of each line.

PROPAGATION

The communication medium to be utilized has a great bearing on th design of the facsimile system. There are essentially three mediums use at present, *viz.*,

1. Radio.
2. Submarine cable.
3. Landline telephone circuits.

Each has its special problems of propagation.

16. Radio Circuits. The radio circuit is affected by multipath trans missions resulting in both general and selective fading, in fading caus by the ionosphere variations of short-time, diurnal, seasonal, and oth phenomena. It is also affected by the signal/noise ratio that can b realized with the equipment available, as well as by interference fro natural and man-made sources.

Amplitude modulation of the r-f carrier cannot be received in a suff ciently stable manner for use in facsimile transmission, and, until recently the only practical system was to utilize a complete telegraphic on-and-o keying of the r-f carrier. The intelligence is conveyed as a "time modula tion," so that the amplitude of the signal could be rigidly limited at th receiver, thus eliminating such variations caused by fading. However even with this limiting, serious distortion occurred in the form of a variable and erratic elongation of the keyed signals due to delayed arrival over the various multipaths from the ionosphere. Although this is in a sense a phase distortion, it is not of the same type experienced on landlines and cannot be compensated by phase equalizers, as known for that purpose. In this method of keying there is no way of minimizing

noise or interference other than narrowing the audio band by means of filters as much as is possible, consistent with maintaining a sufficiently true wave shape of the signals for recording purposes.

The use of f.m. greatly minimizes these difficulties. So long as the received signal amplitude is greater than the peak noise, the noise has but a minor effect. If the signal is two or more times the peak noise, the full benefit of the "noise-improvement threshold" is realized and essentially no noise appears in the recording. The same is true of interfering signals. Further, the only effect of the multipath phenomena occurs at the edges of sudden changes in the picture tonal value, and this appears merely as a raggedness in the recording of such edges—no effect of multipath is to be seen in areas of constant or slowly changing tonal values.

Some work has been done with facsimile transmissions at high speed—240 r.p.m. and better—on u-h-f radio circuits. Here the difficulties with fading and multipath signal variations are not of importance, and available band width is ample. In this case the propagation is affected only by the constancy of the received signal and by the audio phase and amplitude characteristics of the radio and terminal office equipment.

17. Submarine Cable. The attenuation characteristics of even the loaded submarine telegraph cables precludes the useful transmission of frequencies much higher than 100 cycles. For facsimile transmissions the d-c variations in the phototube are d-c amplified and applied directly to the cable rather than as a modulation on a subcarrier as is done in other systems. Therefore the signals are subject to earth currents produced by magnetic variations and are greatly affected by magnetic storms. Correction must be applied to offset the "zero wander"; this can be done successfully for slow variations but becomes more difficult when the rate of these variations approaches the pertinent frequencies of the facsimile signals, as may happen. Earth currents of 9 or 10 volts varying at a rate of 10 cycles, or currents of 50 volts varying at a much slower rate, can be compensated. The transatlantic cable is comprised of two sections, each of which has an attenuation droop of 30 to 90 db at the higher frequencies. These sections must be individually equalized by "signal-shaping" networks. It is also essential carefully to correct a considerable phase distortion existing on the cables. Operations are effected successfully when the total noise and earth currents are 5 per cent or less of the signal swing from black to white.

18. Wire-line Telephone Circuits. Extensive networks utilizing wire lines are in operation in the United States and throughout Europe, as well as in other parts of the world. These all use existing telephone channels and systems and rely on transmitting an a-f subcarrier, amplitude modulated by the picture signals. Both double and single side-band methods are used. In the latter the upper side band is suppressed, the carrier set near the top of the telephone channel (say 2,400 cycles) and the lower side band, extending down to approximately 1,200 or 1,000 cycles, is transmitted.

For circuits of any considerable length it is important to equalize the lines, both for amplitude and for phase distortions. It is essential that the over-all phase characteristic for the pertinent frequency band be nearly linear. Relatively small deviations will delay certain frequencies with respect to others and may produce all sorts of weird effects in the recording.

The over-all gain must also be held within extremely close limits, a variation of 0.1 db can be perceived by the eye and variations great than 1 db cannot be tolerated in high-class commercial service.

Some systems are equipped with networks comprised of special circu which have been carefully equalized and adjusted and are used on for facsimile work. These may be extended by using ordinary telephon lines to interconnect them with other locations, usually on an emergen basis, for a sudden news event. Other systems utilize ordinary telepho toll facilities and take their chances on the quality and stability of t circuit. Some connect their equipment directly to the lines throu repeat coils, and others connect by inductance coils coupled to the ringi box and coil of the ordinary telephone subscriber's station. Most these networks are set up primarily for the handling of news pictures a are therefore designed for the utmost of flexibility so as to meet an emergency of news occurrences.

TAPE-FACSIMILE SYSTEM

Tape equipment is designed solely for message communication a opposed to picture or news matter. It produces a record on a narro tape, much as do the better known telegraph printers. The method recording used to date is that of a rapidly rotating spiral and an ax bar moved by a loud-speaker magnet in accordance with the sign. This is exactly simila to the action shown in Fig. 17. The recordi

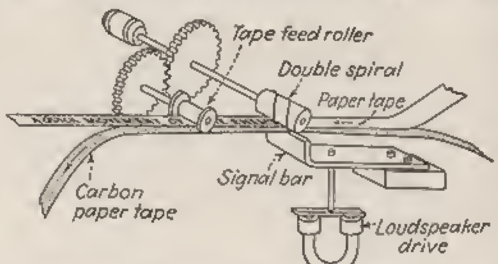


FIG. 19.—Elements of facsimile tape recorder with double spiral.

has been done either with a carbon paper tape, also similar to Fig. 17, by applying ink to the surface of the spiral through the medium of a f roller saturated with the ink. The scanning lines are crosswise of t tape and are made at a rate of about 60 per second. The tape is slow advanced lengthwise so the texture of the lines are about 60 to 100 pe inch.

Two distinct methods of transmission have been developed. In th United States much work has been done to develop a phototube scann along the general principles outlined in Art. 4. This method actual transmits a facsimile copy of written or printed messages placed on a tap at the scanner. It is being developed for mobile services, such as pol and aircraft.

The second method utilizes a special instrument which comprises a large number of cams, one for each character (figure, letter, or punctu

tion mark) to be sent. The cut of the cam is such that a contact operated thereby will send out telegraphic type mark and space signals, which, when recorded as above described, will form the shape of the desired character. Two designs of this instrument have been developed. In one the message is perforated in paper tape as though it were to be sent on a standard printing telegraph circuit. The permutations of the holes in this tape consecutively select and release the proper cams, as the tape is fed through the instrument. In the other type a typewriter keyboard is manually operated and the striking of a type key will release the proper cam. This type of scanning is used extensively in Europe, and the United States rights have been acquired by one of the large companies in this country.

The synchronizing problem is just as pertinent as in the other systems but is possibly slightly easier. This is because the scanning line is so short and the rate so high that the discontinuity between the end of one line and the commencement of the next readily provides a frequency component that may be used for automatic framing and synchronizing.

OPERATING STANDARDS

To date the only effective attempt toward standardization in the facsimile field has been in the adoption of International Standards under the aegis of the C.C.I.T. and the C.C.I.R. The former, in its Opinion 6SI, as amended at Warsaw in 1936, established rules for drum size, line advance, speed of operation, frequencies for synchronizing and carrier, setting up of circuits on the international telephone circuits, tariffs, refunds and rebates, etc. The C.C.I.R. is endeavoring to modify or apply these rules to the needs of the radiophoto service.

A typical proposal to the C.C.I.R. which covers the technical specifications, is cited herewith to indicate the trend:

1. Drum diameter—88.00 mm (3.464 in.).
2. Drum circumference—276.46 mm (10.88 in.).
3. Gripping (framing) loss—15.00 mm (0.59 in.).
4. Phasing loss—5.00 mm (0.196 in.).
5. Maximum skew or hunt—0.08°.
6. Drum length—310.00 mm (12.2 in.).
7. Picture size, maximum—250 × 290 mm (9.8 × 11.4 in.).
8. Drum speed—20, 60 r.p.m.
9. Line advance—4, 5 per mm (101.6, 127 L.P.I.)
10. Index of cooperation—352, 440.
11. Speed stability—0.001 per cent.
12. Screen frequencies—100, 150, 200 cycles.
13. Standard frequency—300 cycles, or multiples.

The International Index of Cooperation is defined by the formula

$$M = \frac{D}{P} = DF,$$

where D = diameter of the drum
P = pace of the scanning line or helix
F = fineness of scanning expressed in the number of lines per unit length of the drum axis.

If two machines have different dimensions but the same index, the picture sent between them will be enlarged or reduced but will not be distorted in its proportions.

APPLICATIONS OF FACSIMILE

Radio Circuits:

Short Wave—at relatively low speeds of 20 to 60 r.p.m. for long-distance transmission.

U.H.F.—at high speeds of 240 to 600 r.p.m. for point to point work.

Medium Wave—at medium speeds 75 to 120 r.p.m. for broadcasting to homes.

Marine Service—broadcast of weather maps, etc., to ships at 20 r.p.m.

Wireline:

Point to Point—for news dissemination or public service, 90 to 120 r.p.m.; for message pickup and delivery (customers' machines or "letter-box" machines) at 180 r.p.m.; for message service on trunk lines at 180 r.p.m.

Submarine Cable—at 20 r.p.m.

Photoengraving—used for preparation of printing plates, either black and white or four-color separation plates for color printing.

Military—for both Army and Navy use in handling maps and documents.

Tape Facsimile or "Hellschreiber"—used extensively throughout Continental Europe for news dissemination to agencies, by radio on 60 to 150 kc—also proposed for aircraft and police-car use.

References

Although a vast amount of work has been done in facsimile, it is fortunate that recent compilations have gathered the various references together so that the few citations given below will permit the reader to follow in detail the developments of facsimile in the radio field and will give him a working knowledge of the wire-line services.

Callahan, J. L.: A Narrative Bibliography of Radio Facsimile, in "Radio Facsimile," RCA Institutes Technical Press, 1938.

Jolley, L. B. W.: "Alternating Current Rectification" John Wiley & Sons, Inc., New York, 1928.

Mathes, R. E., and J. N. Whitaker: Radio Facsimile by Subcarrier Frequency Modulation, *RCA Rev.*, October, 1939, pp. 131–153.

"Radio Facsimile," RCA Institutes Technical Press, 1938. (A group of papers detailing various aspects of the art.)

Schroeter, Fritz: "Handbuch der Bildtelegraphie und des Fernsehens," Verlag Julius Springer, Berlin, 1932.

Wilson, J. C.: "Television Engineering," Sir Isaac Pitman & Sons, Ltd., London, 1937.

SECTION 21

RADIO BROADCASTING

By Carl G. Dietsch, B. Sc.[1]

1. Principal Elements of a Broadcasting System. All the equipment of a broadcasting system extending from the microphone to the radiating antenna of the radio transmitting station will be considered as part of the system. A general circuit layout of typical facilities of the kind used in the larger broadcasting centers for supplying a network of stations with program service is represented by the simplified diagram, Fig. 1. Equipment of a single studio is represented; that of other studios of the usual group would be similar and would be at the point marked on the program bus. Inasmuch as many programs, such as the broadcasting of special events, originate at remote points, in most cases a great distance from the studio, the layout of the facilities for remote pickups, sometimes termed "nemo" programs, has been included to illustrate the use of telephone lines as well as point-to-point radio-telephone communication to complete the circuits necessary.

A list of the essential elements of the system is as follows:

1. Microphones:
   a. Studio.
   b. Remote pickups.
2. Apparatus for controlling and conveying microphone output:
   a. Studio control booth:
      (1) Preliminary amplifier.
      (2) Microphone mixers.
      (3) Studio amplifier.
      (4) Volume control or faders.
      (5) Volume indicator.
      (6) Monitoring speaker.
   b. Remote pickups:
      (1) Preliminary amplifier.
      (2) Volume controls or faders.
      (3) Volume indicator.
      (4) Monitoring equipment.
      (5) Radio telephone or wire-line facilities for intercommunication.
3. Master control-room apparatus:
   a. Volume controls.
   b. Studio amplifiers.
   c. Relays and switching apparatus.
   d. Network channel amplifiers.
   e. Volume indicator.
   f. Monitoring facilities.
4. Telephone-line facilities to local radio transmitting stations and to distant radio transmitters connected to networks.

[1] Engineering Department, National Broadcasting Company.

5. Radio transmitter:
   a. Line amplifier or limiting amplifier.
   b. Volume controls.
   c. Volume indicator.
   d. Radio transmitter.
   e. Monitoring equipment:
      (1) Monitoring rectifier and speaker.
      (2) Modulation-percentage indicator.
      (3) Carrier-frequency monitor.
   f. Antenna.

2. Audio-frequency Range. Perfect reproduction of a sound transmitted through an electroacoustic system requires that the system pass all the audible frequencies of the sound in their relative intensities. Under these conditions of reproduction, the listener would be conveyed acoustically from his loud-speaker to a point near the sound source.

A correlated acoustic chart of the frequency range of various musical instruments within the orchestral range and the different voices which constitute the vocal range is shown in Fig. 2. The shaded keys are not included on a standard piano keyboard. The extreme organ range not shown on the chart is from 16 to 16,384 cycles physical pitch. The extreme frequency-transmission ranges necessary to produce perfect naturalness of speech and orchestral music are shown in Fig. 3. These ranges extend considerably above those of Fig. 2 because they include overtones and noise accompaniment additional to the fundamental tones. These curves were secured as a result of listening tests by a group of observers upon sounds transmitted through an electroacoustic system equipped with electrical filters by means of which frequencies above and below any desired cutoff could be suppressed. Extensive research[1] made during recent years indicates that for perfect reproduction of speech and music a frequency range between 30 to 15,000 cycles is desirable in order that the average ear may appreciate fully all the frequencies produced by the sound sources.

The curves shown in Fig. 4 are an indication of the relative qualities of reproduced orchestral music the frequency range of which was limited by electrical filters. It is apparent from these curves that, where a transmission system has a limited frequency range, such as that which exists in broadcasting technique, an acceptable reproduction of the sound sources may be secured within a band width of between 30 to 9,000 cycles.

The engineering and economic limitations of the frequency range used for broadcasting lie in restrictions of the use of the upper audio frequencies due largely to a limited band width of the modulation spectrum contained between the presently assigned carrier frequencies of 10-kc separation.[2]

An overlapping of the modulation frequencies of a "wanted" station by those of an "unwanted" station of 10,000-cycle separation restricts the range of frequencies to a broadcast listener usually considerably below that which is passed by the broadcasting system itself. The high quality of programs available from broadcasting facilities which have an over-all uniform frequency response from the microphone to the

[1] Snow, W. B., *Jour. Acoustical Soc. Amer.*, July, 1931, p. 61.
[2] Eckersley, P. P., Minimum Frequency Separation, *Proc. I.R.E.*, February, 1933, p. 195.

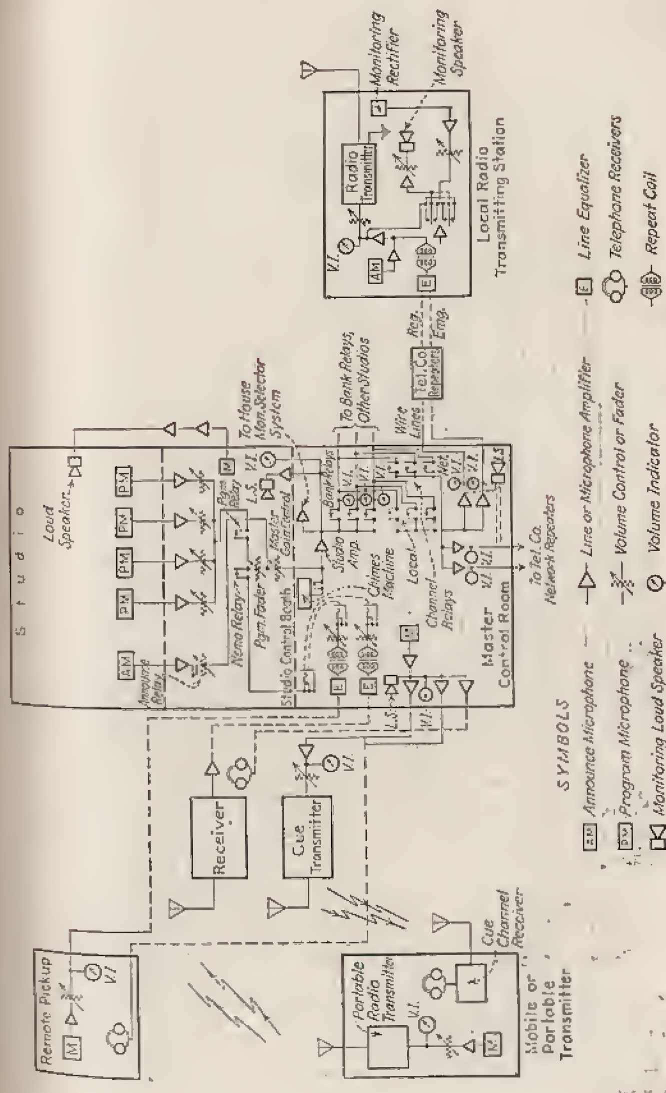


Fig. 1.—Schematic diagram of typical broadcast transmitting system.

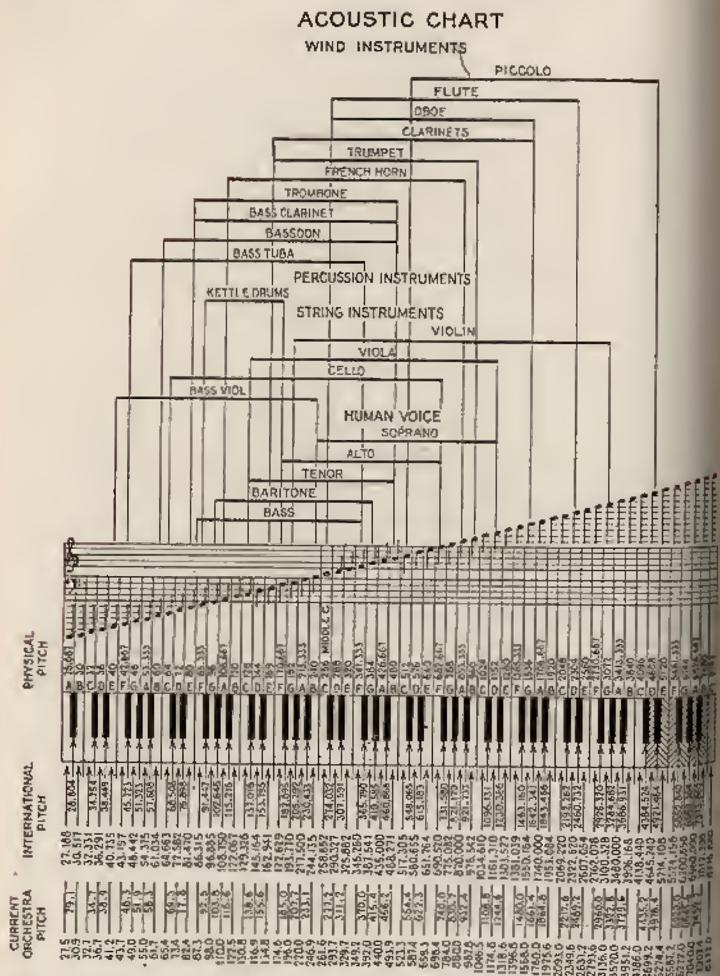


FIG. 2.—Correlated acoustic chart showing the scientific or philosophical scale generally used by physicists, the international equally tempered scale based on $A = 435$ complete vibrations per second. This scale was formerly used by musicians. The current orchestra or symphony scale based on $A = 440$ complete vibrations per second is at present generally used by musicians.

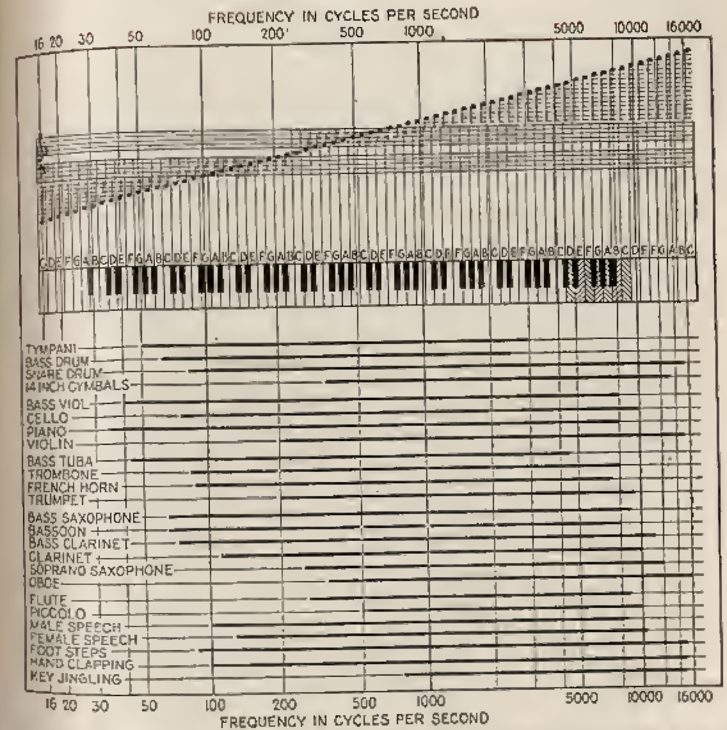


FIG. 3.—Frequency range required for the reproduction of musical instruments, voice, and noise without noticeable distortion.

TABLE I.—PEAK POWER OF MUSICAL INSTRUMENTS
(Fortissimo Playing)

Instrument	Peak Power, Watts
Heavy orchestra	70
Large bass drum	25
Pipe organ	13
Snare drum	12
Cymbals	10
Trombone	6
Piano	0.4
Trumpet	0.3
Bass saxophone	0.3
Bass tuba	0.2
Bass viol	0.16
Piccolo	0.08
Flute	0.06
Clarinet	0.05
French horn	0.05
Triangle	0.05

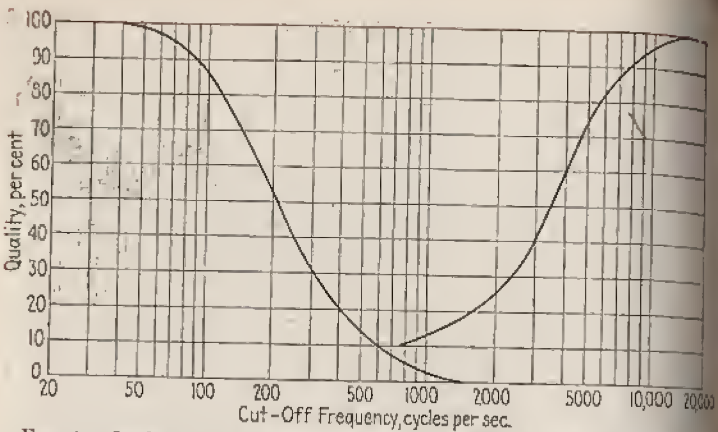


FIG. 4.—Quality of orchestra music as a function of cutoff frequency.

Note	Cycles per second	Organ pipe	Remarks
C_{10}	33,488		
C_9	16,744		Beyond limit of audibility for average person.
	15,000		Considered ideal upper limit for perfect transmission of speech and music.
	10,000		Considered as upper limit for high-quality transmission of speech and music.
	9,000		Considered as satisfactory upper limit for high-quality transmission of speech and music.
C_8	8,372	¾ in.	Highest note on fifteenth stop.
C_7	4,186		Highest note of pianoforte.
G_6	3,136		
E_6	2,637.2		Approximate resonant frequency of ear cavity.
	3,000		Considered as satisfactory upper limit for transmission of speech for ordinary communication.
C_6	2,093		
	2,000		Maximum sensitivity of human ear.
	1,500		Mean speech frequency from articulation standpoint.
A_4	880		
	800		Representative frequency of telephone currents.
E_4	659.3		
A_3	440		Orchestral tuning (see note below).
C_3	261.6		
	200		Considered as satisfactory lower limit for good-quality transmission of speech.
C_2	130.8		
	100		Considered as satisfactory lower limit of high-quality transmission of speech and music.
E_2	82.4		
C_1	65.4	8 ft.	Lowest note of cello.
B_0	61.7		
C_0	32.7	16 ft.	Lowest note of average church organ.
	30		Considered ideal lower limit for perfect transmission of speech and music.
A	27.5		Lowest note of pianoforte.
G	24.5		
C	16.35		Lowest audible sound. Longest pipe of largest organ.

	C	D	E	F	G	A	B	C
Notes of the "Gamut"	C	D	E	F	G	A	B	C
Vibration frequencies proportional to	1	9/8	5/4	4/3	3/2	5/3	15/8	
Intervals between successive notes	9/8	10/9	16/15	9/8	10/9	9/8	16/15	

NOTE: Nearest note is indicated. Scale A = 440 cycles per second based on middle C_1 (symphony pitch) = 261.6 cycles per second.

FIG. 5.—Frequencies to be transmitted on a high-quality system.

antenna within 2 db from 30 to 9,000 cycles and above cannot therefore be appreciated by the average listener because of limitations in the average broadcasting-receiver frequency response and restrictions in the present sound broadcasting band width which can be received free from cross talk and "monkey chatter."

3. Volume Range. Table 1 (page 773) gives the peak power of various musical instruments playing triple forte. A violin playing very softly has an output of about 4 mw, whereas that of a full orchestra has a peak value of 70 watts. The intensity range of the sound sources in this case is about 43 db. Owing to limitations in broadcasting circuits, background noise, and the modulation capabilities of the transmitter, this volume range must be in most cases compressed within the limits which can be handled by the wire lines and their associated equipment, as well as the transmitter where serious amplitude distortion results if modulation peaks, except those of extremely short time duration, exceed the modulation capabilities of the transmitter.

STANDARD REFERENCE LEVELS

4. VU and Decibels. The electrical signal intensity or level of an audio signal passing through each particular circuit of the broadcasting system, including the studio equipment, wire-line facilities, and the

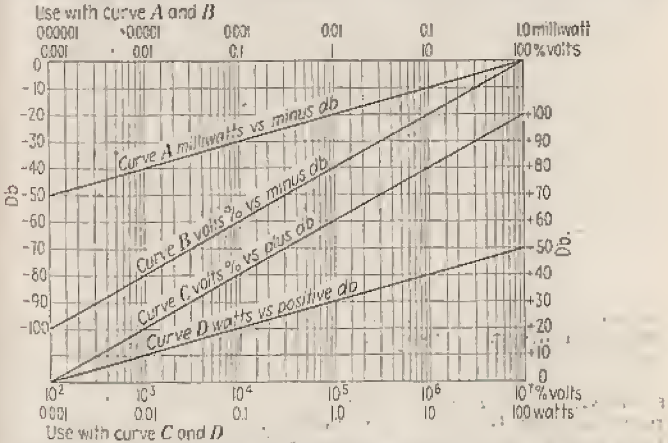


FIG. 6.—Relation between decibels and watts or per cent volts. Zero level = 1 milliwatt across 600 ohms.

broadcasting station, must at all times be carefully adjusted. The adjustment must be such that the transmitted program signals will remain within the limits which every part of the system can transmit without objectionable distortion due to overloading or from interference caused by noise, such as the interference produced by cross talk, induction, rectifier ripple, etc., inherent in equipment as well as associated wire lines. A convenient and consistently accurate method of measuring the amplitude of the signals is required, as well as a reference level common to the entire system. From this reference level, termed the *zero reference point* or zero VU, is based the amplitude of the program waves throughout the system. It also serves as a reference level from which the amplitude of interference or noise may be measured.

For broadcasting technique together with the interconnecting wire lines between studios and broadcasting stations there has been established a standard energy reference level of 1 mw. For the standard line impedance or pure resistance of 600 ohms at the terminals of a piece of apparatus in the system the zero reference level in VU would correspond to $\sqrt{0.6}$ r-m-s volts of 1,000-cycle sine-wave electrical energy as measured by a standard a-c voltmeter across the terminals.

Since program signals have wave shapes that are very complex and because peaks of these complex waves are liable to cause overloading there was developed and standardized a new standard volume indicator for the purpose of measuring program levels at all parts of a broadcasting system so that the correct signal level can be maintained without objectionable overloading. This instrument (see Volume Indicators) is calibrated to read VU on a logarithmic scale. It has electrical characteristics approximately equivalent to those of an r-m-s instrument. For signals having sinusoidal wave shape, the VU readings on this standard instrument should follow the decibel-voltage curve shown in Fig. 6. However, since the instrument is designed and used for measurement of complex program waves, the VU level of a particular program wave is as indicated by this standard volume indicator because of its particular characteristics. The term VU is therefore associated with the reading of this meter whereas the term decibel follows steady-state conditions and mathematical laws.

Simultaneous with the establishment of the new reference level of 1 mw the 0 db level of 12.5 mw was abandoned and the value for standard apparatus and telephone-line termination impedances for broadcasting was changed from the previous value of 500 ohms to the present standard 600 ohms.

AUDIO FACILITIES

5. Microphone Requirements. By means of the microphone, acoustic energy of sound waves produced for broadcasting purposes is converted into those of electric energy, the wave shape of one conforming to that of the other. The principal requirements of a microphone which will produce high-quality conversion are as follows: a relatively high sensitivity with respect to its inherent noise level, a uniform wave response over the frequency range desired, a substantially uniform frequency response over the angles included by its directivity characteristic, and mechanical and electrical ruggedness.

With some reservation, one may say that all forms of acoustoelectric transducers require the introduction of an obstacle into the path of the sound waves. To be effective, the active element of a microphone must either partake of, or otherwise influence, the motion of the air particles or it must respond in some way to the pressure variations on its surface. Some portion of the instrument such as the outside case, regarded as a rigid obstacle, must reflect some of the incident wave energy, whereas

[1] Chinn, H. A., D. K. Gannett, and R. M. Morris, A New Standard Volume Indicator and Reference Level, *Proc. I.R.E.*, January, 1940.

the element which responds to vibration from the sound waves must reradiate some of the energy exciting it. An instrument of high sensitivity and efficiency must, therefore, absorb a considerable proportion of the sound energy reaching it and convert it into electric energy. Faithful reproduction, therefore, is dependent upon the physical size and shape of a microphone. These features enter into the distortion of the true sound field, as well as the characteristics of the elements used to convert the sound into electric energy with a minimum of wave distortion.

Inasmuch as the quality of reproduction of speech and music is dependent upon the acoustic properties of the room[1] containing the sound sources and the placement of the microphone with respect to them, satisfactory results while using even the best instruments require a knowledge of the technique of microphone placement.

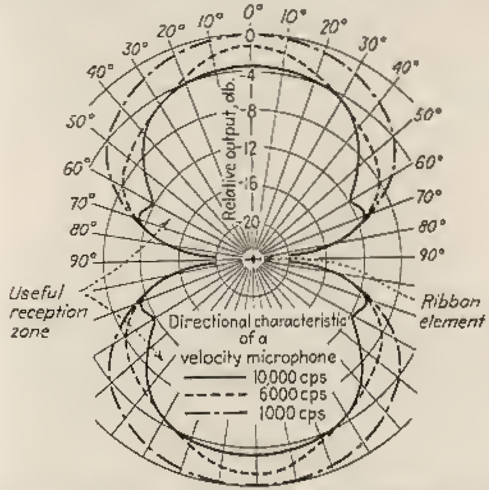


Fig. 7.—Directional characteristics of RCA velocity microphone.

6. The Velocity Microphone. This instrument gets its name from the movement of a metallic ribbon under the motion of air particles impinging upon it, thus setting up by electromagnetic induction an e.m.f. corresponding to the amplitude variations of an incident sound wave.

The commercial form of the RCA type 44BX[2] consists of a thin metallic ribbon suspended between the poles of powerful permanent magnets with the ribbon length perpendicular to, and its width in the plane of, the magnetic lines of force. It is moved from its position of equilibrium by the difference of pressure between its two sides. This pressure difference between the front and back of the ribbon is the same as that produced in a sound field between two points in space separated by this distance.

[1] Hanson, O. B., and R. M. Morris, Design and Construction of Broadcast Studios, *Proc. I.R.E.*, 19, January, 1931.

[2] Olson, H. F., *Jour. Soc. Mot. Pict. Engrs.*, 16, 695, 1931; *Jour. Acoustical Soc. Amer.*, 3, 56, 1931.

The pressure difference between the front and back of the ribbon is proportional to frequency. Since the acoustic impedance of the system is also proportional to frequency and the velocity in a mechanical system is the ratio of the pressure to the acoustic impedance, the velocity of the ribbon is independent of frequency.

With a ribbon constructed to have a natural period below the audible range, the frequency response is free from severe irregularities prominent in some pressure-operated types because of cavity and diaphragm resonance and from pressure-doubling effects produced at the higher frequencies. The ribbon is made light enough so that its motion will conform with the motion of air particles even at very high frequencies, with a result that the response of the velocity microphone is uniform over a wide range of frequencies.

The velocity-type microphone is markedly directional. With a plane-progressive wave the response in front and back of the instrument varies with the cosine of the angle between the direction of the sound wave and the normal to the ribbon. Since these directional properties are practically independent of frequency, they become useful in discriminating against undesired sounds and for obtaining a desired relation between the sounds from different sources and from reverberant sound in a studio.

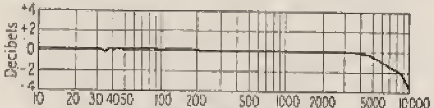


Fig. 8.—Velocity-microphone characteristics.

Its response[1] to reverberant or reflected sound is one-third that of a non-directive system, with the result that it can be used at a distance from a sound source of 1.7 times the distance of a non-directive type and still give the same results with respect to undesired reverberant sounds.

Because of the directional properties of the instrument, its sensitivity is at a maximum in directions in front and back perpendicular to the plane of the ribbon. With an input sound pressure of 1 dyne per square centimeter the unit will normally deliver open circuit across the 250-ohm tap an output level of −74 db compared to a zero level of 1 volt.

7. Moving-coil or Dynamic Microphone. This type of instrument (such as the Western Electric 630A) utilizes a light movable coil contained in a magnetic field to produce an e.m.f. which conforms with the sound waves impinging upon the dome-shaped diaphragm.

The assembly is composed of a coil of fine aluminum ribbon edgewise wound and attached rigidly to a duralumin diaphragm of low mechanical stiffness which supports the coil in a radial magnetic field of a permanent magnet made from high-grade magnet steel. The diaphragm has a rigid dome-shaped center and a tangentially corrugated annulus. It has a high area/stiffness ratio. The diaphragm is cemented to a raised annulus on the outer pole piece. The outer and inner pole pieces are of soft iron and are welded directly to the magnet. The diaphragm is damped by an acoustic resistance which is supported below the coil by a brass ring, which in turn is held in place by rubber gaskets.

[1] Olson, H. F., *Jour. Soc. Mot. Pict. Engrs.*, **16**, 695, 1931; *Jour. Acoustical Soc. Amer.*, **3**, 56, 1931.

When the diaphragm vibrates in response to the sound waves impinging upon its surface, the coil vibrates in a like manner and cuts the magnetic

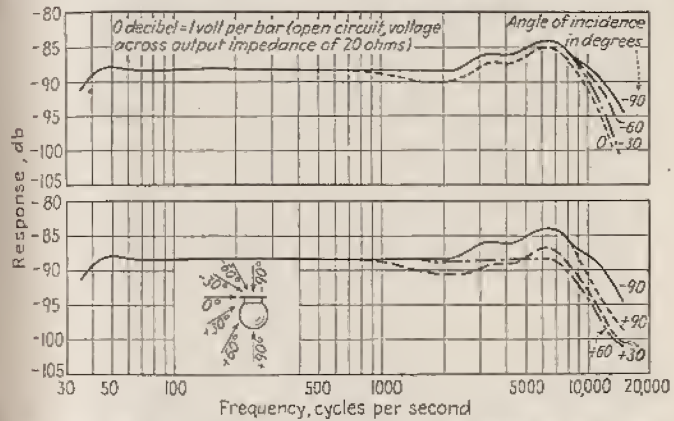


Fig. 9.—Field calibration of dynamic microphone showing effect of angle of incidence of the sound wave.

lines of force. The e.m.f. generated in the coil is substantially proportional to the sound vibrations which cause the diaphragm movement.

The spherical shape of the microphone housing and its size are such that the housing fits closely over the diaphragm and thus produces very little more diffractive effect than the diaphragm itself. To prevent resonance within the spherical case an acoustic resistance baffle is provided to divide the space in two parts. A tube with its outlet at the back of the housing serves the double purpose of equalizing the inside and atmospheric pressures and of increasing the response of the instrument at low frequencies.

This microphone was designed to provide a uniform frequency response in all directions and has been termed a *non-directional microphone*.[1] For this reason the small spherical shape was selected as well as the method of mounting the diaphragm in a horizontal plane. A protective grid is provided over the diaphragm to control the resonance of the cavity in front of the diaphragm. This grid is most useful in the improvement of the frequency response of the instrument at frequencies from 8,000 to 15,000 cycles.

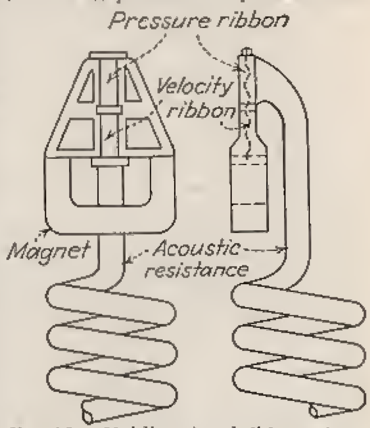


Fig. 10.—Unidirectional ribbon microphone elements.

Wave-response calibration curves of this type of instrument indicate that the frequency characteristics are influenced to some extent at the higher

[1] Marshall, R. N., and F. F. Romanow, *Bell System Tech. Jour.*, July, 1936, p. 405.

frequencies by the angle of incidence from which the sound waves approach the diaphragm. Since the diaphragm is mounted horizontal, the instrument is entirely non-directional with respect to the vertical axis.

In spite of the small physical size necessary to provide the non-directional characteristics, the sensitivity is about −88 db where 0 db is equivalent to 1 volt per bar[1] (open-circuit) voltage across the microphone output impedance of approximately 20 ohms.

The non-directional characteristics of this microphone make it useful as a pickup for large orchestras and choruses where in most cases the sound arrives at the microphone from all directions. Unless the microphone response is uniform in all directions, there is a form of distortion due to discrimination against certain frequencies with directivity.

8. The Unidirectional Ribbon Microphone. In certain forms of studio technique it is desirable to eliminate the pickup of unwanted sound in the rear of the microphone, such as audience noise, room echo, etc. Here the

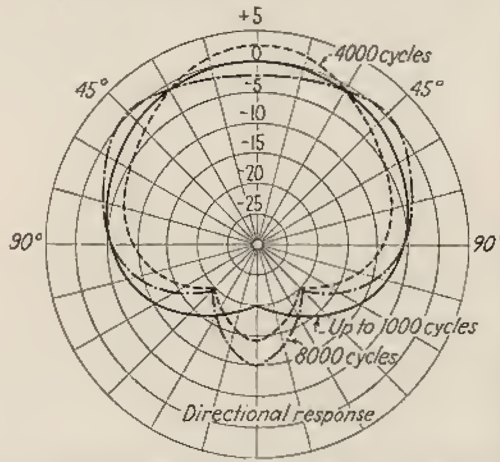


FIG. 11.—RCA 77B unidirectional microphone characteristics.

unidirectional microphone is very useful. The unidirectional instrument utilizes a light corrugated ribbon suspended in a magnetic field in somewhat the same manner as the bidirectional velocity microphone, except that the ribbon is divided into two individual sections, one of which is pressure operated and the other velocity operated.

The field response of the pressure-operated section is very nearly uniform in all directions and may be expressed as $E = E_0$, whereas the response of the velocity section is bidirectional and is equivalent to $E = E_0 \cos \theta$. Since the sensitivity of the non-directional pressure section is made to equal the greatest sensitivity of the bidirectional or velocity section, the combined polar field response characteristic of the two is equivalent to $E = E_0(1 + \cos \theta)$. In three-dimensional space this is very nearly equivalent to a cardioid of revolution. The point of maximum sensitivity is directly to the front of the instrument, while directly to the rear of it the sensitivity approaches zero.

[1] A bar is 1 dyne per square centimeter.

A uniform frequency response in all directions for the pressure-operated ribbon section is approached by allowing the operating face to be freely accessible to the atmosphere while the other side is terminated in an acoustic impedance very nearly equivalent to that of a very long pipe. Since a long pipe is too cumbersome for practical purposes, a short pipe of correct cross section, provided in coiled form, and loaded throughout its length with absorbing material, such as tufts of felt, exhibits a suitable acoustic resistance over a frequency range covering all but the lowest frequencies.

The operating properties of the velocity-actuated ribbon section are quite the same as were described previously for the bidirectional velocity microphone. The ribbon is driven from its equilibrium position by a difference of pressure between the two sides; the pressure difference being due to the difference in phase between the two sides. The vibration of the ribbon caused by the sound waves impinging upon it causes an induced e.m.f. to be generated in the ribbon. The directional characteristics of the ribbon section are practically independent of frequency.

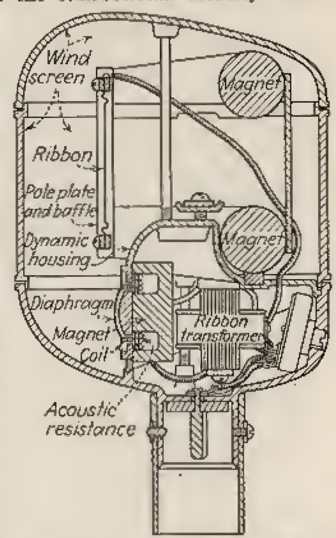


FIG. 12.—Simplified cross-sectional view of the cardiod directional microphone.

The RCA 77-B unidirectional microphone has an open-circuit output level of approximately −81 db based on 1-volt zero reference level for a sound pressure of 1 bar at 250 ohms output impedance.

The Western Electric 639-AA cardioid directional microphone[1] utilizes a ribbon element of special design in combination with a compact pressure type non-directional element to secure a field response having a directional characteristic similar to a cardioid.

The pressure element is of the dynamic type having a dome-shaped diaphragm and constructed in some respects similar to the Western Electric 630A microphone previously described. Commercial instruments of this general type have switches which enable the directional characteristics to be changed at will.

9. Crystal Microphone. This microphone utilizes the piezoelectric phenomenon produced in plates cut from piezoactive crystals. Thin plates cut from Rochelle-salt crystals are used almost entirely for the elements of crystal microphones. In comparison to other crystalline piezoelectric materials, such as quartz, Rochelle salt exhibits greater sensitivity for this purpose and it responds quite readily to mechanical vibrations.

Crystal microphones may be classified under two individual groups: (1) those utilizing multiple sound cells in free space and (2) those utilizing bimorph crystal elements each excited mechanically by an associated diaphragm. In the first of these types utilizing the Brush Development Company assembly, termed the *sound cell*, the elements are plates having dimensions ⅜ by ⅜ by 0.30 in. cut from Rochelle-salt crystals along axes in such a manner that their inherent characteristics tend to cause elongation or contraction when they are subject to an electric field provided by foil electrodes.

[1] MARSHALL, R. N., A Cardioid Microphone, *Bell Lab. Rec.*, July 1939.

By cementing together two such piezoactive plates which have tenden- to act in opposition to each other when a voltage is applied, an assembly produced with a motion analogous to the mechanical motion of bendi a bimetallic thermostatic strip acted upon by variation of temperatu The assembly consists of two plate combinations mentioned above, separa by an air space and held in position by a suitable mounting.

The cell is covered over with a membrane which serves as a pressure and to protect the crystals from the outside atmosphere. When the cell placed in a sound field, pressure acting normal to the outer surfaces of t plates tends to cause bending, with a result that an e.m.f. is genera between the foil electrodes. The two plate combinations are connected parallel. The wave form of this e.m.f. conforms with that of sound wav Because of the small physical dimensions of the plates the frequency mechanical resonance of the system is rather high, with the result th frequency response is quite uniform over a wide frequency range. So models are quite uniformly sensitive up to 15,000 cycles.

Commercial models contain series and series-parallel groups of these soun cells ranging from 2 to as many as 24. The sensitivity of a single sound ce is approximately −90 db, while a multicell microphone has a sensitivity as great as −68 db.

The output impedance (which is purely capacitative) of these instrumen is quite high. This sometimes requires them to be operated directly into th grid of an amplifier tube having a grid leak of about 5 megohms. The smal physical dimensions of a single cell make it practically non-directive. Thi property is also characteristic of multicell units. Figure 13 shows a resistance- capacity-coupled amplifier suitable for use with such a microphone.

The diaphragm type of crystal microphone, such as the Brush Model AP utilizes a hermetically sealed bimorph crystal supported at three points within the microphone housing. Projecting to the center of the specially treated fiber diaphragm is a small drive pin. This engages the remaining corner of the bimorph crystal. Inasmuch as the bimorph crystal is highly sensitive in converting fluctuating mechanical stresses, such as those caused by bending, into corresponding electrical fluctuations, the fluctuations in pressure created by the sound waves impinging upon the microphone diaphragm result in corresponding voltage fluctuations to be produced across the terminal ends of the bimorph crystal plates.

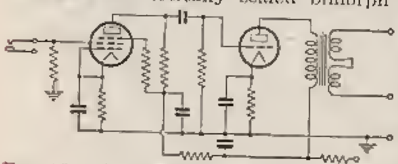


Fig. 13.—Amplifier for use with crystal microphone.

Similar to most pressure-operated microphones, this instrument is semi-directive. However, the smallness of the instrument assists in securing rather uniform frequency response with direction. By placing the diaphrag facing the ceiling of the room, the instrument is essentially non-directional in a plane through the diaphragm and parallel with the ceiling.

The output impedance of the crystal element is approximately 250,000 ohms. This permits long cables to be employed when either high or low impedance connections are used. By means of a high-quality matching transformer this microphone can be satisfactorily operated into circuits exhibiting 50, 200, or 500 ohms impedance. The frequency response of this instrument is substantially flat from 100 to 5,000 cycles. It has a variable control to allow manual adjustment of frequency-response characteristic while in operation. It has an output level of approximately −48 db based on a zero reference level of 1 volt per dyne per square centimeter.

The crystal microphone shunted across the input resistor of the amplifier tube attenuates the higher frequencies of background noise such as those caused by thermal agitation. The properties of the crystals themselves are such that they are liable to damage such as a change of frequency character-

istics and output if the instrument is subjected to temperatures in excess of 120°F. to 125°F. particularly for periods of several hours.

10. Condenser Microphones. The condenser microphone utilizes the principle of mechanical variation of thickness of the air dielectric of a charged electrostatic capacity as a medium to change acoustic energy into electrical energy of corresponding wave shapes. One form of this microphone consists essentially of an electric condenser formed by a thin, tightly stretched duralumin diaphragm spaced approximately 0.002 in. and insulated from a flat brass disk called the *back plate*.

A polarizing potential difference is applied between the condenser electrodes formed by the diaphragm and the back plate. The varying pressure upon the very thin diaphragm by the sound waves causes the electrostatic capacity of the condenser to vary by an amount in the order of 0.01 per cent of its normal value of 200 $\mu\mu$f.

The microphone has an aluminum alloy diaphragm 0.001 in. in thickness. The edges are clamped between threaded rings, the requisite stiffness being obtained by advancing the stretching ring until the desired resonant frequency, usually about 5,000 cps, is obtained. The space between the diaphragm and the back plate is hermetically sealed to prevent dust and moisture from entering and resulting in noise. The thin rubber auxiliary diaphragm, together with a small air-vent hole in the center of the back plate, is provided as an equalizing system for changes in atmospheric pressure.

On account of its inherent high-impedance characteristics, it is usual to incorporate an amplifier in the microphone housing to reduce to a minimum the length of the lead and the corresponding shunting capacity between microphone and associated amplifier grid. Sometimes a compact amplifier is placed on the floor alongside the microphone, the two being connected with low-capacity cable. A d-c polarizing voltage in excess of 180 volts has been used, but this should never exceed 500 volts.

Developments upon the early Wente[1] models by using duralumin as a substitute for steel as diaphragm material brought the sensitivity of modern instruments to about ten times that of early models.

Since this is of a pressure-operated type, there are inherent irregularities in its characteristics from acoustic and mechanical phenomenon. The microphone diaphragm is subject to certain resonance frequencies as well as the cavity. These tend to disturb the smoothness of the response characteristic. The sound waves striking and being reflected from the flat surface of the diaphragm cause pressure doubling especially at high frequencies.

Below 500 cycles this instrument is practically non-directive whereas at frequencies above 2,000 cycles the directivity is very noticeable. This directivity has a tendency to discriminate against h-f noise and reverberation, and, under certain conditions where the studio does not accentuate the low frequencies, it has an advantage since the human ear responds more easily to background noise of higher frequencies than to lower frequencies.

The sensitivity of the condenser microphone on the basis of an input sound pressure of 1 bar is approximately −60 db below 1 volt as measured at the output of the preamplifier.

The Western Electric 640A miniature condenser microphone unit[2,3] contains a diaphragm a fraction of an inch in diameter. The condenser unit is mounted in one end of a tapered shell housing, of dimensions approximately 2¼ in. in diameter and 7 in. long, which also contains the preamplifier. The weight of this microphone and preamplifier unit is 1¼ lb.

The output level of the complete instrument is −61 db below zero level of 1 volt per bar open circuit at the preamplifier output impedance of 50 ohms.

[1] Wente, E. C., *Phys. Rev.*, 19, 498, 1922.
[2] Harrison, H. C., and P. B. Flanders, An Efficient Miniature Condenser Microphone System, *Bell System Tech. Jour.*, July, 1932, p. 451.
[3] Hopper, F. L., *Jour. Soc. Mot. Pict. Engrs.*, September, 1939, p. 278.

Published field-response curves indicate a rather uniform frequency-respon characteristic from 40 to 10,000 cycles with some dropping off of the high frequencies as the angle is increased from an axis line normal to the diaphrag

11. Carbon Microphones. These devices use the variation resistanc of carbon granules to produce electric waves from sound waves. A typical example of a "double-button" carbon microphone is shown in Fig. 14. The diaphragm of this microphone is made from duralumi 0.0017 in. in thickness and is clamped securely around its outer edg Stretching of the diaphragm to give the desired resonant frequency

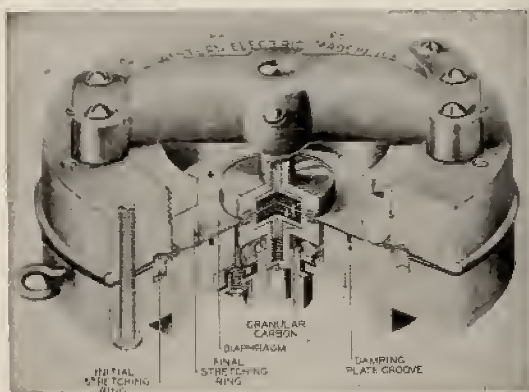


Fig. 14.—Carbon microphone.

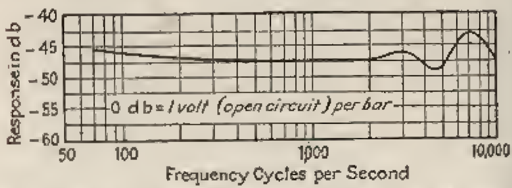


Fig. 15.—Response of air-damped duralumin diaphragm.

usually about 5,700 cycles, is done in two steps by means of two stret ing rings. To ensure uniformly low contact resistance, the portions the diaphragm which are in contact with the granular carbon are cover with a thin film of gold deposited by cathode sputtering. The carb granules will pass through a screen having 60 meshes per inch but will retained on a screen having 80 meshes per inch. Each button conta about 0.06 cc of carbon corresponding to about 3,000 granules.

The use of an air-damped stretched duralumin diaphragm has resul in uniform response over a wide range of frequencies.

The operation of a carbon microphone may be affected by cohering (som times called *caking*) of the granules. Severe cohering causes a large re duction in resistance and sensitivity which persists for an extended per unless the instrument is tapped so as to agitate mechanically the granule

One of the common causes of cohering is breaking the circuit when current is flowing through the microphone. Experience has shown that the use of a simple filter consisting of two 0.02 μf condensers and three coupled coils, each having a self-inductance of 0.0014 henry, will effectively protect the microphone button without introducing an appreciable transmission loss; a potentiometer switch also serves to prevent caking.

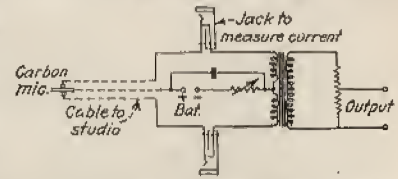


Fig. 16.—Carbon microphone connections.

The quality of transmission obtained with a double-button carbon microphone compares favorably with that secured with a condenser microphone; the carbon microphone has the disadvantage, however, of a high noise level or "microphone hiss." Figure 16 shows the manner in which the carbon microphone is connected to its associated amplifier. The current through each button is usually in the neighborhood of 10 to 20 ma.

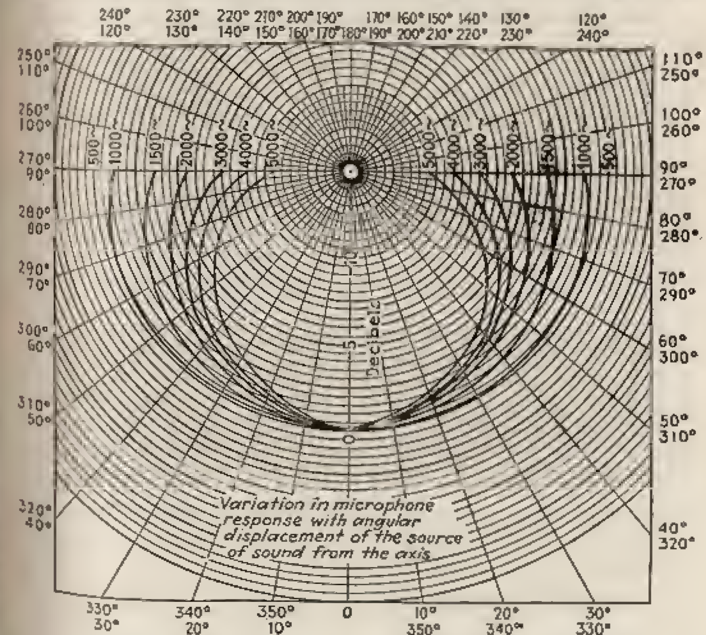


Fig. 17.—Directional characteristic of carbon microphone.

The sensitivity of the carbon microphone is somewhat higher than the other types. The average sensitivity is about −40 db.

Wave-response curves[1] for a carbon microphone show that response at normal incidence is quite uniform from 60 to 1,000 cycles. Above 1,000 cycles

[1] Ballantine, Stuart, High-quality Broadcasting, *Proc. I.R.E.*, **22**, 576, May, 1934.

it increases rapidly, becoming about 15 db higher at 2,500 cycles than at 1,0 cycles. This increase extends rather uniformly from 2,500 to 6,000 cycl where there is a marked falling off.

12. Parabolic Reflector Microphone. The use of a large conca reflecting surface mounted behind a microphone has been found to g the instrument pronounced directional characteristics in the recepti of sound waves. The system gets its name from the shape of the refle ing surface, a cross section of which contains a section of a parabol By virtue of the microphone placement at the focus of the parabola revolution or hollow paraboloid section, the sound waves striking t reflecting surface are concentrated upon that microphone diaphrag facing the inside of the paraboloid resulting in increased sensitivity the instrument in line with the axis inside of the paraboloid.

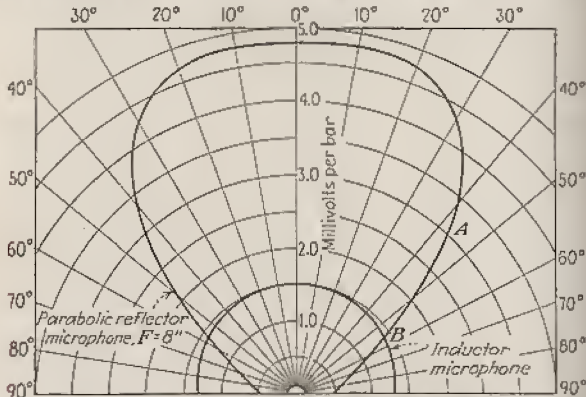


FIG. 18.—Comparative axial response at 1,000 cps in millivolts per bar. parabolic reflector; B, inductor microphone.

The use of the reflector, therefore, makes possible the placement of t instrument sufficiently far from the sound source so that it is practical equidistant from all the instruments or voices, with a result that t problem of securing proper balance and volume control is simplifi The directional characteristic makes it possible to swing the microph and its reflector as one would a searchlight and in this manner follow t action on the stage of an auditorium or on the field of a sporting even There is an increase in sensitivity along the line of axis of about 4 to due to the use of the parabolic reflector.

Since the reflector increases the sensitivity and makes it possible to loca the microphone at a greater distance from the source of sound, it is desira that the output of the microphone should fall off rapidly if the sound ori nates at a point displaced more than 30 deg. from the axis of the instrume if this characteristic is obtained, reverberation and reflections in the stu or auditorium will have very little effect.

The h-f response may be increased by as much as 15 db over the respo at low frequencies by varying the position of the microphone in the reflect However, in focusing the microphone, care must be taken to select the mo

useful frequency range, because at certain points of focus there is a tendency for sharp irregularities in the l-f response due to cancellation between the directly received and reflected sound from the paraboloid reflector. In certain instances where the h-f absorption is considerable, the ability to accentuate the highs by refocusing proves very helpful.

Another distinct advantage of the directional microphone is its ability to disregard to quite an extent the acoustics of the room as it responds almost entirely to the sounds upon which it is directly focused. In some cases another microphone without a reflector has been used with the parabolic microphone so that it may be faded in at certain times to make the reproduction sound more realistic. The parabolic microphone has been used to pick up sound from a certain section of a large crowd or audience of a sports event or to pick up the voice of a single individual at a time in an audience.

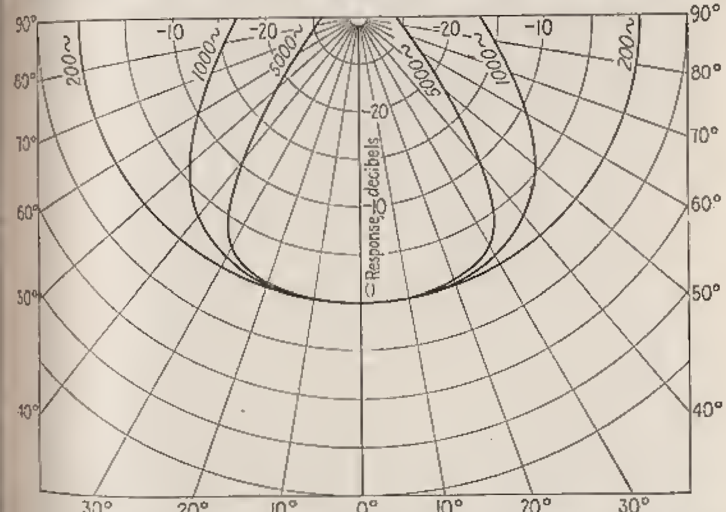


FIG. 19.—Axial frequency response of parabolic microphone with a focal length of 8 in.

13. Microphone Calibration and Testing. The sensitivity of a particular microphone is generally expressed as the open-circuit output voltage generated at the microphone terminals for a unit sound pressure against its active element. The intensity of the sound waves impinging upon the active element may be evaluated as a pressure or force. This is usually expressed in dynes per square centimeter or bars, where one bar of sound pressure is equivalent to one dyne per square centimeter of the surface area.

The actual voltage generated by the microphone being very minute (a very small fraction of 1 volt for a sound pressure of one bar against its element), the sensitivity may be expressed in minus decibels below a standard reference level usually taken as 1 volt. When it is desired to correlate this value with the amplifier gain one would have to assume the microphone to be loaded with a matching impedance. This would result in dropping the output voltage to one-half of the open circuit value or a corresponding 6 db decrease in output.

The sound pressure at a particular point where a standard microphone is set up is generally measured by the Rayleigh disk method. The instrument consisting of a light circular mirror suspended by a fine quartz fiber at an angle of 45 deg. to the axis of the tube through which the sound waves pass. The torque produced on the disk mirror by the sound field is measured by the deflected beam of light focused upon it. For small angles of deflection, the rotation of the disk is proportional to the sound intensity in the tube and consequently to the intensity of the undisturbed field. The actual value of torque may be determined by a torsion head which has a tendency to return the mirror back to its original position.

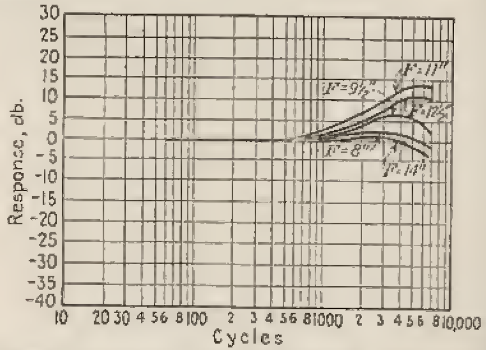


FIG. 20.—Frequency response in axis of parabolic reflector microphone at various focal lengths.

Where a sound chamber having suitable acoustic properties to prevent reverberation, at the lower frequencies especially, is not available, the response calibrations are made in open air in a quiet atmosphere. From a standard microphone calibrated in this manner, other instruments may be compared to it for characteristics.

In determining the response characteristic of a diaphragm-type instrument such as a condenser microphone, use has frequently been made of the thermophone method, the thermophone consisting of two strips of gold foil mounted on a plate and fitted into the recess in the front of the microphone, the recess being entirely enclosed and filled with hydrogen. A d.c. upon which is superimposed an a.c. is passed through the foil and causes fluctuations in the temperature of the foil and the gas immediately surrounding it. These fluctuations in temperature cause changes in the pressure on the microphone diaphragm, and the magnitude of the pressure developed on the diaphragm can be computed from the constants of the system. Thermophone calibration is often referred to as a pressure calibration, since it depends entirely upon the actual pressure developed on the diaphragm and hence does not take into account any effects which may occur when the microphone is used for actual pickup purposes. The response obtained by placing the instrument in a sound field of constant pressure is termed a *field calibration*.

The effect of the diffusion of the sound field and the tendency for most acoustic materials to be more absorbent at high frequencies appear to cause the microphone actually to respond more closely to the field calibration rather than to the pressure calibration.

Previous to the use of any microphone in an actual broadcast or rehearsal, it is carefully tested by speaking into it and having another trained individual listen to the quality of the sound reproduced through a high-fidelity amplifying and loud-speaker system. As compared to the results secured from a standard microphone of known high quality, the condition of the microphone under test can be determined.

STUDIO TECHNIQUE AND MICROPHONE PLACEMENT

14. Studio Problems. A problem of vital concern to a broadcasting system is that of providing favorable acoustic conditions within its studio or auditorium facilities in order that the effects of reverberant sound from the walls of the enclosures may be kept within desirable proportions in comparison to the sound reaching the microphones directly from the source. Of even greater concern are the problems involving correct placement of microphones with respect to the sound sources within the enclosures, to assure faithful reproduction of each voice or musical instrument, their significant overtones, and a pleasant blending of the groups of voices or instruments.

It is, therefore, by virtue of the selection of a microphone which will faithfully transmit all the actual sounds that occur within its range as well as the correct placement of it within a studio or auditorium having suitable acoustic characteristics that high-quality programs can be produced. Under optimum conditions of reproduction a broadcast listener would hear the same acoustic naturalness of the program from his loud-speaker as he would if he were to be transported to a favorable spot in the studio or auditorium where the sounds originating therein would afford a sensation most pleasing for him to hear.

The major considerations involved in proper studio design such as sound proofing, ventilation, optimum dimensions, and suitable acoustical treatment of the walls have been given.[1] At present we shall be concerned only with the problems of microphone placement, assuming that favorable studio and auditorium conditions exist. Normally, these considerations would be as follows: adequate soundproofing that would prevent undesired extraneous noises from entering a given enclosure, and suitable acoustical treatment of the walls and floor to provide equal absorption over a wide frequency range and give the enclosure in itself a uniform frequency characteristic. It is of considerable importance that the frequency characteristic of the studio or enclosure be considered for high-quality transmission because this characteristic is actually superimposed upon that of the microphone under conditions where the reverberant sound received by the microphone is appreciable as compared with that received directly from the source.

15. Single versus Multiple Microphone Usage. During the first years of broadcasting, it was a usual procedure to use more than one microphone to pick up a program, especially under conditions where the broad-

[1] HANSON, O. B., and R. M. MORRIS, Design and Construction of Broadcast Studios, *Proc. I.R.E.*, 19, January, 1931; SIVIAN, L. J., *Bell System Tech. Jour.*, 10, 108, 1931; MORRIS, R. M., and G. M. NIXON, Broadcast Studio Design, *RCA Rev.*, October, 1936.

casting group was rather large. This was necessary on account of rather low microphone sensitivity and the inherently high noise level of the carbon microphones used during that period requiring a placement of those instruments sufficiently close to the sound sources to overcome the inherent background noise of these carbon types. The combination of more than one microphone for making a pickup had a disadvantage in that the outputs from the several microphones used were not in proper phase relation with respect to the sound sources. This resulted in considerable distortion when the microphone outputs were combined and fed into a common amplifier.

Improvements in microphones to secure higher sensitivity as compared to inherent instrument noise level has resulted in the use of only one microphone at a time. The microphone is located at a sufficient distance from the sound sources so that more than one microphone is not necessary to obtain a good acoustic balance from a group. The practice of using more than one microphone at a time has, therefore, been discouraged whenever possible because of the phase distortion of the sound field resulting.

16. Microphone Placement. The carbon microphone, has been practically abandoned for use in broadcasting pickup work. The directive characteristics of the carbon and condenser types at the higher frequencies make necessary the placement of the broadcasting group within an area contained within an angle of 30 deg. either side of the microphone axis.

The frequency characteristics of any diaphragm type of microphone are dependent upon the relative positions of the microphone and the source of sound. When the sounds approach at right angles to the plane of the microphone diaphragm, a uniform response over the desired range may be obtained. But, if the sounds approach from any other point, it will be found in general that the response will fall off with frequency. This characteristic is illustrated by Fig. 17, which indicates how response varies with the angular displacement of the sound source from the microphone axis. It will be noted that there is a high loss at the higher frequencies for high angular displacements. Since the majority of musical instruments depend for their quality or timbre upon the presence of overtones, it is obvious that, if these overtones are discriminated against, the quality will be changed materially. If, in considering the loss in the higher frequencies with angular displacement, we apply the limitation that the loss at 5,000 cycles shall not be more than 2 db, then Fig. 17 indicates that, in using a single microphone of the diaphragm type, all the musical instruments of a group should be kept within an angle of 30 deg. either side of the microphone axis.

An individual source of sound such as a speaker, announcer, or musical instrument should not be placed closer to the microphone than 1 ft. Greater distances are determined by the volume range of the voice or instrument and the relative volume desired with respect to the accompanying instruments.

One must consider that in different selections and different arrangements of the same selection the relative importance of the particular instruments may be changed considerably. Where desired prominence cannot be given to a particular group at a certain time using a single microphone, it may be necessary to fade-in another located near the group to be emphasized. A number of microphones can in this way be

used, each for the particular interval desired, to emphasize the particular musical instrument or instrumental group, the soloist, or the announcer. The control of individual microphone circuits for this arrangement is performed in the control booth by suitable mixing and switching devices. Extensive rehearsing accompanied by listening tests at a remote point are generally required to secure the desired balance for a particular setup previous to an actual broadcast.

In general, the volume range of certain instruments adjacent to one another permits their alliance into natural groups, each instrument of a group being placed approximately equidistant from the microphone. One group may contain violins, violas, and cellos; a second group, the piano, harp, flutes, and clarinets; a third group, the oboes, bassoons, and French horns; a fourth group, the string bass, tuba, timpani, and traps; a fifth group, the trombones and trumpets. In dance orchestras the guitar is usually placed in the first group, the saxophone in the third group, and the banjo with the fourth group.

There are many factors involved in securing the proper placement of vocal sound sources or musical instruments before a microphone particularly before a pressure or diaphragm type. While certain rules have been set up, they may serve only as a guide. Most satisfactory results are obtained by a combined study of the instruments as well as an actual setup of them before a microphone in a given enclosure. The results of actual listening tests by means of a high-fidelity speaker and monitoring system performed by one who has a trained ear for music or sound naturalness is a final check upon the proper placement.

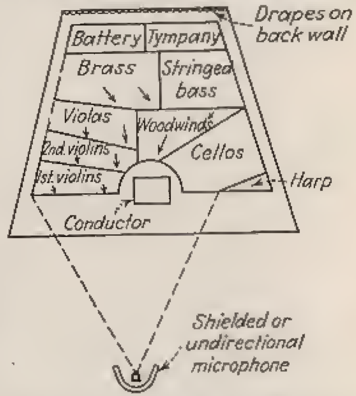


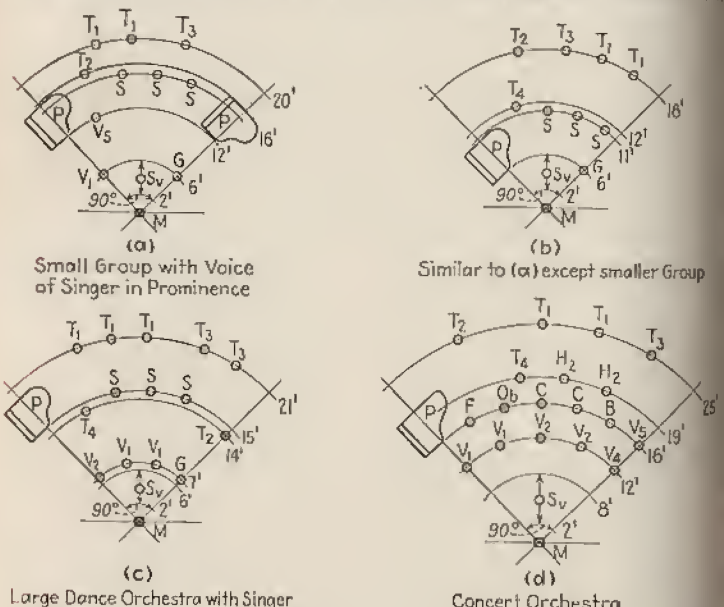
Fig. 21.—Setup of 110-piece symphony orchestra.

17. Typical Studio Arrangement. A typical setup of a large symphony orchestra before a condenser microphone is shown in Fig. 21.[1] The instruments are placed so as to obtain the desired balance for theater or auditorium work and to obtain the proper harmonic balance allowing for the microphone directional characteristics on higher frequencies.

The microphone is acoustically shielded to prevent reverberation from the auditorium behind it. Present-day microphones, such as the unidirectional and cardioid types, could be used to advantage without the shield since their response in the rear is very small. The string instruments in this setup, being the least powerful ones, are concentrated in the foreground of the group. The wood winds are next in line followed in the background by the powerful brass and percussion instruments. In this arrangement the string tone of the orchestra is given a favorable position to produce a softness to the music which will not be overpowered acoustically by the heavy brasses and percussion instruments.

[1] Hanson, O. B., Microphone Technique in Radio Broadcasting, *Jour. Acoustical Soc. Amer.*, 3, No. 1, July, 1931.

Figure 22 shows various arrangements of instruments and voices before the inductor or diaphragm type of microphone. The characteristic of this type permits the placement of the musical instruments within an area contained by an angle of 45 deg. on either side of the microphone axis. In using this type of instrument the source of sound, speaker, announcer, or



LEGEND

B - Bassoon
C - Clarinet
F - Flute
G - Guitar
H_2 - French Horn
M - Microphone
O_B - Oboe
P - Piano
S - Saxophone
S_V - Soloist (voice)
T_1 - Trumpet
T_2 - Tympani, Traps
T_3 - Trombone
T_4 - Tuba
V_1 - First Violin
V_2 - Second Violin
V_4 - 'Cello
V_5 - String Bass

FIG. 22.—Orchestra arrangements for use with a single type 50A inductor microphone.

musical instrument should not be placed closer than 1 ft. from the face of the microphone.

The bidirectional characteristics of the velocity microphone are advantageous in that the performers can be distributed on both sides of the instrument in a manner shown in Fig. 23. The uniform frequency-response characteristic of the instrument with directivity is an advantage in that the intensity of some instruments may be decreased without discriminating against their higher frequencies, simply by moving them at a larger angle with respect to the microphone axis

An orchestral arrangement involving the use of a velocity microphone as suggested by LaPrade[1] is shown in Fig. 24.

The orchestral group in this arrangement was conveniently located on one face of the instrument. To prevent reflection from a wall directly in back of the microphone, the instrument is tilted at an angle of approximately 30 deg. toward the orchestra. An exceedingly well-balanced pickup has been accomplished by this method.

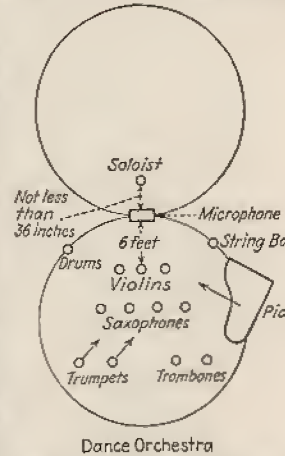


FIG. 23.—Dance orchestra microphone arrangement.

18. Volume Controls or Faders. Volume controls or faders used in high-quality broadcasting circuits should have frequency characteristics which are uniform between 30 and 15,000 cycles to prevent them from causing frequency distortion. Also essential is a very low noise level. This is normally −150 db or better. Proper shielding for protection against dust and dirt is necessary to maintain a low noise level, as well as to act as a shield against any stray r-f electromagnetic fields.

In Figs. 25 to 31 are shown various types of attenuating structures used in broadcasting technique. The type shown in Fig. 25 is frequently used

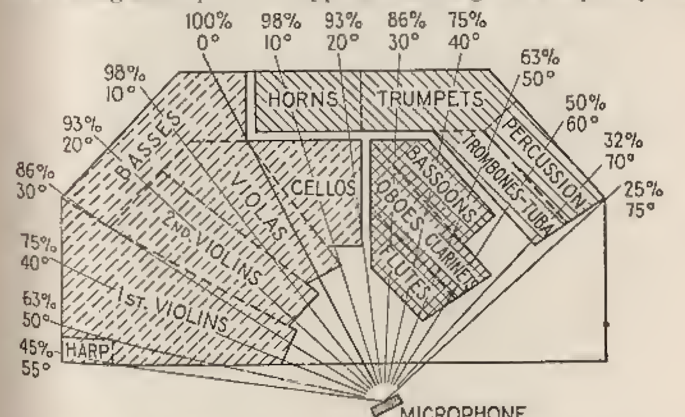


FIG. 24.—Velocity-microphone setup for large orchestra group.

as a microphone fader and is commonly known as the LT structure. When used in multiple such as for mixing several microphone outputs.

[1] LAPRADE, ERNEST, National Broadcasting Co., The Technique of Broadcasting Instrumental Groups, address at North Central Music Educators Conference, Indianapolis, March, 1935. See *Proc.* Music Educators National Conference, 1935.

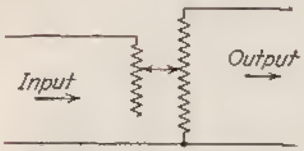


FIG. 25.—LT attenuator.

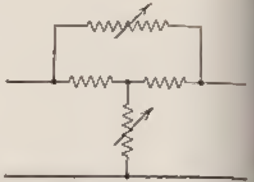


FIG. 26.—Bridged-T attenuator.

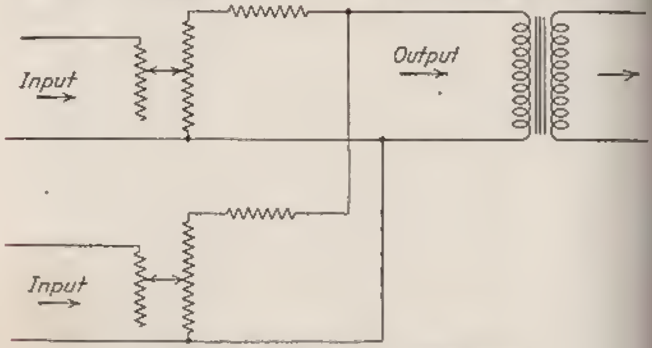


FIG. 26a.—Multiple-type LT attenuator.

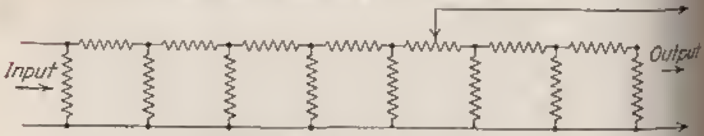


FIG. 27.—Single-ladder attenuator.

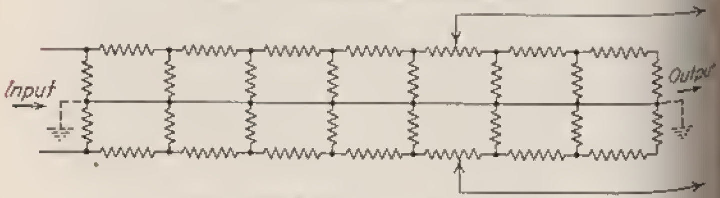


FIG. 28.—Balanced-ladder attenuator.

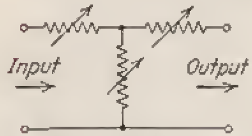


FIG. 29.—Type-T attenuator.

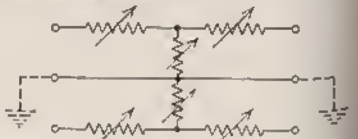


FIG. 30.—Balanced-H attenuator.

as in Fig. 26a, sufficient resistance is inserted in one output lead from each attenuator to maintain correct circuit matching. The bridged-T structure shown in Fig. 26 is used extensively for the same purposes.

The ladder attenuators maintain an impedance that remains practically constant in both directions through the middle of the attenuation range. Important features of this type of attenuator are its simplicity of design requiring fewer contacts and switches. The minimum attenuation setting of a ladder pad normally corresponds to its insertion loss which amounts to approximately 2.5 db. Where an attenuation range is required extending from zero upward, the H or T structures are used. They are usually constructed with a minimum attenuation setting of zero.

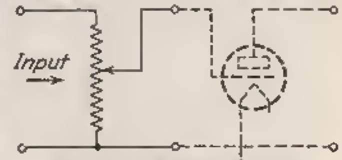


FIG. 31.—Voltage divider.

The T and balanced-H structures maintain a constant impedance in both directions when properly terminated. The balanced-H and ladder

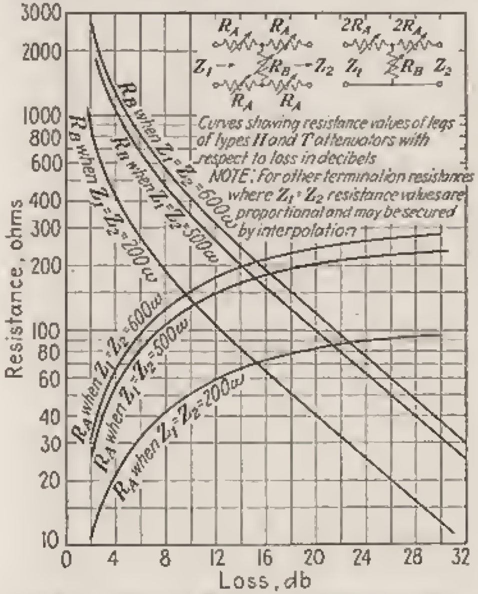


FIG. 32.—Chart for H and T attenuator design.

structures are used where the transmission circuits must be balanced to ground. They are frequently used in broadcasting circuits as master gain controls. Figure 31 shows a high-impedance voltage divider usually in the form of a gain control in the input circuit of a vacuum tube. This is a common type of gain control used on speech amplifier units.

Microphone fading is usually accomplished at the outputs of the preamplifiers and beyond where programs originate in studios. For field pickups the fading is in most cases accomplished directly at the outputs of the microphones. This, of course, requires attenuators of very low noise level. Microphones of the moving-coil dynamic and the velocity-ribbon types have constant low-impedance output over a wide frequency range and for this reason can be faded directly at their outputs.

The curves in Fig. 32 give resistance values of the branches of an H-pad suitable for a channel having an impedance of 200, 500, or 600 ohms, the range of attenuation being between 2 and 30 db. Similar curves for other impedances may be determined from the formulas published previously.[1] (See also Sec. 6 of this book.)

19. Volume Indicators. The volume level of an audio signal at any particular point in a broadcasting system is normally measured by means of a standardized instrument called the *volume indicator*. The components of the complete instrument consist essentially of a sensitive high-resistance voltmeter of the copper oxide type, an associated T attenuator for extending the range of the meter to higher readings, and a variable resistor accessory to the attenuator provided for calibrating the instrument. The instrument scale is marked in a logarithmic fashion, and superimposed upon this is an associated percentage scale. Two scales are provided, the A type tending to emphasize the VU readings and the B type in which the percentage readings are more prominent.

While an oscillograph placed across the circuits at a particular point in the system would give a true picture of the rather complex wave shapes present from program signals, it would be a rather cumbersome and expensive method of indicating the characteristics of the signal, although it could be used if properly calibrated against a standard. However, through coordination between the leading broadcasting systems and the telephone company whose facilities were also involved, there was developed[2] a standard instrument of the indicating-needle type having characteristics most suitable for the purpose of indicating signal volume.

The standard volume indicator (Fig. 33*a* and *b*) utilizes a d-c instrument with a non-corrosive full-wave copper oxide rectifier mounted in its case. Arranged for bridging, as in Fig. 33*a*, across a line, it has an impedance of about 7,500 ohms measured with sinusoidal voltage. Of this impedance 3,900 ohms is in the meter and about 3,600 ohms is external for the purpose of securing required dynamic characteristics.

The dynamic characteristics are such that if a 1,000-cycle sine wave voltage of an amplitude to give a steady reading of 100 on the voltage scale is suddenly applied, the pointer will reach 99 in 0.3 sec. and then overswing the 100 point by at least 1.0 and not more than 1.5 per cent. The frequency response of the instrument is very good as is indicated by the fact that it does not depart from its 1,000-cycle reading by more than 0.5 db between 25 and 16,000 cps.

The standard volume indicator is calibrated to read 0 VU when it is connected to a 600-ohm resistance in which is flowing 1 mw of sine wave power at 1,000 cps or *n* VU when the calibrating power is *n* db above 1 mw. However, owing to limitations in the present art, it has not been found practical to make an instrument of sufficient sensitivity to be calibrated to read 0 VU across 600 ohms with 1 mw, and therefore the instrument is normally calibrated by the application of 1.228 volts r.m.s. (4 db above 1 mw in 600 ohms) to the instrument in series with the proper external resistance to cause a deflection to the 0 VU or 100 scale point. The instrument therefore has sufficient sensitivity to be read at its normal 0 VU point on a volume level of +4 VU, which is the minimum setting of the attenuator and volume indicator; for this reason the 1-mw calibration is correct.

[1] Johnson, K. S., "Transmission Circuits for Telephone Communication," D. Van Nostrand Company, Inc., New York; and Lanterman, W. F., The Design of Attenuating Networks, *Electronics*, February, 1931.

[2] Chinn, H. A., D. K. Gannett, and R. M. Morris, A New Standard Volume Indicator and Reference Level, *Proc. I.R.E.*, January, 1940.

For routine checking of the calibration of volume indicators, a "reference" instrument properly calibrated may be used in a simple comparison method. To the terminals of a source of a-c voltage of adjustable output, the reference-volume indicator and volume indicator to be calibrated are connected in parallel. The attenuators of both indicators should be set at +4 VU. The applied voltage is then adjusted until the reference-volume indicator

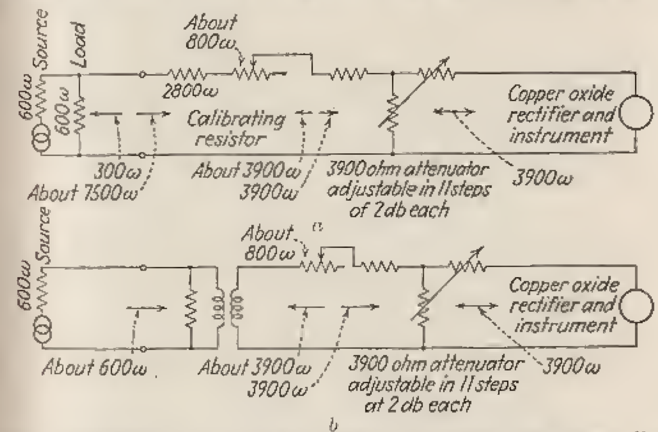


Fig. 33.—Standard volume-indicator circuit, *a*, for bridging across a line; *b*, low-impedance arrangement such as line termination.

pointer is at the 0 VU or 100 mark. If the pointer of the volume indicator being checked is not then on the 100 mark, its calibration resistor should be adjusted until it reads the same as the reference-volume indicator.

Inasmuch as the standard volume indicator has been developed and standardized as a method of checking volume of signals of complex wave shapes, it has associated with it the term VU. This term has been restricted to its intended use; hence, whenever a volume level reading is encountered expressed in so many plus or minus VU, it will be understood that the reading was made with an instrument having the characteristics of this standard instrument and expressed with respect to the reference level. Most previous types of volume indicators, even when recalibrated to a 1-mw basis, will not give indications corresponding to those of the new instrument on all types of program waves owing to the particular characteristics of the new instrument.

20. Speech-input Amplifiers. These amplifiers are sometimes termed *preamplifiers* or *microphone, line, and program amplifiers*. They comprise the apparatus necessary to increase the electrical energy output of the microphone or transcription reproduction to a sufficient level to permit its transfer by means of wire lines to the broadcast transmitter. The normal energy level of programs entering the wire lines or program loops is approximately +8 VU (+14 VU delivered from the line amplifier with a 6-db isolating pad). In Fig. 31 is shown the arrangement of preamplifiers and line amplifiers between the microphone and the wire lines. Other equipment shown are the microphone controls, volume indicators, monitoring amplifiers, and relay-switching systems.

Speech-input equipment is designed to have a substantially uniform response from about 30 to 15,000 cycles and above. The maximum gain of such a two-stage amplifier from input to output is approximately 48 db. The input impedances are 67.5/250 ohms, and the output impedances are 250/500 and 600 ohms.

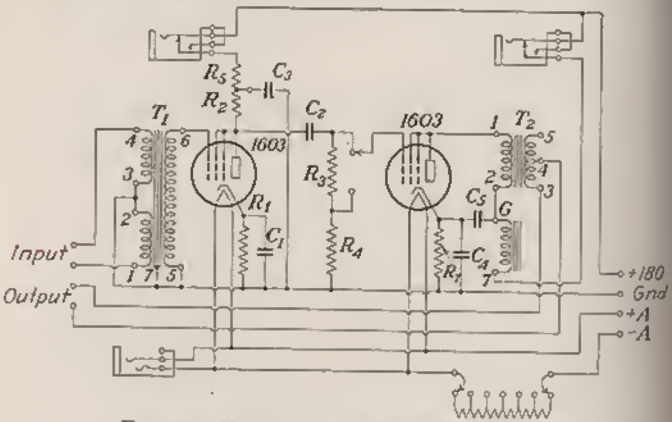


Fig. 34.—Microphone or preamplifier circuit.

PROGRAM RECORDING FACILITIES

The essential parts of a large broadcasting system usually include the facilities for recording programs for the following reasons:

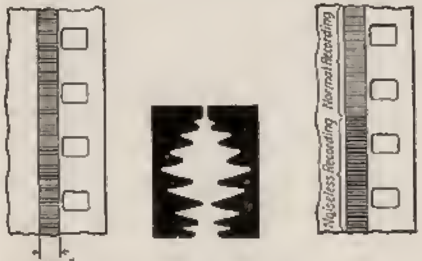
1. To have an accurate record or log of the program material actually broadcasted from a station. This is known as *reference recording*.
2. To secure a record of a studio or special events program at some distance over wire lines and thus be able to reproduce the program at a time most convenient for an audience which may be in a time zone a number of hours different from that in which the event takes place.
3. For production of recordings for use at small stations where wire line facilities are not available.
4. The recording of an audition of a person or group of persons qualifying for a program part.
5. Production of sound effects such as crowd noise, etc., for convenient use and introduction into a particular program.

21. Recording Equipment. The essential equipment required for producing high-fidelity recordings on disk records consists of the following: For bridging a program bus by means of a multiple point switch there is a limiting amplifier of the type similar to that described under Radio Facilities. It is the function of this limiting device to prevent overcutting of the record on high peaks. Following this are duplicate recording channels, each a program amplifier having linear amplitude

characteristics and a uniform frequency response over a wide range. This amplifier normally has audio power-handling capabilities up to +46 VU so that high audio peaks are not distorted before reaching the cutter head. There is a standard volume indicator across the line following this equalizer since the cutting head is placed after the equalizer, the output of which drives the cutter head.

22. Methods of Recording. It is possible to secure high-quality recording for broadcasting either by recording sound on disk records or on film. Disk records are used most extensively in broadcasting technique.

The principle methods of recording sound on film are more commonly used in sound motion-picture technique at present than for radio broadcasting. These methods include the following: (1) variable density, which may be accomplished by using either a light valve or glow lamp; (2) variable area, accomplished by using a galvanometer "vibrator"; (3) recording with a Kerr cell; (4) film engraving; and (5) a vibrating ribbon (used abroad).



0.100"
(a) (b) (c)

Fig. 35.—(*a*) Variable-density sound track produced by light-valve ribbons or glow lamp; (*b*) variable-area noiseless track produced by vibrating mirror; (*c*) noiseless recording showing greater density during periods of low modulation.

1. *Variable-density Recording.* The light-valve method uses a light of constant intensity; the ribbons of the valve move in response to a voice current and cause a sound track of variable density to be recorded on the film. When using a glow lamp to produce a sound track, a light source, whose intensity is varied, is focused on a film through a slit of fixed dimensions. Sound tracks produced by these two methods are similar. Variable-density sound tracks are shown in Fig. 35*a* and *c*. The average density of the sound track in this case acts as a "carrier" on which the modulations of the sound waves are recorded in less or greater density variations than the mean.

2. *Variable-area Recording.* In general this is accomplished by using a light of fixed intensity, which is modulated through the operation of a galvanometer, or vibrator. This produces serrations on the sound-track area of the film, as shown in Fig. 35*b*.

3. *Recording with Kerr Cell.* In this method the light-valve unit or oscillograph unit is replaced by a Kerr cell. The appearance of the sound track is similar to the variable-density sound track.

4. *Film-engraving System.* In this method an electric-cutting stylus actuated by a power amplifier is used to engrave the ound record directly on the face of the film. The position of the sound track may be inside or outside the sprocket holes. The depth and shape of the groove are similar to those used for cutting disk records (*i.e.*, from 2 to 2.5 mils in depth, and 4 to 6 mils in width).

5. *Vibrating-ribbon Recording.* Several methods developed abroad make use of a vibrating ribbon to cast a fluctuating shadow upon the sound track. One such ribbon valve, developed in Soviet Russia, can be rotated 90 deg., so as to yield at will either variable-area or variable-density recording.[1]

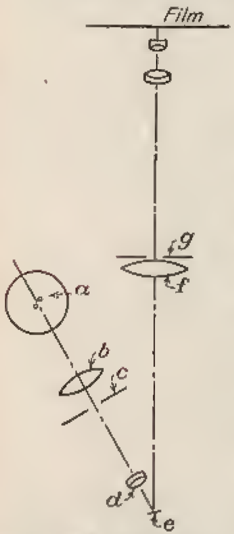


FIG. 36.—Schematic diagram of RCA Photophone recorder. *a*, recording lamp; *b*, condenser lens; *c*, triangular aperture; *d*, lens; *e*, galvanometer mirror; *f*, condenser lens; *g*, mechanical slit.

The RCA Photophone recorder,[2] used for variable-area recording, is shown in Fig. 36. Two coils actuate the galvanometer. One carries the voice current to be amplified; the other, a portion of that current which has been rectified and is used as bias. In the absence of modulation a very narrow transparent line is produced down the center of the sound track. A speech signal causes the mirror to vibrate about a central position determined by the bias current and hence to reflect to the film a varying width of the triangular aperture.

A variation of this method is *push-pull recording*, in which the sound track carries two images side by side but 180 deg. out of phase. The optical system of the reproducer focuses each recording separately on one cathode of a double-cathode "full-wave" photocell.[3]

The Western Electric light-valve recorder consists essentially of a duralumin ribbon "hairpin" in a plane at right angles to a strong magnetic field. The ribbon is approximately 6 mils wide and ½ mil thick. This ribbon is stretched by means of an adjustable spring over a bridge having a narrow slit for passage of the light from the recording lamp through the optical system to the film.

Setscrews are provided to center the ribbon accurately over the slot, which is approximately 8 mils wide and 250 mils long.

The ribbon is tuned after proper spacing on the valve to 9,500 cycles or higher, so that its natural period will be outside the range of the frequencies being recorded. A diagram of the optical system using a light valve for recording is shown in Fig. 38. The light source is provided by a special lamp having a horizontal filament. The lamp socket mounting is so adjustable that the filament can be focused properly on the light-valve slit. The sound track produced is shown in Fig. 35*a*.

[1] *Jour. Soc. Mot. Pict. Engrs.*, March, 1934, p. 158.
[2] *RCA Rev.*, October, 1936, p. 3.
[3] *Jour. Soc. Mot. Pict. Engrs.*, July, 1932, p. 51.

A portion of the speech input is detoured through the *noise-reduction amplifier* and used to control a bias current which flows through the hairpin ribbon and in turn controls the ribbon spacing. The result is a *noiseless recording* as shown in the lower half of Fig. 35*c*. The increase in sound-print density reduces the ground noise (and consequently increases the volume range of the record) to the extent of about 12 db. A *dialogue equalizer*[1] is sometimes used with wide-range recording to

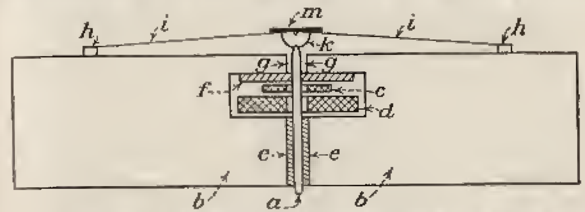


FIG. 37.—Schematic diagram of the galvanometer used to actuate the mirror of Fig. 36. *a*, silicon-steel armature; *b,b*, silicon-steel pole pieces; *c*, voice coil; *d*, bias coil; *e,e*, non-magnetic spacers; *f*, rubber pad for damping at resonance frequency (9,000 cycles); *g,g*, air gaps; *h,h*, prongs providing tension for galvanometer ribbon; *i,i*, galvanometer ribbon; *k*, mirror plate; *m*, mirror The mirror vibrates rotationally about a center through the ribbon.

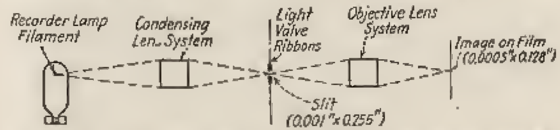


FIG. 38.—Optical system used in light-valve recording.

reduce the l-f response during dialogue and especially for intimate close-ups.

23. Glow-lamp Recorder. This consists of a two-element gaseous-discharge tube which varies its illumination in accordance with the voice currents impressed on its circuit. This produces a variable-density sound track similar to the light-valve track. The Aeolight, used by Fox Film Corporation, is one of the recorders in this class. The lamp is not focused upon the film, but a portion of its illumination is allowed to pass through a quartz slit which is in contact with the film.

The recording level for the Aeolight is approximately +12 db above zero reference level. All lamps have a steady d-c component impressed, which causes them to burn at a predetermined exposure. This exposure is modulated by an a-c component due to the introduction of voice currents from the recording amplifier. The resulting output is a variable-density sound track similar to that shown in Fig. 35*a*. The illumination from a glow lamp is approximately proportional to the amount of current flowing through it, within the normal recording range.

24. Sound on Disk Recording. The direct method of disk recording utilizes aluminum disks usually 16 in. in diameter and 0.050 to 0.060 in. thick, coated with a cellulose nitrate compound (usually miscalled

[1] *Jour. Soc. Mot. Pict. Engrs.*, April, 1934, p. 254.

"acetate"). The cellulose nitrate coating is used as the medium for recording sound modulations. These disks are suitable for immediate playback.

For indirect recording it is the usual procedure to use soft wax records approximately 17 in. in diameter and from 1 to 2 in. thick. These records are later processed to produce a hard record approximately 16 in. in diameter and ¼ in. thick.

The sound record is cut in the highly polished surface of the wax disk by means of an electromechanical recorder. The technique of cutting wax records is similar to making standard electric phonograph records. The standard speed for common phonograph records is 78 r.p.m., while for broadcasting records it is usually 33⅓ r.p.m. This speed with a 16 in. disk gives a playing time from 10 to 15 min.

Both types of records are cut with the spiral proceeding from the outside edge of the record toward the center, similar to making standard electric phonograph records.

25. Variation of Frequency Response on Disks. In recording on a disk revolving at a constant angular velocity, the cutter stylus is placed near the outside edge of the record, and, as the engraved spiral of the sound track progresses toward the center of the disk, the velocity with which the stylus travels on the disk is decreased. This velocity is directly proportional to the radius between the center of the disk and the position of the stylus. Now, for most satisfactory reproduction of the higher frequencies, the stylus must travel with sufficient velocity over the disk to provide sufficient space in the groove to permit satisfactory engraving of the h-f pulsations of very short duration. Therefore there is a tendency for more satisfactory engraving of the higher frequencies near the outer edge of the disk than toward the center. In other words, for a given cutter engraving on a disk of constant rotational speed, the frequency response one may reproduce from a disk is more satisfactory near the outer edge of the disk and is less satisfactory, especially to high frequencies, as the stylus moves toward the center. For this reason, if high-fidelity results are to be obtained, the groove containing sound modulation should not be closer to the center of the disk record than 5 in. for 78-r.p.m. recording and 8 in. for 33⅓-r.p.m. recording. Medium to good results are obtained with the groove containing the sound modulation at a radius on the disk of not less than 2½ in. for 78-r.p.m. and 4 in. for 33⅓-r.p.m. recording. For a given playing time it is sometimes possible to keep this minimum radius, cutting more grooves per inch, sometimes as many as 160 in lateral and vertical disk engraving.

26. Lateral and Vertical Disk Engraving. In the lateral system the groove depth is kept constant, and the engraving stylus moves in a horizontal fashion to produce undulations in the sides of the groove. The groove spacing therefore must be sufficient to prevent the stylus from cutting into adjacent grooves at the low frequencies.

The vertical system utilizes an engraving stylus moving in a vertical direction. The groove depth varies with the mechanical modulation, whereas the groove width is kept more or less constant, with a result that the groove spacing can be kept closer with a correspondingly greater duration of playing time.

In Fig. 39 are illustrated waves produced in disk-record grooves under conditions of "constant-amplitude" and "constant-velocity" recording. The wave marked 1 illustrates constant-amplitude engraving produced by

a constant sound level regardless of frequency at the cutter. In this case the recorded amplitude is the same for all frequencies.

The constant-velocity system utilizes constant vibrational velocity of the stylus in the record groove under influence of the cutter head. In this case the amplitude of the wave is inversely proportional to the frequency. The wave marked 2 illustrates undulations in the record groove produced by constant-velocity recording, producing an increase in amplitude with a decrease in frequency for a constant sound-level input (assuming that the entire system from the microphone to the cutter head has a uniform frequency characteristic). In this instance, where amplitude $= v/kf$, the amplitude of a wave frequency f of 100 cycles would be one-half that of 50 cycles for a constant velocity v. At the lowest frequencies, therefore, the amplitudes would be excessive if sufficient amplitudes of the higher frequencies are to be

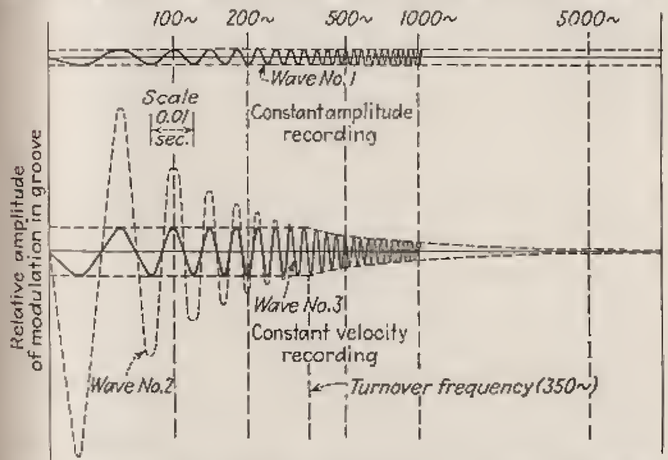


Fig. 39.—Characteristic of waves produced with the constant amplitude and the constant velocity systems of recording.

produced in the record groove. Since the groove spacing would have to be made considerable to avoid "groove crossover" or "echo" effects in adjacent grooves, due to excessive amplitudes at the lower frequencies, it is customary to cut records constant amplitude at frequencies below some point between 350 and 800 cycles and constant velocity for frequencies above this point. This is illustrated as wave 3, a solid line. The transition frequency between constant-amplitude and constant-velocity recording normally some point selected between 350 to 1,000 cycles is called the *turnover point*.

To produce a constant-amplitude cutting characteristic up to the turnover point and a constant-velocity cutting beyond, it is necessary to utilize corrective equalizers depending upon the particular type of cutter head used. In some cases the response characteristic of the electrodynamic cutter head itself is a contributing factor in the production of the constant-amplitude and constant-velocity range as well as the turnover point. This is illustrated in Fig. 40.

Commercially, direct recording has become of great importance because of its advantages of immediate playback and cheapness when producing disks in small quantities. While the nitrocellulose coating is essentially softer than the pressed records manufactured by the electroplated soft wax pro- as many playbacks as 100 may be secured from a nitrocellulose disk with well-designed lightweight reproducer. A substantially flat frequen- response may be recorded on, and reproduced from, these disks over a ran of between 50 and 10,000 cycles and higher near the outside portion of t disks. It is good practice where very high-fidelity reproduction is requi on 33⅓-r.p.m. disks to use the outside portion of the disk to compensate the loss of the higher frequencies in reproducing as the pickup moves tow the center of the disk or to divide time into two or more disks, thus permitt reasonably high linear recording velocity of the cutter stylus. A volum range of approximately 55 db has been obtained from nitrocellulose d using the lateral system of recording and reproducing. With satisfacto operating conditions over-all distortion of the combined recording and pl back operations is less than 5 per cent. This over-all distortion is also function of engraving velocity, decreasing as the velocity is increased, a decreasing with a decrease in engraved depth.

Flutter is a term used to describe vertical modulation produced in t recording groove due to the bounding of the cutter head at a frequency approximately 30 cps. It is normally caused by mechanical response of t recording head and its associated supporting-arm mechanism under excitat

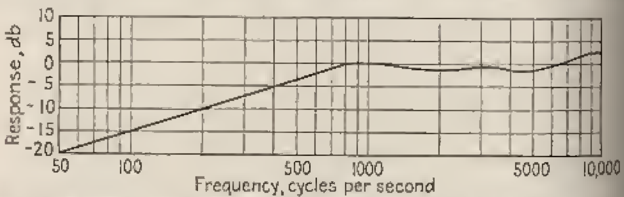


FIG. 40.—Frequency response of cutter head based upon optical measurem of the stylus tip motion for constant input.

from building noise and other l-f rumble. In observing reflections from rec grooves created by a single source of light, the effects of flutter can be note in the form of spokes or long spiral patterns extending from the inside ou the recorded surface to the outside. Under a microscope this vertical mod lation may be seen as a varying width of the cut groove. Manufactur supply stabilizers which assist in the elimination of flutter.

When recording on nitrocellulose disks, an air-suction nozzle is provi near the cutter to remove shavings or shreds so that they will not interf with the engraving process and also to provide for safe disposal of this hig inflammable material. Care must be taken to avoid dust, fingerprints, grit from entering the engraved surfaces of the disk. Otherwise there i tendency for increased noise. It is customary to engrave 120 grooves pe inch on these disks, although 96 and 112 and as high as 160 grooves per i have been used. This number is fixed by the lead screw of the recordi machine. The groove depth engraved on this type of disk is normally ab 0.0015 to 0.002 in. Commercially, it has been possible to secure recordin of this type having a noise level 50 to 60 db below the maximum modulat signal, although the average record has only a 35- to 40-db spread betw noise and modulated signal. By the method explained below for process soft wax from which pressings are made of a hard material, nitrocellu disks may be similarly processed for the purpose of making a large number pressings.

The indirect recording method requires considerably more equipment an time to manufacture the pressed disks than the direct method described above. However, for mass production, pressings can be made consideral more cheaply than single records by the direct process.

27. Necessary Equipment. Equipment necessary for wax disk recording consists essentially of a machine lathe especially designed to turn the wax record clockwise at a uniform speed, which is 33⅓ r.p.m. for broadcasting work. The carriage of the lathe is driven with a lead screw carefully machined to move the recorder holder at a predetermined rate while cutting the wax record. The lead screw is driven through a gear train which regulates the number of grooves cut per inch, usually 86, 92, 98, 112, or 120. A recorder holder provides the necessary support for the electrical recorder.

A horizontal turntable, driven through a vertical shaft, is provided for supporting the wax record. The vibration of the driving motor is eliminated on different lathes by various methods. The Western Electric lathe uses an oil dashpot placed below the lathe bench, and through which the vertical shaft of the turntable is driven. This dashpot provides the necessary damping to ensure smooth recording on the record.

The RCA machine utilizes a motor on a rubber isolating mounting. The table is driven by means of a rubber roller, the shaft of which is belt driven from the motor pulley.

The details given below refer to lateral-cut records, this being the most common type of record that has been used for broadcasting. Vertical-cut records are made by some studios for playback purposes. Both types have their particular advantages.

28. Sound-recording Channel. A schematic diagram of a typical recording setup is shown in Fig. 41 which represents a Western Electric system.

29. Preliminary or Booster Amplifier. This amplifier (see Fig. 41) is mounted between the mixer panel and the volume-control panel. It is used to amplify the output of the mixer before passing through the volume-control panel. Amplification is desired at this point to raise the recording level sufficiently high to prevent undesirable pickup from stray electric currents or other sources entering the voice-transmission circuit. It also eliminates possible noise when operating the volume-control potentiometer. This amplifier differs in detail for various systems. In the Western Electric system, it is a three-stage resistance-coupled amplifier using three 264-A tubes.

30. Volume-control Panel. The outputs from the individual mixer panels are connected in parallel, and leads from them are connected to the input of the preliminary or "booster" amplifier. The output from the preliminary amplifier is fed into a control potentiometer, which permits simultaneous adjustment of the total volume without changing the relative adjustments of individual mixer values. This panel also mounts an extension volume indicator to give a visible indication of the volume level maintained at the bridging bus.

31. Main Amplifier. This amplifier is so designated that it amplifies the output from the volume-control potentiometer and delivers the amplified current to the bridging bus circuit (or in simpler installations, directly to the power-control panel and recording machine). It is the amplifier furnishing the largest gain in the recording channel. The main amplifier differs in details for the several recording systems. In the Western Electric system it may be an impedance-coupled amplifier with input and output transformers, *i.e.*, the first stage using a Western Electric 102-type tube and the second and third stages, 205-type tubes. The total gain of this amplifier is approximately 70 db. The gain control of the amplifier is provided by a potentiometer in the input circuit. One *bridging* amplifier is required for each recording machine, its principal function being to prevent variation in individual recording circuits from introducing any loss or distortion to other circuits.

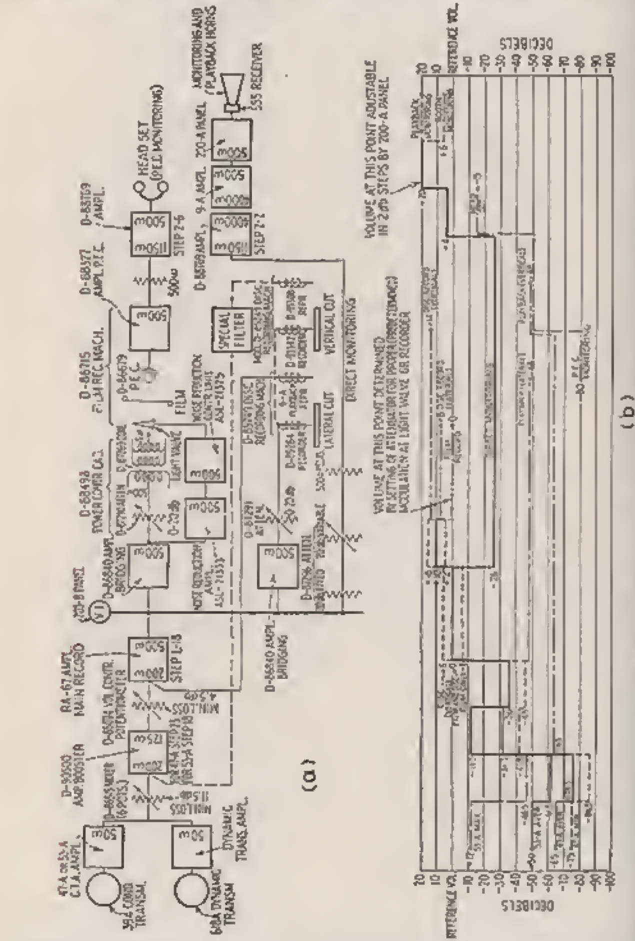


FIG. 41.—(a) Schematic of Western Electric recording system; (b) transmission-level diagram of system.

divides the electrical circuit output from the main amplifier, depending upon the number of amplifiers connected to the bridging bus. It is essentially a power amplifier, with the input transformer arranged for high input impedance, making the bridging of several of the amplifiers across the main bus practical.

The bridging-amplifier outputs are connected to the film and wax recording machines in the recording room. The wax recorder requires approximately +8-db volume level, and the film recorder around 0 db.

32. Disk Records. The grooves of a disk record are ordinarily spaced 92 to 160 per inch. For 92 grooves per inch this allows about 0.011 in. from center to center of the groove, of which 0.006 in. is the width of the groove itself. The maximum lateral motion of the stylus is thus limited to about 0.0025 in. on either side. Generally, 0.002 in. should not be exceeded. Cutters usually used are designed as constant-velocity devices. In practice such cutters have this characteristic only above 300 cycles or higher. Below this point the amplitude is independent of frequency. If the maximum amplitude for a 300-cycle wave is equal to 0.002 in. on either side of the center, then a 1,500-cycle amplitude for the same electrical input level would be 0.0004 in.

The shape of the groove varies somewhat in commercial practice, but it is approximately 0.006 in. wide and 0.0025 in. deep. The pitch of the groove is generally 0.010 to 0.011 in., leaving a space between grooves of about 0.004 in. With only this space available, the maximum safe amplitude is something less than 0.0025 in., if the walls of the groove are not to be cut too thin.

Cutting stylus consists of a sapphire, synthetic ruby, or other hard point fastened to the lower end of the stylus arm. One end of the sapphire has a rounded point about 0.002-in. radius and a cutting angle between 86 and 88 deg. for the sides.

The advance ball is a small cylindrical sapphire, ground spherically at one end and held in an adjustable mounting attachment to the recorder. This ball supports the weight of the recorder, and the arm, being adjustable, permits regulation of the depth of the groove on the wax.

Playback reproducer is provided to permit playing back the wax record immediately after it is cut for rehearsal work and test. This usually renders the wax unsuitable for processing, and for this reason two wax records are usually provided for each recording channel, one of which can thus be used for playback and the other for processing. The pressure of the needle on the wax is generally adjusted to between 15 and 20 g.

A needle provided for playback from the soft wax is designed differently from the ordinary needle used for the finished hard record. The Western Electric type has a point 0.003-in. radius. The needle is constructed on a mandrel, ground to a smooth finish, and the point given a chromium plate to improve wearing quality.

Checking Speed. The periphery of the turntable is usually divided with vertical lines, so that a neon lamp, operating from a 60-cycle source, may be used as a stroboscope to observe the turntable motion. The lines on a standard turntable are usually arranged so that with 60 cycles on the lamp, as the turntable rotates at exactly 33⅓ r.p.m., the lines will appear to be stationary. If faster than 33⅓ r.p.m., the lines will advance slowly, and, if slower than 33⅓ r.p.m., the reverse will be the case. This check of the speed is usually made with the wax record on the turntable.

Checking the Damping Action. A method of checking the instantaneous constant speed may also be used to check correct damping of the turntable. With the turntable rotating at normal speed, the oscillator for supplying 60-cycle source to the neon lamp may be adjusted until the vertical lines appear stationary. If the disk is now touched lightly by hand, the line or spot observed will appear to shift its position owing to momentary load. As soon as the hand is removed, the line or spot observed should come back to its original position. Observing the movement will determine whether the turntable has insufficient damping or too much damping.

Determining the Starting Point. Disk records for radio broadcasting use are cut in clockwise rotation from the outside in, similar to ordinary phonograph disk records. To obtain a definite starting point for the records when in use, the first groove is spaced an appreciable distance from the rest of the cut. This is obtained by a coarse speed cam actuating the lead screw at the start of recording. As the lead screw makes its first complete revolution, it moves the recorder under the influence of the cam until the recorder is in its normal cutting position.

33. Wax-suction Equipment. This equipment is provided to furnish a means of removing the shavings from the wax record during recording. The suction tube is so placed that the shavings thrown off by the stylus are carried away from the face of the wax. A central suction system is usually provided in studios having several recording channels. This usually consists of a turbine suction pump with pipe lines leading from a central suction point to a separator tank placed in each recording room. In some smaller installations, an individual bell jar, with a small suction motor, is used for each recording machine.

34. Wax Preparation. Two types of waxes are generally used in sound recording, those having a working temperature of 75°F., and those with a working temperature about 90°F. Matthews type M 75°F. working temperature, is perhaps most commonly used. It is considered good practice to maintain the room temperature for the type M wax around 75°F. when recording.

The procedure for preparing the wax consists briefly of the following steps:

1. At the center of the wax, which is usually indicated by a cross mark, a $\frac{9}{32}$-in. hole is drilled to a depth of $\frac{1}{2}$ in.
2. A course cut is made for a depth of about $\frac{1}{8}$ in. on one face of the wax and repeated as necessary to obtain a perfectly flat surface. The wax is later reversed, the first cut surface becoming the base for the finished wax.
3. On reversing the wax, a hole is cut from the other side to meet the hole drilled on the bottom.
4. A coarse cut is now made on the top surface and repeated where necessary to produce a smooth and flat surface. The wax is now ready for the final shaving or polishing cut, which is done with a sapphire or ruby cutting tool.
5. The face of the shaving knife is usually set at an angle of between 45 and 50 deg. to its line of travel, depending upon the particular design of the knife. Its rounded end is toward the center of the wax. The cutting face of the knife is set at an angle of 90 deg. to the surface of the wax. The turntable revolves in a counterclockwise direction.
6. The suction nozzle is placed close to the cutting knife, about $\frac{1}{8}$ in. from the front face and $\frac{1}{32}$ in. above the cutting edge.
7. The best finishing speed is usually determined by experience, but generally ranges from 150 to 160 r.p.m. The finished cut on the wax should give a perfectly polished surface free from ripples or blemishes of any kind.

35. Record Processing. Briefly, this consists of the various steps after obtaining the soft wax record, to produce the final hard record

for commercial use. A complete description of each step would go beyond the limits of this section. The following are the essential steps in this process:

1. The surface of the engraved soft wax disk is rendered conductive by spreading a very thin, extremely fine conducting powder, such as metallic powder, over its surface; by the finer processes of depositing silver from a solution of silver nitrate; or by sputtering pure gold of very minute thickness on the surface. This metal coating is for the purpose of forming one electrode in the electroplating process.
2. Electroplating of this record with a sheet of copper $\frac{1}{32}$ to $\frac{1}{16}$ in. in thickness deposited on the wax. The negative electroplate obtained is separated from the wax and used to hot-press a molding compound, such as shellac, mixed with a finely ground filler. The first electroplate obtained is called a *master*.
3. Two test pressings are made from the first master, after which it is electroplated with a positive.
4. From this positive, sometimes referred to as an *original*, a metal mold or *stamper* record is made.
5. From the record, duplicate originals may be made and, from them, duplicate molds or stampers. By thus making a number of duplicates, it is possible to protect the original master from injury.
6. From each stamper it is possible to obtain as many as 1,000 finished pressings.

Generally, it may be said that the duplicating process reproduces everything on the original wax engraving to such a fine degree that the only difference one may observe is in the materials, one soft wax, and the other a harder, more durable plastic, composed of shellac, vinyl, or acetate compounds mixed into a filler having very little abrasive properties. The surface of these manufactured records is considerably harder than the nitrocellulose coating on metal-covered disks used for direct playback and, with a sufficiently light reproducer, will reproduce with good quality up to 1,000 playings.

36. Re-recording. It is common practice to select desired portions of a sound record by a process of re-recording. This is done with both disk and film records. Either can be played on standard reproducing equipment, which then serves as the input to the recording system, in place of the microphones. Special re-recording equipment is also used; one type consisting of a film reproducer and a film recorder combined in a single instrument and actuated by a single motor. The output of the reproducer photocell is, of course, returned to the recorder light valve in the same casing only after it has passed through an external amplifier. This instrument is used to copy on 16-mm film a sound track that was originally recorded on 35-mm stock; optical reduction, however, is also used for that purpose. Duplication of records and films is often called "dubbing."

Re-recording is used to superimpose special sound "effects" upon a record. For this purpose two or more reproducing systems are connected as a parallel input to the recorder amplifier. The method offers superior control over the relative volume of such sounds as gunshots, background music, storms, etc., and, moreover, tends to reduce the cost of production. A library of "effect" records is maintained at many studios.

Originals intended for re-recording are sometimes made abroad by cutting a lateral track in discarded film, which is reported to be entirely serviceable for this purpose and to withstand many playbacks with damage.

37. Electrical Recording Machines. It is essential that a record machine of a precision type should have a constant speed. For reason it is usually driven by a synchronous motor. The mechan inertia of the revolving table assists in keeping the rotational sp constant, the speed regulation of the disk being usually better t 0.3 per cent. It is customary to mount the driving motor on vibrati dampers in such a manner as to prevent the motor vibration fr reaching the revolving table. Vibration from the motor shaft is ke from reaching the shaft of the revolving table either by using belt dri rubber differential speed rollers, or both. The spacing of the groo cut on the disk is controlled by gear trains and the lead screw wh moves the cutter head toward the center of the disk. The number grooves engraved per inch can be set by means of the gears. A suct tube is provided for removing the shaving or thread produced wi engraving. A microscope and groove illumination lamp facili examination of the engraved grooves. A playback pickup arm generally provided in addition to the engraving cutter mechanism permit playback of the record for quality checking.

38. Recording Heads or Cutters. The essential requirements of recording head suitable for producing high-quality recordings are follows: (1) freedom from amplitude distortion in producing undulati on the disk record, (2) suitable frequency-response characteristic over range of 40 to 10,000 cycles to produce constant-amplitude and consta velocity recording over the frequency ranges required, (3) freedom fr mechanical resonance which would tend to cause overcutting, and reasonably good efficiency in transformation of complex electric wa energy into mechanical vibration of the cutting stylus.

There are numerous types and designs of cutting heads manufactur for recording sound on disk. The most common in present-day usage a the electrodynamic and the piezoelectric crystal types.

Electrical recorder heads provided for disk recording are generally designe so that the average linear velocity of the stylus (which may be expressed a a constant × the frequency × amplitude) is proportional, over a w range of frequencies, to the impressed voltage, or $v = kfc$. The method damping the moving system varies with different records. The Wester Electric recorder uses a rubber tube about ¼ in. in diameter and 8 in. long one end of which is fitted to the armature assembly and the other end fr Oil is sometimes used to damp the armature movement in other types recorders.

A drawing of an electrodynamic type of recording cutter is shown in Fig 42*a*. With a modulated current passing through the winding of this instru ment, the armature produces and transfers to the cutting stylus mechanic undulations conforming with those in the electric wave, except that amplitude is altered somewhat by mechanical and electrical means. In F 42*b* is illustrated the RCA MI-4887 high-fidelity recording head. Th cutter head utilizes a band-pass mechanical network terminated in a mechanical resistance material. The balanced armature is centered means of a tempered steel spring. It is supported on knife-edge beari upon which the lateral stylus motion is centered. Nicaloi is used for pole pieces of the permanent magnet.

The frequency-response characteristic of this cutter head is shown Fig. 40. Below 800 cycles, frequencies are controlled to hold amplitu constant, the stylus velocity decreasing as the frequency is reduced. Abov

the 800-cycle point the response curve shows constant-velocity motion well over a frequency of 10,000 cycles. It is possible by electrical means to move the turnover point in this curve from 800 cycles to a lower frequency of, say, 500 cycles if desired.

While the electrical input impedance of the cutter head itself is approximately 5 ohms, an electrical impedance compensating network can be secured

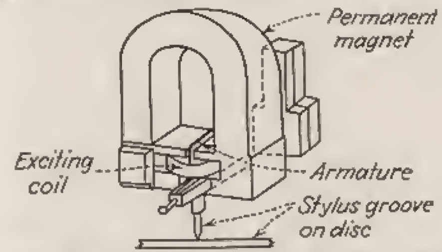


Fig. 42*a*.—Electrodynamic type of recording reproducer.

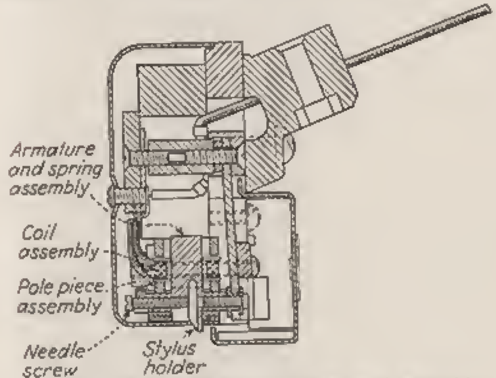


Fig. 42*b*.—High-fidelity recording-head assembly.

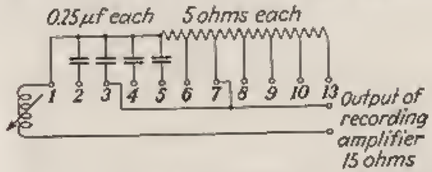


Fig. 43.—Circuit for correcting characteristic of recording head.

to retain the total input impedance of 15 ohms throughout a wide frequency range. A high-quality amplifier having a power output of 10 watts or more is recommended for driving this cutter head.

39. The Crystal Cutting Head. This type of recording head, utilizes a ...ly bimorph Rochelle salt crystal to drive the sapphire stylus to engrave sound waves laterally on disk records. For the constant-amplitude recording range the voltage applied to the crystal of the cutter head is normally volts r.m.s., while for the constant velocity range of recording it is about volts r.m.s. Since the internal impedance of the head is rather high normally 159,000 ohms at 100 cycles, the actual power consumed by the crystal is rather small, being less than 1 watt, although the power output recommended for the driving amplifier is considerably more.

A corrective equalizer is required with the cutter for constant-velocity recording above 350 cycles. Under correct operating conditions the manufacturer shows that this cutter has a frequency characteristic substantially flat within ±3 db between 30 and 10,000 cycles.

A sapphire cutting stylus is recommended for use with the cutter head. For most conditions of recording the groove depth is 0.0025 in. for cutting soft wax and 0.0015 to 0.002 in. for nitrocellulose records.

40. Measurement of Frequency Response. By examination of the frequency-response curves of the various component parts of a recording system the over-all performance of the system can be checked. The program microphones and amplifiers which feed the recording head are measured in a conventional manner with a standard sound source, beat-frequency oscillator, output meter, or cathode-ray oscillograph. Under these conditions the output of the amplifier at the terminals of the cutter head is usually flat within ±1 db between frequencies of 40 to 10,000 cps.

The recorder cutting head, however, usually has a sloping frequency characteristic (Fig. 40). The response of the cutting head alone has been measured by supplying constant level tone at various frequencies to the head and, by means of a tiny mirror attached to the stylus, reflecting a beam of light into a phototube. It is usual practice to measure the response of the cutter and disk material together.

This consists of making a recording of the output of a beat-frequency oscillator held at constant voltage at the cutter terminals. Frequencies usually recorded in order from outside to inside are as follows: 10,000, 9,000, 8,000, 7,000, 6,000, 5,000, 4,000, 3,000, 2,000, 1,500, 1,000, 800, 500, 300, 200, 150, 100, 80, and 50 cps. The completed record is then removed from the turntable; and under a concentrated single source of light, the reflection of the light source as seen in the grooves shows peculiarly patterned shapes similar to their descriptive name "Christmas tree." The pattern is symmetrical about the radius of the disk. It is actually a graphic representation of the frequency responses of the cutter and disk material together. The radius of the disk is the axis of frequency, the end of the pattern nearest the center being the lowest frequencies. The width of the pattern measured perpendicular to the disk radius is proportional to the undulations of the groove. This in lateral recording corresponds exactly to modulation depth. The phenomenon is due to the reflection of light over a wider band, the greater the ratio of modulated groove width to depth.

Inasmuch as good reproducing equipment usually has flat characteristics, the Christmas tree pattern may be produced with straight sides from the turnover frequency, of say, 500 to 7,000 cycles. Below this, it is customary to compensate the loss of low frequencies by boosting them with electrical filters in the reproducer. If it is noticed that pronounced peaks are in the pattern, the cutter head may be adjusted or filters inserted to produce the response characteristics required.

41 Record Reproducing Facilities. Transcribed programs generally originate in studios located separately from those in which recording is done. It is quite evident that, if full advantage is to be taken of the high-quality program material recorded on disk records, the transcription or reproducing equipment must also be of the precision type.

The transcription turntable is generally driven by a high-torque synchronous motor cushion-mounted within the console or cabinet. The motor shaft is flexibly coupled to the main turntable spindle. Speed regulation is reduced to a very small value for both rotational speeds of 33⅓ and 78 r.p.m. by means of flywheel inertia and a mechanical filter in the drive shaft.

Speed reduction of the RCA type 70C turntable is accomplished by means of a heavy-duty ball-bearing speed-reduction mechanism operated by a button located at the rim of the turntable disk. Noise and vibration pickup is kept at a minimum by cushion-mounting the motor and spindle housing and cushioning the suspension arms.

Special consideration is generally given to the design of a satisfactory tone arm and reproducer head for high-fidelity reproduction. The reproducer head must be light in weight and in pressure on the groove of the disk. Normally the pressure exerted by the diamond point stylus as measured by means of a spring balance or postal scale should not exceed 2 oz. A more desirable weight is less than 1½ oz. A lightweight tone arm and reproducer head assists in the reduction of record hiss or scratch noise and also the reduction of high frequencies especially near the center of the disk. Lightness also assists in securing more playbacks from a record since a lateral reproducer having a stylus which operates too stiff or having too great a pressure on the disk tends to erase the higher frequencies from the record groove.

Commercial reproducer heads generally utilize electrodynamic or piezoelectric principles as electric generators to convert mechanical force supplied by the groove modulation through the stylus assembly to the electrical generator element.

The RCA MI-4856 reproducer (Fig. 42*a*) is equipped with a permanent diamond point, the radius of which corresponds to the 0.0023-in. standard for lateral cut non-abrasive high-fidelity records. The armature is of the damped-reed type. The two upper air gaps are filled with non-magnetic material and are inactive. A linkage having a 6:1 leverage ratio is provided since the armature impedance is too high to be directly coupled to the record groove through the stylus. A diamond point is secured in the lower end of an extremely light pivot-arm spring supported vertically but rigid laterally. The pivot arm is thus permitted to rise without lifting the entire head. In the direction of useful motion transmitted to the armature the linkage has a minimum of compliance with a resultant cutoff of about 9,000 cps. This peak is reduced by means of a block of loaded rubber arranged as a selective damper approximately adjusted for the resonant frequency.

A shunt capacity located within the tone arm is generally connected across the pickup coil to react broadly with the inductance, increasing the response through the upper frequency range. An equalizer may be placed directly in the output of the pickup head to compensate for losses in the record modulations.

The piezoelectric type of lateral disk-record pickup head utilizes a bimorph crystal under torsional strain to convert mechanical modulations of the record groove into electrical waves. The sapphire stylus used with this reproducer is set in a small screw which fits the thread of a hollow magnesium block. The motion of the chuck is converted into a torsional strain in a bronze wire. This in turn conveys a twisting force to the bimorph crystal sealed hermetically within a compartment. The e.m.f. produced at the electrodes of the crystal is developed from the twisting force produced by the stylus and attachment mechanism.

This type of reproducer head is normally rather light in weight, resulting in a stylus pressure of approximately 1 oz. on the disk. It may be used for reproducing either constant-amplitude or constant-velocity recordings, the type of electrical compensating network required being dependent upon the particular characteristics of the recordings.

42. The Orthacoustic System. There is a limitation in the amplitude of the lower frequencies recorded upon a disk. This is corrected by a sloping characteristic in the response curve below the turnover point brought about either mechanically in the recording head or electrically with a suitable network. The undesirable needle h-f hiss is another limitation to be overcome for a satisfactory recording and reproducing system, as well as the l-f rumble caused by the turntable and building vibration, etc.

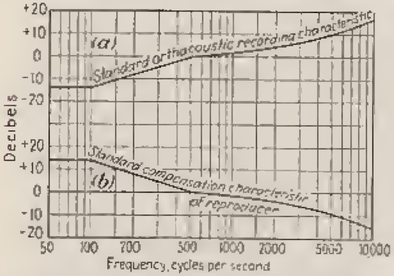


FIG. 44.—Orthacoustic recording characteristic which gives preemphasis to high frequencies.

In the RCA-NBC Orthacoustic system, recording and reproducing units are individually compensated to offset characteristics of each other and thus create reproduction which is very nearly the equivalent of the original sound. Below 100 cycles the characteristic of the recording system is made constant velocity by electric means. This tends to give preemphasis to the low frequencies. Then it rises from 100 to 500 cycles on a constant amplitude basis in accordance with the mechanical and electrical characteristics of the cutter.

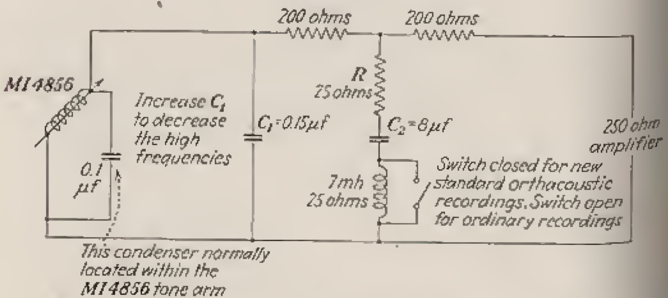


FIG. 45.—Compensation filter for Orthacoustic reproducer.

Above 500 cycles a preemphasis above a constant velocity is given to the high frequencies especially over the noise frequency range.

The necessary characteristic for reproduction is the inverse of the curve, or Fig. 44, secured by electrical and mechanical means, especially those of the transcription head itself (see Reproducer or Playback System). An over-all response curve is produced which is flat over the desired range.

43. Wire Lines. Wire telephone systems are employed almost exclusively for the national distribution[1] of programs to the various stations connected on a network.

The frequency band which is transmitted over long-distance program circuits extends from about 100 cycles to about 5,000 cycles; to transmit music with improved fidelity a wider band than the above is desirable. A few circuits are at present available which extend the band down to 30 or 50 cycles and extend the higher range by 2,000 or 3,000 cycles. Program transmission circuits must be designed to handle wide ranges of volume. At present the volume range is limited to some 25 or 30 db, from about +8 VU down to about −22 VU. Obviously, since the dynamic range of a symphony orchestra is about 60 db, the wire-line circuit necessitates some compression of the dynamic range especially on long network circuits.

44. Standardization of Transmitting Levels. To obtain optimum conditions from the standpoint of noise and cross talk, it is desirable to transmit program material into loops at as high volumes as practicable. Telephone-company experience has demonstrated that in general +8 VU is about the highest volume of program material that can be tolerated in a local cable plant of the kind in which broadcasting loops are routed, from the standpoint of interference to other circuits. In view of these conditions, therefore, +8 VU (+14 VU output of amplifier followed by a 6-db pad) shall be the standard volume level for transmitting to loops in local telephone cables. This isolating pad is for the purpose of isolating the amplifier from the telephone company loops.

RADIO FACILITIES

45. Audio-frequency Equipment. The process of transferring programs from the main control room of the studios to the broadcast transmitting station is generally accompanied by a considerable reduction in the program signal level. Attenuation caused by the wire line upon which is added that caused by the line equalizer lowers the signal intensity as much as 25 db. A line equalizer consists of a specially designed network containing correctly proportioned values of L, R, and C. Irregularities in the wire-line frequency characteristics are smoothed out by the equalizer to produce a uniform frequency response of the wire line over as wide a range as practicable.

To increase the level of the incoming signals to a sufficient intensity to drive the first tube of the speech amplifier of a broadcasting transmitter, a line amplifier is required. This amplifier is usually of a high-quality limiting type having sufficient gain to raise the audio program signal to a level of approximately +15 VU. At this level it enters the first speech-amplifier stage. The line equalizers, line amplifiers, variable attenuators, volume indicators, monitoring amplifiers, microphone for making local announcements, together with their switching equipment and jack panels, are normally mounted on racks in a shielded room called the *control room*. The shielding consists of an outside-grounded copper screen containing within it a floating copper screen.

46. Limiting Amplifier. A special type of amplifier normally used in the speech-input layout at the broadcasting transmitter is of the com-

[1] CLARK, A. B., Wire Line Systems for National Broadcasting, *Proc. I.R.E.*, **17**, 1998, November, 1929.

pressing or limiting type. This amplifier automatically reduces channel gain whenever the program peaks become excessively h... Thus it tends to prevent overmodulation. As a result, distortion d... to transmitter overmodulation can be avoided while at the same ti... the average modulation can be raised with a corresponding audio p... gain at the receiver. This is noticeable especially at low passage... program material where background noise may become objectionabl...

By rectifying a small portion of the program signal output, a ... voltage control is provided on a program signal amplifier. This act... does not just cut off the program peaks, but it reduces the gain and ... allows it to again rise slowly to normal.

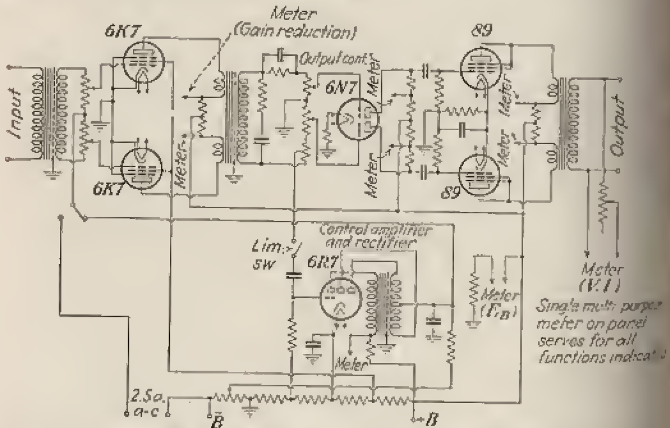


FIG. 46.—Simplified schematic diagram of RCA 96A limiting amplifier.

The signal voltage is amplified and then rectified in a diode with a r... that a variable d-c bias voltage appears across a resistor in series with ... bias voltage to the grids of the first stage of the amplifier. With an increas... signal, the bias becomes more negative and the output of the amplifie... reduced. This action does not occur, however, until the audio signal ... applied to the control tube exceeds the fixed bias of this tube.

A potentiometer across the secondary of the input transformer is util... as a variable-input control from which the corresponding input level at whi... the compression takes effect is varied. Owing to the high gain of the amp... (58 db), the beginning of the compression may be as low as −40 VU. ... vision is also supplied for adjustment of the output of the amplifier by m... of a potentiometer in the input of the second amplifier stage. By means ... this control the output level can be set anywhere within the range of −10 V... to +18 VU.

To compress sudden peaks of the program wave, the control circuit m... function very quickly. The time constant of the circuits involved is s... that the reduction in gain occurs in 0.001 sec. To prevent the gain fr... fluctuating at low audio or syllabic frequencies, there is a slow discharge ... delay circuit provided to allow the compression bias voltage applied to ... grids of the tubes in the first stage to leak off slowly and return the ampl... gain to normal in about 7 sec. This delay has been set by actual listen... tests to prevent introduction of distortion or destroy speech inflections.

...The amplifier has an output of +29 VU with 18 VU compression. The ...quency response is flat within ±1 db from 30 to 10,000 cps.

47. Program monitoring facilities are a very essential part of broad...st station equipment. In broadcasting technique, *program monitor ...dio* refers to a monitoring check on the audio signal input to the ...ansmitter, whereas *program monitor radio* refers to a check on the ...modulated signal secured by rectification of the carrier envelope as ...duced at the broadcast transmitter output. By switching from the ...*ut signal* to that produced by rectification of the modulated trans...tter carrier, the station personnel can determine by listening tests and ...easurements the relative amount of distortion produced in the broad...sting station equipment. For monitoring the outgoing program the ...rsonnel normally listens to the program monitor radio as produced ...y demodulation of the signal at the antenna system. This ensures that ...l portions of the audio and radio transmitting equipment, as well as ...he antenna system, are functioning. This is indicated by monitoring ...ud-speakers or oscillographs.

Facilities for program monitoring are provided in a room suitably ...nstructed and acoustically treated to provide a favorable place for listening tests in the judgment of quality. This may be either the transmitter room itself or an adjoining room called the control room

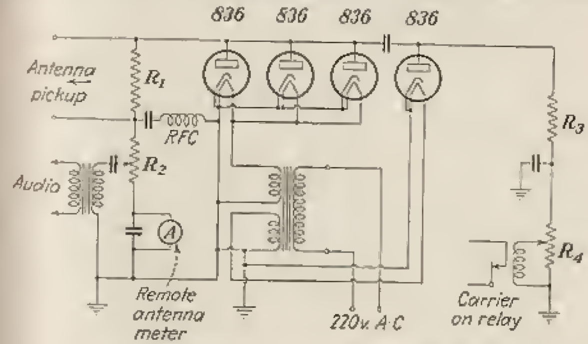


FIG. 47.—Antenna monitoring rectifier circuit.

...here the speech input is normally located. The equipment for moni...ring the audio signal consists of high-quality audio amplifiers, the gain ...f which can be regulated for proper signal volume; high-quality loud-...peakers; and associated switching equipment. The frequency response ...f the entire system should be flat over a range of between 30 to 12,000 ...ps and higher. Additional to this equipment for program monitoring ...dio is a well-designed monitoring rectifier capable of demodulating the ...arrier signal as picked up at either the output tank circuit of the radio ...ansmitter or at the antenna, preferably the latter.

...chematic diagrams of two types of antenna monitoring rectifiers, shown ...n Figs. 47 and 48, illustrate single-ended and push-pull types, respectively. ...hese rectifiers are equipped with circuits enabling them to be used as the ...mote antenna current-meter rectifier, to close a carrier-on relay or timeoutage clock relay as well as the monitoring signal for oscillograph or lo speaker. In coupling such rectifiers as shown to an antenna circuit, p

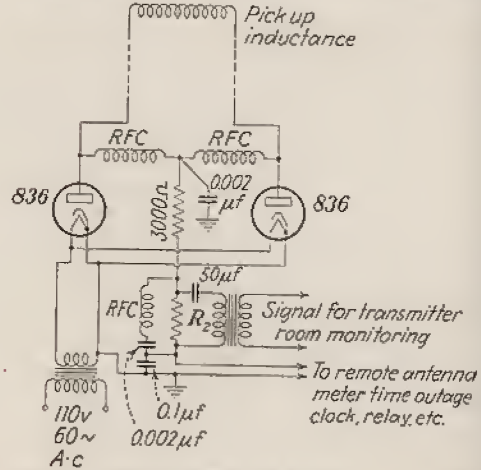


FIG. 48.—Push-pull antenna monitoring rectifier.

cautions are usually taken to prevent the generation of even and odd harmonics into the antenna circuit as produced by rectification. U certain conditions such harmo generation and radiation from t antenna system may create int ference on the harmonic freq cies. For this reason, the push- type when inductively coupled a high current point of the ant system has considerable advant over single-ended types, in even harmonics are not pronounced.

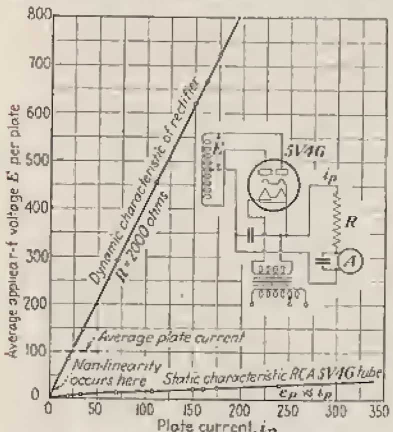


FIG. 49.—Characteristics of monitoring rectifier. (i_p refers to total plate current of two tubes.)

For rectifying the envelope of carrier wave to secure a signal loud-speaker monitoring or modulation measurements with oscillograph, it is essential that linearity characteristics of monitoring rectifier betwe the impressed voltage and the pl current is substantially strai throughout the operating ran The unit must also have a unif frequency-response characteri to provide reproduction of the nal without frequency distort

Diode rectifier tubes are used tensively for monitoring radio te phone signals. As an individual element of the monitoring rectifier, the di

self is not a linear device since the internal resistance of the diode decreases as the anode voltage is increased. The selection of diode tubes having low internal voltage drop and the introduction of sufficient resistance in the plate circuit are required in the design of a monitoring rectifier of satisfactory linear characteristics. Linearity may be further improved by application of a constant positive bias in the plate circuit so that the diode draws steady plate current over the most non-linear lower portions of the curves. In Fig. 49 these design features are illustrated for a 5V4G diode, which is a particularly good type for monitoring rectifier use due to its low internal voltage drop. Tubes of higher inverse peak voltage are often required for rectifiers of higher

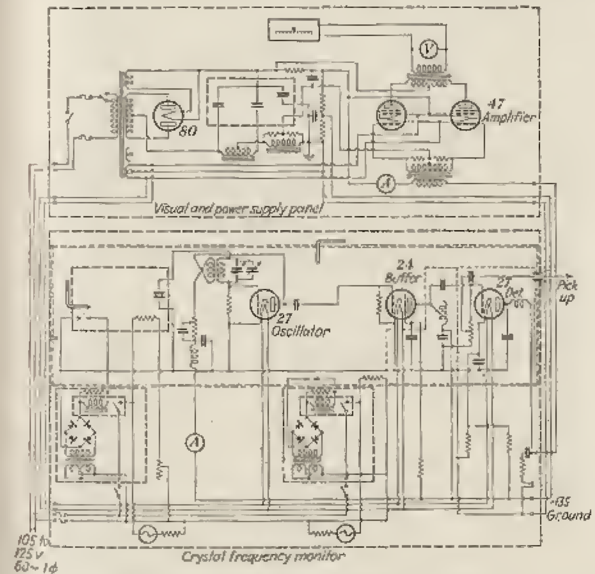


FIG. 50.—Circuit of frequency monitor.

power handling characteristics and to withstand voltage surges (such as those caused by lightning) from an antenna circuit.

The percentage distortion of a rectifier may be approximately calculated from the dynamic characteristic by using a similar formula to that used in calculating percentage distortion of three-element tubes as audio amplifiers.

48. Frequency Monitor. This instrument is required at a radio broadcasting station for the purpose of measuring the carrier frequency deviation of the transmitter. The FCC rules under Sec. 3.59 state that the operating frequency of each broadcasting station shall be maintained within 50 cycles of the assigned frequency until Jan. 1, 1940; thereafter the frequency of each new station or each station where the new transmitter is installed shall be maintained within 20 cycles of assigned frequency; and after Jan. 1, 1942, the frequency of all stations shall be maintained within 20 cycles of the assigned frequency. Under Sec. 3.60 (FCC rules) the frequency monitor is subject to FCC[1] approval and must hav stability and accuracy of at least 5 p.p.m.

The frequency monitor type EX 4180 (Fig. 50) is an approved type operates within the limits specified. The instrument contains a freque standard oscillator utilizing an accurate quartz crystal together with 27 tube operating in a special circuit having excellent frequency stabili This precision oscillator drives a 24 buffer amplifier stage with very li coupling between them. The output of the buffer is coupled to the grid a power detector stage. A few watts of r-f carrier energy, picked up from t broadcasting transmitter at some stage below the one modulated, is capa tively coupled to the mixing potentiometer connected to the grid of t detector.

Inasmuch as the standard crystal oscillator stage is adjusted to a f quency 500 cycles from the carrier frequency of the transmitter, the detec produces in its plate circuit a 500-cycle beat note. This is in turn ampl by a 47 stage to a sufficient level to operate an indicating frequency met

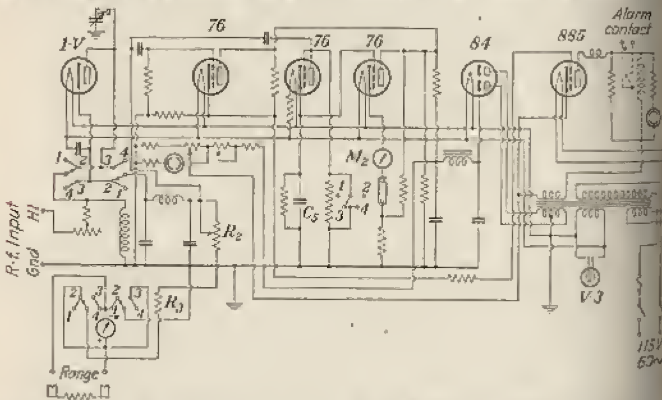


FIG. 51.—Schematic diagram of RCA 66A modulation monitor.

This meter, indicating directly any frequency from 450 to 550 cycles, ha scale calibrated from 0 to −50 cycles and to +50 cycles. With the inst ment calibrated exactly from an approved measuring station, the transmit when operating on exactly its assigned frequency, will cause the mete read zero deviation. A negative or positive drift of the carrier is indica by direct reading of the instrument. The electrical elements and circu of the oscillator, amplifier, and detector stages are contained in a const temperature cabinet at about 50°C. Contained within this cabinet another containing the quartz crystal which is kept at a constant temperat of about 60°C.

49. Modulation Monitor. Section 3.55 of the FCC rules requires th each broadcasting station shall have an approved modulation monitor operation at the transmitter to measure the degree of modulation of t transmitter and for furnishing instant warning when the degree modulation exceeds a selected specified value.

[1] "Rules and Regulations and Standards Applicable to Standard Broadcast Stat as Promulgated by Federal Communications Commission," *Broadcasting*. Nati Press Building, Washington, D. C. Also for sale by Superintendent of Docum U. S. Government Printing Office, Washington, D. C.

Referring to Fig. 51 illustrating circuits of a modulation monitor, the modulated r-f signal to be monitored impressed across diode tube 1-V is thus rectified. The resultant rectified pulsating d.c. flows in diode load resistor R_2, R_3. The average value of d.c. is indicated on the carrier meter. This is proportional to the average carrier voltage. The average component of voltage across this load resistor excites two indicating devices: (1) the meter M_2 calibrated to read modulation percentage and decibels directly and (2) a flasher and alarm circuit for providing a warning when the degree of modulation is exceeded.

The modulation indicating meter is excited in the following manner: The audio component secured from the first diode is rectified by the second 76 diode detector tube and charges condenser C_5. The voltage across this condenser is impressed across the grid circuit of the 76 vacuum-tube voltmeter stage which has the indicating meter M_2 in its cathode circuit. Circuit constants are made such in this instrument that the a-f peaks on the r-f carrier are as indicated by meter M_2. The neon flasher is operated by the relay tube, an 885, which is in turn driven from the first 76 tube under the same audio component from the tube 1-V that is used for operating the indicator meter system. If desired, the instrument may be used to operate an auxiliary alarm when the modulation peaks rise to an excessive value.

Modulation indicators are usually calibrated by means of a pure sine wave modulating signal applied to an accurate cathode-ray oscillograph and checked against the indicator. The frequency response must necessarily be flat over the audio range used to ensure accuracy of measurement over the range.

According to Sec. 3.55 of the FCC rules, a license of a broadcast station will not be authorized to operate a transmitter unless it is capable of delivering satisfactorily the authorized power with a modulation of at least 85 per cent. When the transmitter is operated with 85 per cent modulation, not over 10 per cent combined a-f harmonics shall be generated. Under Sec. 3.46 (FCC rules) design recommendations call for the total a-f distortion from microphone terminals, including microphone amplifier, to antenna output should not exceed 5 per cent harmonics (voltage measurements) when modulating from 0 to 84 per cent and not over 7.5 per cent harmonics (voltage measurements of arithmetic sum) when modulating 85 to 95 per cent (distortion shall be measured with modulating frequencies of 50, 100, 400, 1,000, 5,000, and 7,500 cycles up to the tenth harmonic or 16,000 cycles or any intermediate frequency that readings on these frequencies indicate is desirable).

The operating percentage of modulation of all stations is normally maintained as high as is possibly consistent with good quality transmission and good broadcast practice.

RADIO BROADCASTING TRANSMITTERS

Production of a broadcasting signal that will afford a means for conveying speech and music to the receiving set of a broadcast listener involves the generation of a constant r-f carrier upon which there are superimposed audio frequencies, the intensities of which conform as nearly as possible with those contained in the sound produced in the studio. The production of such a signal may be accomplished by several methods[1] of modulation.

In American broadcasting technique the amplitude system of modulation is used exclusively in the present standard broadcasting band of 550 to 1,600 kc. The advantage of a.m. for transmission in this band lies in the production of a modulation envelope containing but a single pair of side bands, thus permitting station channel separation of 10 kc.

[1] RODER, HANS, Amplitude, Phase and Frequency Modulation, *Proc. I.R.E.*, **19**, Dec., December, 1931.

Compared to the a-m system, the phase and frequency methods of modulation produce an infinite number of side bands. It is evident that greater channel separation is needed and for this reason f-m stations have been assigned to the u-h-f part of the spectrum.

The primary requisites of a radio transmitter satisfactory to operate under the present rules of the FCC for producing radio broadcasting signals are as follows:

1. Satisfactory carrier frequency stability well within the allowable FCC tolerance of ±20 cps maximum deviation.

2. Amplitude and frequency characteristics providing low over-all signal distortion.

3. Suitable safety devices to avoid hazards to operating personnel and electrical circuits and equipment complying with the National Electric Code.

4. Minimum carrier noise level, approved electrical metering facilities, minimum r-f harmonic frequency power output; and freedom from parasitic frequency emissions.

5. Low operating costs requiring an over-all high operating efficiency with respect to power input, low approved power tube operating expenses, and low expenses for operating personnel.

6. Durability, simplification of adjustment, and maintenance (requiring accessibility for repairs).

7. Reliability of service providing for continuous operation with a minimum of interruptions at rated carrier power output, modulated within legal limits.

8. Satisfactory dimensions for given power output providing for minimum installation and building costs.

9. Low initial transmitter and installation costs.

10. A pleasing appearance.

A recent trend is toward transmitters having high-level modulated and high-efficiency linear power amplifiers for the purpose of producing the desired high-quality broadcasting signal with a minimum of operating expense.

60. Typical Transmitting Equipment. In Fig. 52 is illustrated a simplified diagram of a radio broadcasting transmitter of recent design rated at 5-kw carrier power output. It is commercially known as the RCA type 5 DX.

The emitted carrier frequency of this radio transmitter is maintained well within a tolerance of ±20 cycles by a crystal-controlled oscillator unit. The present FCC regulations provide for a frequency deviation of not more than ±20 cycles for all newly licensed stations and, effective Jan. 1, 19[illegible] for all broadcast stations. This is accomplished through the development of V-cut quartz crystals having a temperature coefficient of about 1 part in 1,000,000 per degree centigrade. The mounting of the crystal is surrounded by a heater in close thermal contact with the bimetallic thermostat. The effects of changes in the ambient temperature are thus compensated and the crystal is maintained at constant temperature. There is no tuned circuit associated with the crystal input circuit. Thus it is effectively "electron coupled" to the output. A small trimmer capacitor is provided in shunt with the crystal to adjust it to "zero beat" or exactly to the desired carrier frequency.

Two crystal oscillator units are provided, one being a spare, which may be switched into use instantaneously. The output power of the crystal oscillator in use is amplified to the full 5-kw carrier output by a single 802 buffer, an intermediate stage utilizing an 805 tube to drive the push-pull 805 intermediate power amplifier stage. This drives the 892R power amplifier stage. The modulated power amplifier is adjusted for plate-modulated class C r-f operation. The output of the power-amplifier stage is normally conveyed

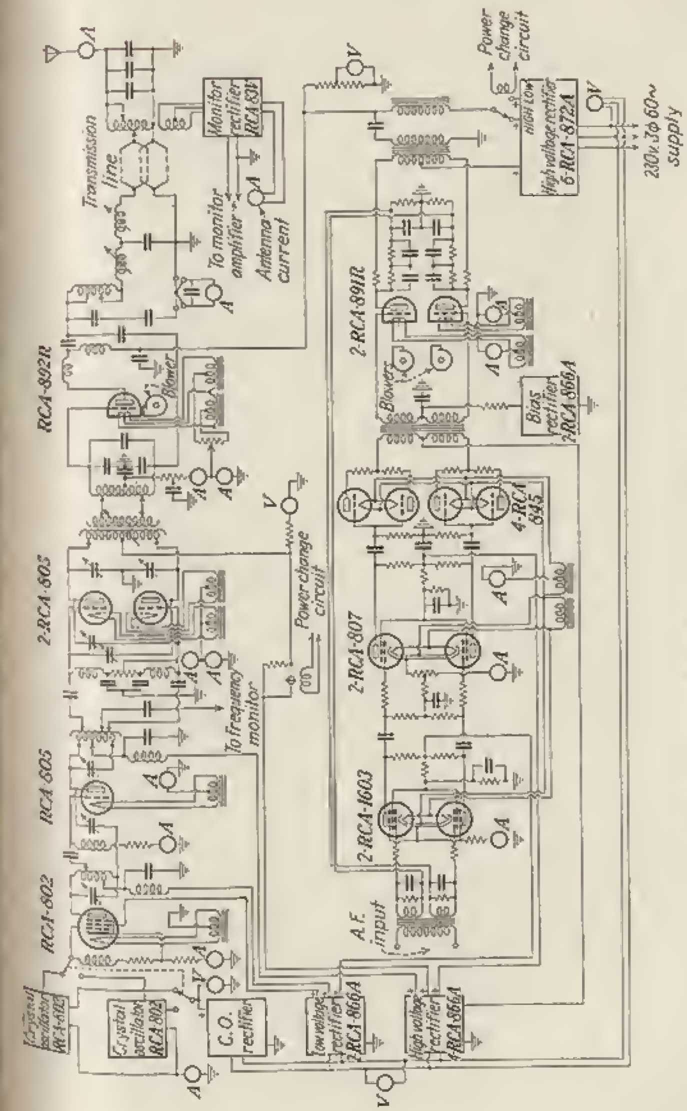


Fig. 52.—Simplified diagram of 5-kw broadcast transmitter.

to the antenna by means of a concentric or four-wire open transmission l
through network circuits reducing r-f harmonic content to a very low val

The transmitter utilizes high-level modulation, *i.e.*, the 892R stage plate-modulated by a push-pull stage containing two 891R tubes. T modulator tubes are biased for class B audio operation for the purpose securing high efficiency. The modulator is coupled by means of a mod lation transformer to the plate supply voltage of the modulated amplif The modulator tubes are driven through an input transformer by 845 tu operated push-pull as class A audio amplifiers. The 807 and 1603 stag are also operated as class A audio amplifiers.

Elaborate precautions have been taken in the design of the audio stag to control the phase rotation with respect to the frequency characterist in those circuits to which degenerative feedback has been applied. A circuit elements, especially the audio transformers, must have a minimu of phase shift over the a-f range to realize advantages from the applicati of degenerative feedback. As illustrated in Fig. 52, a potentiometer ac the primary of the modulation transformer provides a signal voltage that introduced out of phase into the input of the audio system. Hum or no generated in the r-f power amplifier appears across the modulation transform and is thus also introduced out of phase to the speech amplifier input. There fore, with regeneration, the over-all carrier noise level is very low. Measur ments indicate this to be 65 to 70 db below the signal level of 100 per ce modulation. The amplitude distortion is maintained by this system w below 3 per cent r.m.s. over the a-f range of between 30 to 10,000 cps, a the over-all frequency response of this transmitter is substantially f within 1 db over this audio range.

Features of this transmitter which merit consideration are its simplicit brought about through the use of a-c filament supply for all tubes, th eliminating filament motor-generator sets. This points toward a considerabl saving in power and vacuum-tube operating costs as well as on transmitte space requirements and initial installation costs. Reduction of carrier noi level of this transmitter to an extremely low level is accomplished through th use of indirectly heated cathodes of tubes in the low-level stages and the u of degenerative feedback. The transmitter requires no water-cooling syste since all power tubes are air-cooled.

A very small portion of the modulated r-f power produced by the pow amplifier is introduced with proper phase rotation into the first audio stag of the transmitter to reduce carrier hum and noise. Design features requ minimum phase shift in all circuits involved to permit satisfactory operatio of this system of reducing carrier noise and envelope distortion.

A transmitter is normally supplied with a phantom antenna for use durin transmitter warm-up and test periods. Switches are provided for tra ferring r-f carrier power from the output stage of the transmitter to either th radiating antenna or the phantom. The latter is designed to act as a effective resistance-load equivalent to the characteristic impedance of th transmission line. It must necessarily be capable of dissipating 75 kw of r- energy in a 50-kw transmitter when modulated 100 per cent with a sustain audio signal having sinusoidal wave shape.

Figure 53 illustrates a simplified circuit of a Western Electric 405S-1 5-kw transmitter. The modulation system consists of a low-level grid-bias modu lation applied to the Western Electric 241-B driver stage for the high-efficien power amplifier output stage. In view of the small amount of audio pow requirements from the modulator for this system, the audio and modulator stages are quite simple and of low power. Stabilized feedback is utilized between the power amplifier stage and the first audio stage to reduce over distortion and carrier noise. The value of r-m-s a-f harmonic distortion ove the range from 40 to 5,000 cps is less than 2 per cent at 85 per cent modu lation and less than 3 per cent at 100 per cent modulation. The r-m-s noise level is normally 60 db below a signal produced by a 100 per cent modulated carrier. The frequency response is flat within 1 db from 30 to 10,000 cps.

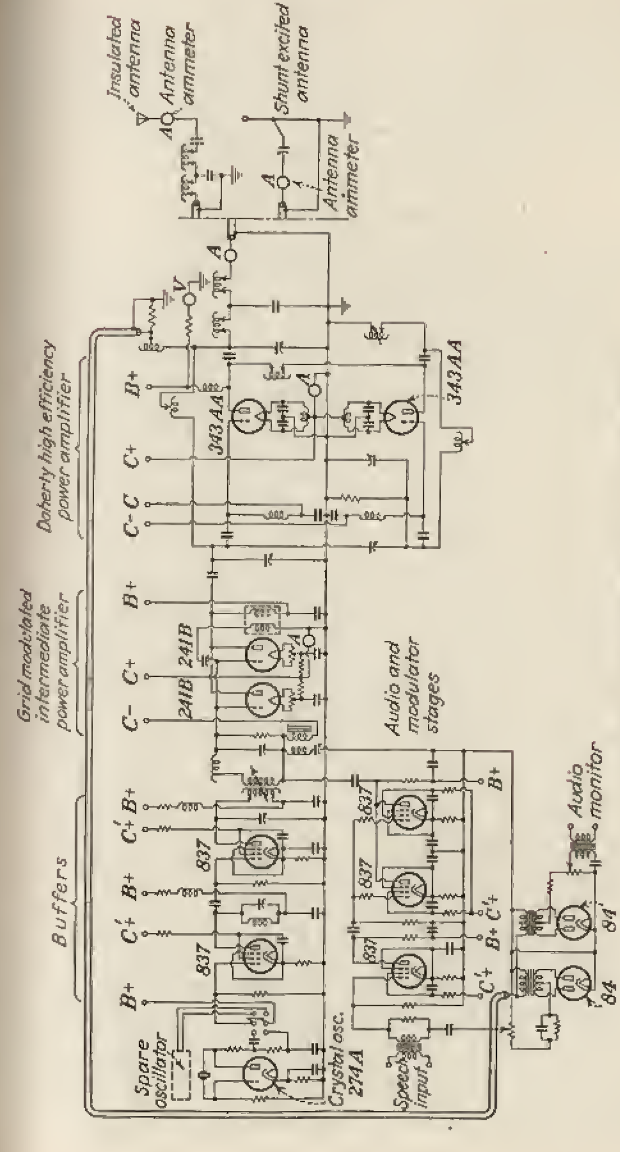


Fig. 53.—Western Electric 5-kw transmitter.

The over-all efficiency is about 15½ kw for carrier only, 16 kw for the average program, and 19.5 kw for 100 per cent modulation from a single sinusoidal frequency.

51. International Broadcasting. Transmitters for this service are operated at high frequencies and for this reason are considerably different in design from transmitters operated in the 550-kc to 1,600-kc band. They are used with directive antennas having a power gain of 10 or more and have carrier powers up to 50 kw 100 per cent modulated.

THE R-F CIRCUITS

52. Radio-frequency Amplifier Neutralization. One of the essential adjustments in an r-f amplifier circuit to obtain stability and prevent self-oscillation is accomplished through neutralizing the electrostatic capacitance of the grid-to-plate electrodes in the triode power tubes.

For the purpose of neutralizing an amplifier stage such as the r-f amplifier shown in Fig. 52, first remove plate voltage from it and apply normal r-f excitation to the grid circuit. Tune the grid circuit to resonance in the usual manner. Next connect a low-power (5- to 10-watt) high-resistance lamp across one or two turns of the plate tank inductance. The leads to the lamp should be very short and provided with clips for convenience. Next tune the plate tank circuit to resonance with the grid exciting voltage frequency as indicated by maximum brilliance of the lamp. It is to be noted that the circulating current in the plate tank circuit which lights this lamp includes the coupling effect of the grid-plate capacitance of the tube.

The neutralizing voltage of opposite polarity is obtained by connecting to the opposite end of the grid or plate tank circuits, as the case may be. The magnitude of the voltage used to neutralize the grid-plate capacity current is regulated now by adjusting a neutralizing condenser. As the neutralizing condenser is varied, the lamp will change brilliancy, and, when correct balance is obtained, the lamp will be at practically zero brilliancy. As neutralizing capacitance is changed, some slight corrections in plate tank tuning and in grid tuning may be necessary, due to interactions of these two circuits. Always tune to resonance by maximum lamp brilliancy and neutralize for minimum brilliancy.

When best results are obtained by the lamp method, remove it from the plate coil, and, if more accurate adjustment is required, a low range r-f ammeter should be inserted in series with the tank circuit. By using a meter, maximum accuracy is obtained by tuning the circuit to obtain absolute minimum current.

Since the effect of coupling between successive stages greatly affects the neutralizing, the adjustment should be made with all circuit conditions and couplings as nearly final as possible.

The power-amplifier circuit in Fig. 53 is equipped with an entirely different neutralizing system. This consists of an effective inductance shunting the interelectrode grid-to-plate capacity of the power tube. Suitable d-c blocking capacitors are provided to prevent the plate voltage from reaching the grid through this neutralizing inductance.

Neutralizing adjustments with this shunt inductance may be accomplished with a high resistance lamp or thermo-milliammeter attached to the output tank circuit in much the same manner as was described for capacitor neutralization except that neutralization is accomplished by adjustment of the shunt inductance. This system has great advantages over the neutralizing capacitor method especially where it is desirable to keep circuit tank capacity and the corresponding kva/kw ratio to a low value. This is the case where stabilized degenerative feedback is applied through an amplifier stage where a minimum phase rotation with frequency is required.

53. Class B Linear R-f Amplifiers. The operation of a push-pull class B r-f amplifier may be understood by a study of Fig. 54. Here it is shown that plate current drawn by the tubes is very closely a linear function of the grid-voltage swing. The associated output-circuit loading is adjusted so as to realize from the tube a maximum conversion efficiency. Some curves showing how plate-current efficiency varies with effective

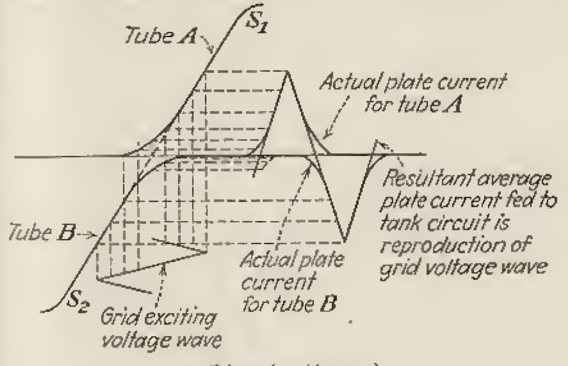


FIG. 54.—Theoretical curves showing push-pull class B r-f amplifier operation.

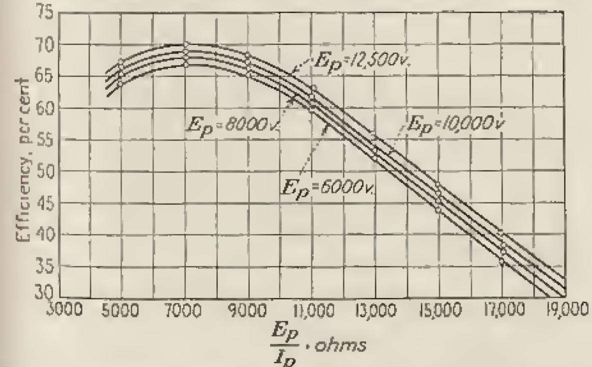


FIG. 55.—Load characteristic curves of two RCA 892 power amplifier tubes.

load impedance are shown in Fig. 55. The crest position on these curves depends upon the tube characteristics and the power factor of the circuit into which it operates. These curves were taken at a broadcast frequency by varying the load upon the output circuit of a linear amplifier stage and measuring the efficiency of the stage at various d-c plate voltages.

Under conditions where the conversion efficiency is a linear function of the grid swing, the power output is necessarily proportional to the square of the grid swing. Hence the peak power output at 100 per cent modulation is four times that at which the modulation is zero. The steady power output under conditions of sustained 100 per cent modulation is 1.5 times the output of zero modulation. Therefore, in considering power-tube requirements for a class B linear-amplifier stage, provision must be made with respect to filament emission and plate dissipation so that the tubes are capable of supplying peak power output of four times that of the nominal carrier-power output rating of the transmitter. This assumes that the modulation capability of the transmitter is 100 per cent.

In adjusting a push-pull linear amplifier, both sides of the circuit must necessarily have very nearly identical operating conditions with respect to grid swing and circuit adjustment, so that equal plate currents are measured on the individual tubes identified as A and B in Fig. 54. The grid-bias adjustment depends necessarily directly upon the plate voltage used since the position of the characteristic curve is moved with each corresponding change in plate voltage.

As illustrated in Fig. 54, with a simple triangular wave form, the method of determining optimum grid bias depends upon the point where an extension of the straight portion of the curve intersects the horizontal axis. The dynamic curves of tubes A and B have their straight portions in direct alignment. Distortion due to the lower bend in each characteristic curve is averaged out together with the kva/kw inertia effect in the output-tank circuit. On the other hand, it is illustrated in the curves that, for maximum modulation peaks with output increasing as the excitation voltage is increased, there is a limit to the output as represented by the upper bends, points S_1 and S_2, on the curves where the tube saturation points begin.

Linearity is therefore dependent upon grid bias, grid-exciting voltage, and output-tank loading. The procedure for setting taps for correct output-tank loading consists of first saturating the grids of the amplifier tubes with sufficient r-f grid driving power. Then with one-half normal, class B operating plate voltage applied, the amplifier is loaded until it delivers rated carrier power normally to the antenna, the plate efficiency being usually between 65 to 70 per cent. Then the grid-exciting voltage is reduced (usually by means of grid-loading resistors) until the amplifier stage with full plate voltage applied delivers the same rated carrier output with a corresponding plate efficiency of very nearly 35 per cent. This is the plate efficiency for class B r-f amplifier as specified by Sec. 3.52 of the FCC rules in the determination of carrier output power by the indirect method. Under this same section the plate efficiency for plate-modulated class C r-f operation of the last radio stage as measured by the indirect method is 70 per cent for transmitters having carrier power output up to 1 kw and 80 per cent for 5 kw and over.

54. High-efficiency linear-power amplifiers are a result of work to reduce the expense for operating power of broadcasting transmitters of 5-kw carrier output and above. The limitations of the class B r-f linear amplifier as previously discussed illustrate that for satisfactory operation of this system the plate power efficiency ranges from 30 to 35 per cent for the stage. Considering the driver and modulator stages and the transmitter auxiliaries, with this system the over-all efficiency from power mains to carrier power output may range from 20 to 25 per cent. The high-efficiency amplifier circuit[1] provides a plate operating efficiency of as high as from 60 to 65 per cent to be realized from a linear power amplifier.

[1] Doherty, W. H., A New High Efficiency Linear Amplifier for Modulated Waves, *Proc. I.R.E.*, September, 1936.

The amplifier circuit (Fig. 56) has been divided in block form into individual units. Voltages at points in the circuit are as indicated by symbols E_x, E_1, E_2, E_3, and E_0. The exciting voltage E_x passes into two branches. The one leads into a negative 90-deg. phase-shifting circuit, thus transforming it to the proper amplitude for grid excitation of the *carrier amplifier* tube. This grid voltage E_2 is amplified by the carrier tube, the a-c components of plate voltage becoming E_1 (180 deg. out of phase with E_2). The output voltage E_1 in passing through the impedance-inverting network shown has its phase retarded an additional 90 deg. at the output of the network. Therefore, in

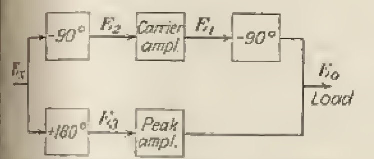


Fig. 56.—Block diagram of high-efficiency power amplifier.

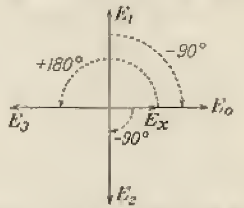


Fig. 57.—Phase relations in high-efficiency amplifier.

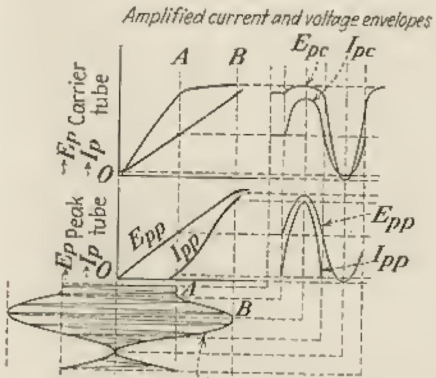


Fig. 58.—Amplifier operating characteristics.

turning through 360 deg. in this path, the resultant E_0 is in phase with the exciting voltage E_x. In the lower branch of the circuit the 180-deg. phase reversal of E_x in passing through the grid network and the phase reversal produced in passing through the *peak amplifier* tube results in a correct phase of E_0 at the load. The phase shifts may be further clarified by the vector diagram of Fig. 57, where the output voltage produced by both the carrier and peak tube are illustrated as acting in phase to produce E_0 at the load.

Figure 58 illustrates in graphical form the theoretical individual and combined operation of the carrier and peak tube branches of the power amplifier unit as produced by a modulated r-f exciting voltage (assuming sinusoidal variation of r-f voltage with modulation). For the carrier amplifier tube the plate voltage rises very nearly linearly over the region O to A, flattening off at this point due to saturation; beyond this point any increase in grid-exciting voltage for this tube produces practically no further increase plate voltage. The carrier amplifier tube plate current on the other hand rises quite linearly from O to B. Thus from O to A the operation is similar to that of a class B r-f linear amplifier operating into a load impedance of constant value; whereas from A to B there is a progressive reduction in plate impedance under influence of positive delivery of power from the peak amplifier tube on upward modulation swing, as observed through the impedance-inverting network and the plate-current rises. The plate voltage of peak amplifier tube rises linearly from O to B, where the curve flattens because of saturation. This tube is biased to a point where little positive power is delivered for grid-exciting voltages below carrier amplitude. However, owing to coupling to the carrier amplifier tube output circuit through the impedance-inverting network, a voltage exists in its plate circuit during this idle stage for the tube. Therefore, a linear variation in plate voltage for the carrier amplifier tube between O and A causes a corresponding linear voltage variation between O and A in the plate circuit of the peak amplifier tube because it is in parallel with the load. Owing to grid-bias conditions with respect to the excitation voltage for the peak amplifier tube, appreciable plate-current flow begins when exciting voltage assumes amplitude greater than that necessary for an unmodulated carrier condition. Over the region A to B, plate current rises very nearly linearly to the ... at B. It is evident that at the crest of the modulation cycle corresponding to B both the carrier and peak amplifier branches are delivering equal power outputs in phase to the load.

Adjustments required for satisfactory operation of the high-efficiency power amplifier consist of correct neutralization of the interelectrode tube capacities, correct grid biasing of the carrier and peak amplifier tubes, adjustment of the grid load resistors of both amplifier tubes and their grid and output circuits to resonance, as well as obtaining correct phase-inverting characteristics from the circuits involved. For the purpose of correct loading of the amplifier into a load, the r-f transmission line should be properly terminated to permit operation of the amplifier into a resistive load. It will be noted that for the purpose of securing the impedance-inverting characteristic required, a 90-deg. phase shift is also secured. All other phase-shift networks are utilized to compensate for this undesired phase shift. Compensation for phase shift must be effective over all useful side-band frequencies and also at the carrier frequency. The 90-deg. phase-shifting circuit in the grid of the carrier amplifier tube and the 180-deg. phase-shifting circuit in the grid of the peak amplifier tube are utilized for compensation purposes ...

55. Stabilized degenerative feedback as applied to radio-broadcasting transmitters reduces the audio-harmonic distortion and noise created within the transmitter equipment, thus providing high-fidelity performance. Reduction of carrier-noise level may be carried to as low as 65 db below 100 per cent modulation signal by utilizing degenerative feedback, even with a-c applied to the filament of all tubes. The FCC, Sec. 3.46, recommends that the carrier hum and extraneous noise (exclusive of microphone and studio noises) level (unweighted r.s.s.[1]) be at least 50 db below 100 per cent modulation for the frequency band 150 to 5,000 cycles and at least 40 db down outside this range. Harmonic distortion may be reduced to well below the FCC requirements, and in some cases the measured value of r-m-s a-f harmonic distortion in the range 50 to 5,000 cycles is less than 2 per cent at 85 per cent modulation and less than 3 per cent at 100 per cent modulation even with a high-efficiency power amplifier unit as a part of the system.

The application of stabilized degenerative feedback to audio amplifiers is described in another section. A thorough treatment is also covered

[1] Root sum square.

in the literature.[1] In the application of degenerative feedback to the transmitter (Fig. 52) it is evident that the principles as applied to audio amplifiers also apply to the circuits shown.

Application of feedback to radio transmitters is, in general, more complex than when applied to amplifiers. Theory shows that, if a part of the output of an amplifier or radio transmitter is fed back to the input and combined with the input signal in reverse phase, the effective gain is reduced. However, if the signal fed back contains noise and distortion components not present in the input signal, these components will be amplified by the full gain of the amplifier and, in traveling through the system to the point where they were picked up, will tend to neutralize the distortion and noise in the system provided that the fed-back signal is exactly 180 deg. out of phase with the input signal and the phase shift through the system is small over the range of the distortion frequencies. Under such conditions the distortion will be reduced in amplitude by the amount of gain reduction.

In operating a transmitter with feedback the over-all gain of the audio system is reduced by the amount of feedback used. For example, if 30 db of feedback is employed and the feedback voltage is removed suddenly by some fault, the program input will be 30 db too high, and readjustment of the program input level must be made instantaneously to prevent overloading. In the transmitter of Fig. 53 the feedback voltage is secured by rectifying a small portion of the power output of the power amplifier unit by means of a feedback rectifier designed for minimum phase shift. This voltage is introduced into the first speech amplifier audio stage together with the audio input signal.

With the application of degenerative feedback to cascade r-f amplifiers it becomes extremely difficult to maintain the phase of the rectified signal picked up at the output of the transmitter sufficiently close to the 180-deg. rotation required throughout the entire a-f range. Unless all networks in the entire cascade system are correctly designed, the kva/kw ratio of all r-f tank circuits are kept to a very low value, and stray capacities are minimized, there is an accumulative phase shift through the feedback loop wherein the degenerative system is active.

Under conditions where the voltage fed back to the audio input of the transmitter after passing through the feedback loop is other than 180 deg. out of phase with the input signal, less noise and distortion cancellation result. This is especially true under conditions where the phase shift of the feedback loop becomes less than 90 deg. or more than 270 deg. At frequencies where the phase shift approaches zero and 360 deg. from that of the input signal, stabilizing circuits are necessary to prevent violent oscillation of the entire transmitter at these frequencies, provided, of course, that the amplification around the loop is at least unity. These frequencies are sometimes referred to as those at which the phase "turns over." For the h-f turnover point, say around 25 kc, an adjustable stabilizing filter may be utilized in one of the low-power speech amplifier audio stages. This prevents oscillation or singing of the transmitter at the particular high a.f. where the condition exists and for this reason is called the "anti-sing" circuit. In addition there may be required a l-f (below 100 cycles) stabilizing circuit in one of the low-power audio

[1] Black, H. D., Stabilized Feedback Amplifiers, *Bell System Tech. Jour.*, January, 1934.

stages to prevent oscillation at the l.f. at which another unstable condi-tion exists.

By correct proportioning of all constants of the a-f and r-f stages and associated networks throughout the entire section of the transmitter containing the feedback loop and by application of the stabilizing circuits as mentioned together with careful transmitter adjustments, an effective amount of feedback can be normally secured for cancellation of noise and distortion.

MODULATION EQUIPMENT

56. The Speech Amplifier. An audio-amplifier unit employing power tubes is usually necessary as the preliminary part of the audio system of a transmitter to raise the audio-signal intensity to a sufficient amount to swing the grids of the modulator tubes. Resistance coupling is frequently used in speech-amplifier circuits. In Figs. 52 and 53 are shown simplified circuit connections of typical transmitter speech amplifiers.

Amplitude modulation provides a means for reproducing a signal containing a distortion not exceeding a few per cent with the carrier fully modulated. In broadcasting transmitters it can be effected by either plate or grid modulation. When grid modulation is applied to a power amplifier tube, either by bias-voltage or r-f grid-voltage change, the efficiency of the power amplifier is rather low, ranging from 30 to 35 per cent. A plate-modulated radio stage operating as a class C amplifier has a comparatively high efficiency ranging from 70 to 80 per cent. This advantage of higher efficiency, however, is offset by the low efficiency of the plate modulator unless a class B audio amplifier is used for modulating. Therefore there is not much difference in the two systems, in so far as efficiency is concerned, with respect to power and vacuum-tube costs except under conditions where modulating power for a class C r-f output stage is supplied from a modulator of rather high efficiency.

When the power-amplifier stage of the transmitter is plate-modulated the setup is called a *high-level system* of modulation; whereas a transmitter modulated in a low-power stage of the transmitter and followed by a class B r-f power amplifier is termed the *low-level system* of modulation.

57. Modulators and Modulated Amplifiers. In Fig. 59 is shown a constant-current system of modulation due to Heising.[1] The modulator and modulated amplifier are connected in parallel with a constant-current source of supply. This is connected to the common plate lead through a large inductance L_1 called the *modulation choke*.

The dynamic modulating characteristics can be determined with a fair degree of accuracy from the static characteristics of the modulator tubes in a method illustrated in Fig. 60. The modulated amplifier is assumed to be a pure resistance load in parallel with the plate resistance of the modulator tubes and both assumed to be supplied with power through a modulation choke of infinite impedance. The sum of the instantaneous currents in the amplifier and modulator in this case is a constant. An approximation is made of the number of modulator tubes required to modulate a given r-f amplifier. The plate-current ordinate for a single tube must be multiplied by the number of modulator tubes before the load line BA can be plotted, the slope in amperes per volt which depends upon the load resistance produced

[1] HEISING, R. A., Modulation in Radio Telephony, *Proc. I.R.E.*, **9**, 365, August, 192[illegible]

by the amplifier. Line BA was chosen for two modulator tubes operating at 3,000 volts plate into an amplifier of 2,000 volts and 150 ma or an effective resistance of 13,333 ohms. The mean modulator plate current I_0 is chosen from allowable plate dissipation and load line BA drawn in about operating point C. The modulator grid voltage swings from $-\frac{1}{2}E_f$ (filament voltage) to equal grid voltage on the other side of the operating point. By taking

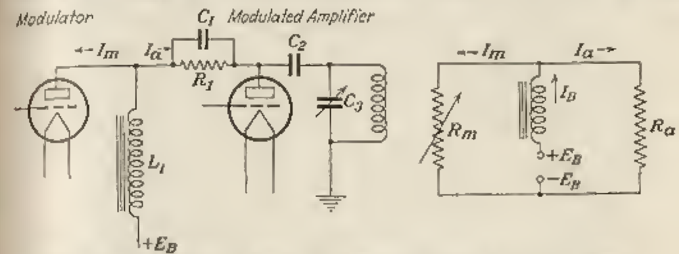


FIG. 59.—Heising constant-current modulator and equivalent.

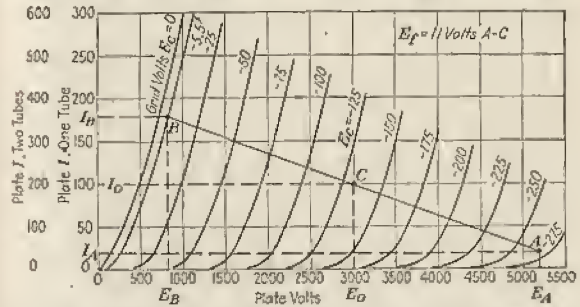


FIG. 60.—Method of determining modulator characteristics.

readings of plate current and voltage from end points of the load line, the following information becomes available:

$$\text{Modulation factor} = \frac{E_A - E_B}{2E_0}$$

where E_A = maximum plate-voltage swing
E_B = minimum plate-voltage swing
E_0 = d-c plate voltage at operating point C.

$$\text{Per cent 2nd harmonic distortion} = \frac{\frac{1}{2}(I_A + I_B) - I_0}{(I_B - I_A)} \times 100$$

where I_A = maximum plate-current swing
I_B = minimum plate-current swing
I_0 = plate current at operating point C.

$$\text{Power output in watts} = \tfrac{1}{8}(E_A - E_B)(I_B - I_A)$$

58. Design for High Audio Fidelity. In the design of the modulated amplifier circuit of the above system certain elements of the circuit must be properly proportioned to afford a uniform frequency characteristic. The capacitance of C_1 (Fig. 59) should be large enough so that its impedance at the lowest frequency to be transmitted is less than one-third R_1, or the plate-dropping resistor.

The capacitor C_2 provides an r-f path from plate to filament of the amplifier tube and at the same time breaks the path for the d.c. It must also break the path for higher frequency a-f current and permit it to flow through the amplifier tube. It should, therefore, be no larger than necessary to conduct the r-f plate current without producing excessive phase shift in the plate current under conditions where C_2 is less than $2C_3$.

Sufficient impedance of the modulation choke over the a-f range is another important factor in circuit design. Its impedance at the lowest a.f. should be at least two times the effective resistance load produced by the r-f amplifier tube. The choke should be free from inherent self-capacitance defects over the frequency range to maintain a sufficiently uniform high impedance at the higher frequencies.

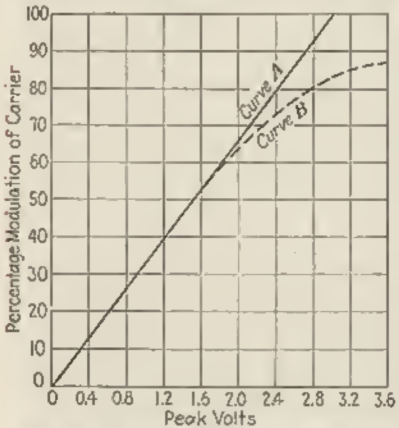


FIG. 61.—Amplitude curves taken on a modulated amplifier. Curve A taken on stage with sufficient driving power applied to saturate grid. This shows negligible amplitude distortion. Curve B taken on stage with insufficient grid excitation to cover positive peaks. Amplitude distortion becomes noticeable at 60 per cent modulation and increases with higher levels.

High-quality signal reproduction requires that amplitude distortion should be kept at a minimum. A common cause of amplitude distortion is due to underexcitation of the grid of a modulated amplifier tube when plate modulation is applied. This results in insufficient driving voltage during periods of high plate-voltage swing and consequently peak-output limiting. Trouble from this cause shows up quite clearly upon an amplitude curve or upon an oscillograph in the form of chopped-off positive peaks. In Fig. 61 are shown amplitude curves taken on the modulated carrier of a stage the grid of which was excited to saturation as shown in A and underexcited in B. It is a custom to have available a surplus of driving power for a modulated amplifier to prevent any possible occurrence of amplitude distortion.

The constant-current or Heising system of plate modulation is often designated as a class A system, since the modulator tube performs under conditions similar to those encountered in a class A amplifier. Conditions of operation of a tube in a class A system may be defined as those under which the plate current of the tube does not pass through zero at any time during a grid-voltage cycle.

A vacuum tube performing as a class B audio amplifier or modulator operates with a negative bias voltage fixed at a condition approaching plate-current cutoff. Therefore plate current of the tube increases with

a positive grid-voltage swing, but, as the grid voltage passes through the positive half of the cycle and swings negative, the plate current is cut off and remains so until the grid again swings positive.

Operation of a tube as a class B amplifier may be defined as that under which the plate current for the tube flows for one-half of a grid-voltage cycle. By virtue of a push-pull circuit arrangement shown in Fig. 62 it is possible to develop a combined output plate current from two tubes which conforms with the grid-driving voltage throughout the cycle.

A properly designed class B system permits a much higher plate efficiency to be secured from a given set of tubes and correspondingly a much greater output from them than with a class A system. This

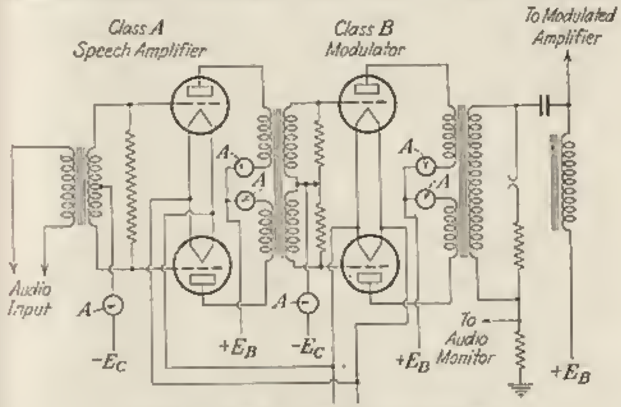


FIG. 62.—Class B push-pull modulator.

efficiency has been made to reach as high as 66.6 per cent with a small percentage of audio harmonic distortion.

Inasmuch as it is often necessary to drive the grids of class B audio amplifiers into their positive grid-current region to obtain maximum power output, it is important that the driver-amplifier stage for the modulator stage should have a good output-voltage regulation. This calls for driver tubes having a sufficient output capacity to deliver an undistorted voltage to the grids of the class B stage, even though there is a non-uniform increase of load on the driver stage caused by the class B tubes as they are driven through the positive grid-current region of their dynamic operating characteristics.

FREQUENCY MODULATION SYSTEMS

The method of program signal transmission by means of f.m. utilizes a frequency variation or deviation at the audio rate, the deviation frequency being a small percentage of the unmodulated carrier frequency.

Assume the existence of an f-m transmitter operating on 42.6 Mc and that a maximum deviation of ± 75 kc is desired. Then a sustained sine wave of, say, 1000 cps may be applied to the modulator audio input, the amplitude of the audio signal adjusted to provide ± 75 kc deviation.

This would result in utilizing the full modulation capabilities of the transmitter. With a complex program input, the frequency deviation at any instant corresponds to the amplitude of the complex wave at that instant.

Channels for f-m transmissions have been assigned 200 kc apart, which has been found to be a sufficient carrier separation to allow a frequency deviation of as much as ±75 kc. The width of the band required[1] in the frequency spectrum is *at least* twice the value of the highest modulating frequency or twice the frequency deviation, whichever is greater. Important side-band components may occur outside these limits, however.

59. Methods. There are diverse methods of producing f.m. on an r-f carrier. Two rather different systems have been classified as (1) direct f.m. and (2) indirect f.m., accomplished primarily by phase modulation.

Direct f.m. is produced by frequency modulating directly the master oscillator stage, as illustrated in Fig. 63, which has a normal unmodulated

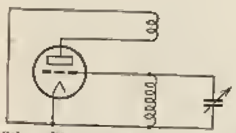


Fig. 63.—Frequency-modulated oscillator.

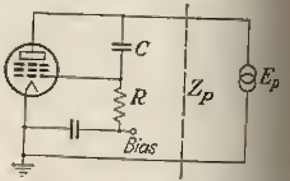


Fig. 64.—Reactance-tube modulator.

carrier frequency of either the transmitter output frequency or a convenient subharmonic thereof. In papers[2,3] giving a mathematical treatment of f.m., it has been illustrated that, if the tank circuit constants of the oscillator stage are varied in accordance with the audio-input frequencies, there will be produced a resultant f-m output signal. With some device operating as a condenser microphone varying the capacity of the tank circuit of the master oscillator, there may be produced an r-f carrier frequency modulated to conform with the sound undulations vibrating the microphone diaphragm. This illustrates f.m. by the direct method.

A modified form[4] of the direct system of frequency modulating a transmitter is accomplished through the use of a tube (Fig. 64), employed as a variable reactance. Here a variable reactance is caused to exist between the cathode and anode of the reactance tube by grid-bias variation at an audio rate. By supplying the grid of the reactance tube with r-f voltage previously passed through a phase-shifting circuit of suitable resistance and capacity, the grid-excitation voltage is caused to

[1] Carson, Notes on the Theory of Modulation, *Proc. I.R.E.*, February, 1922; Van der Pol, Frequency Modulation, *Proc. I.R.E.*, July, 1930.

[2] Roder, Amplitude, Phase and Frequency Modulation, *Proc. I.R.E.*, December 1931.

[3] Crosby, M. G., Frequency Modulation Propagation Characteristics, *Proc. I.R.E.*, **24**, No. 6, June, 1936.

[4] Crosby, M. G., Frequency Modulation Noise Characteristics, *Proc. I.R.E.*, April 1937.

be in phase quadrature with the plate voltage. Then the a-c portion of the reactance-tube plate current i_p will be very nearly

$$i_p = g_m e_g$$

However, since

$$e_g = jKe_p$$

then, under influence of the phase-shifting network,

$$i_p = jKg_m e_p$$

and

$$Z_p = \frac{e_p}{i_p} = -\frac{j}{Kg_m}$$

From which the equivalent capacity produced by the tube

$$C_e = \frac{Kg_m}{2\pi f}.$$

The reactance tube may be caused to appear as an equivalent variable capacity across the oscillator tank circuit.

For satisfactory transmission of the f-m signal it is essential that, in addition to producing the modulated wave, there must be present a satisfactory carrier-wave stability. For this reason it is necessary to add a stabilizing circuit by means of which the average frequency of the wave is compared to that of a precision crystal oscillator and thus to supply a compensating voltage to the grid of the modulator tube. The compensating voltage supplied is proportional to frequency deviation from the crystal oscillator standard. The action of the circuit is somewhat similar to that described under a-f degenerative feedback except that the improvement in frequency stability in this case is proportional to the loop gain or $\mu\beta$, where μ is the frequency compensation of output frequency resulting from 1 volt change of modulator-grid voltage and β is the volts produced by the frequency comparison circuit for unit (kilocycle) frequency change.

The circuit[1] required to produce the d-c compensating voltage from a given frequency deviation consists of the precision crystal oscillator standard, the mixer tube or converter stage, and the discriminator and detector stages. An i.f. of about 1,500 kc is produced by mixing the transmitter and crystal output frequencies. This in turn is applied to the discriminator, utilizing practically the same circuit as that in f-m receivers when connected to a double diode rectifier. This produces a d-c output potential proportional to deviation of the applied frequency as compared to that to which the circuit is adjusted. Since the feedback voltage utilized for frequency stabilization is caused to pass through a low-pass RC network of sufficient time constant, the circuit has practically no effect on the audio modulating frequencies. Stabilization of the average carrier frequency is for the purpose of preventing a change of the mean carrier frequency during modulation and permits the same output carrier frequency regardless of whether or not modulation is applied. The discriminator is provided with a linear characteristic as broad as the maximum frequency swing produced in the output frequency. A compromise on the band width is necessary, however, to maintain a steep characteristic in the discriminator circuit, and thus, to provide a sufficient amount of frequency stabilization, the band width of the discriminator should not be too great. In commercial transmitters there is normally a linear characteristic over the range of ±100 kc with the discriminator peaks separated by 400 kc, thus providing good over-all stability either with idle

[1] Weir, I. R., Comparative Field Tests of Frequency Modulation and Amplitude Modulation Transmitters, *Proc. Radio Club Amer.*, **16**, July, 1939.

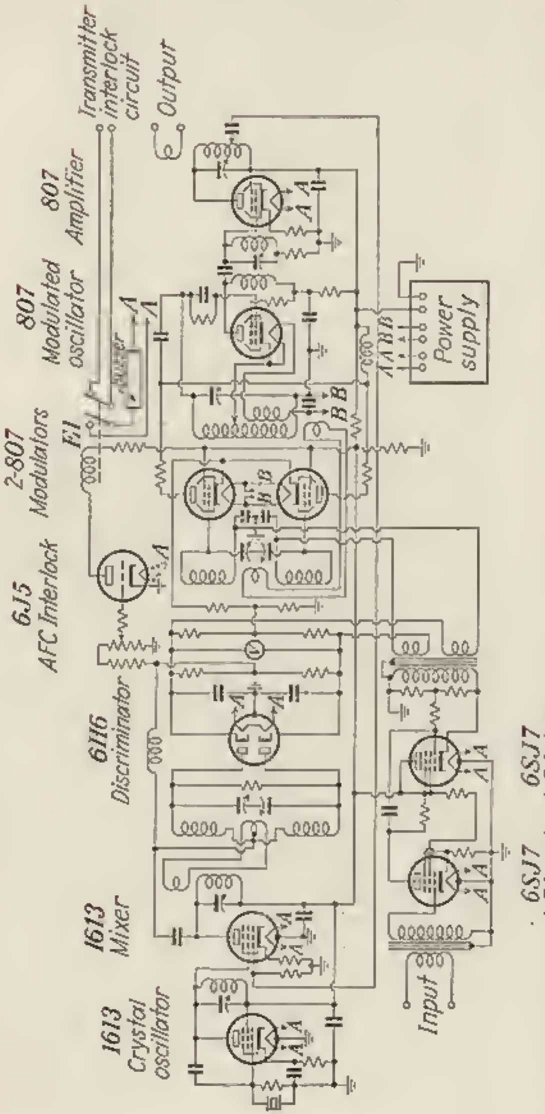


FIG. 64a.—Complete f-m transmitter circuit producing a compensating voltage from a frequency deviation as used in RCA type f-m 50A transmitter.

carrier or under full modulation. It also provides full frequency control over sufficient range to prevent the oscillator from drifting out of the control range.

An important advantage of the f-m transmitter over one utilizing the amplitude system lies in the fact that the efficiency of the r-f amplifier stages can be as great as it is for class C telegraph service. At ultrahigh frequencies this may be between 50 and 70 per cent.

A typical f-m transmitter operating on a carrier frequency of between 40 and 44 Mc may have an over-all distortion below 1½ per cent at all modulating frequencies from 30 to 7,500 cycles. The frequency characteristic may be flat within ± 1 db from 30 to 15,000 cycles, and the carrier noise level may be better than -60 db below a signal produced by full modulation swing of the transmitter. It may be designed to operate with a normal maximum modulation frequency deviation of ± 60 kc and to be linear within a deviation range of ± 75 kc for use under the present 200-kc assigned channels. The power output of this transmitter may be increased by additional class C r-f stages up to 50 kw.

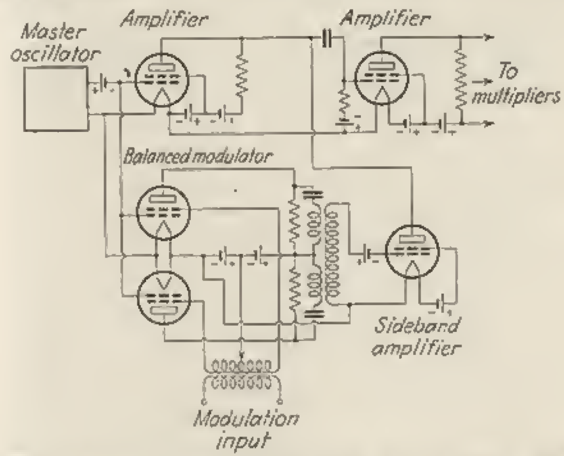


FIG. 65.—Schematic diagram of Armstrong frequency modulator.

The indirect or phase-modulation method of producing frequency modulation[1] consists in general of a constant-frequency oscillator, a modulator (the function of which is to change the phase of the oscillator output as illustrated in Fig. 65), and a series of multipliers to increase the amount of phase modulation sufficiently to secure the frequency shift or modulation required in the radiated signal. Results are secured by splitting the oscillator output into two paths. One path contains a phase-shifting network which shifts the phase 90 deg. and, in the other, a balanced modulator generating side bands with a suppressed carrier.

A combination of these two signals produces a phase-modulated signal with a phase-shift modulation capability up to ± 30 deg. with satisfactory linearity. A frequency modulated wave is derived therefrom by transmitting the signal through frequency multipliers. A multiplication of several thousand times is required to obtain deviations of ± 75 kc.

[1] ARMSTRONG, E. H., A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation, *Proc. I.R.E.*, May, 1936.

To produce f.m. and at the same time maintain a constant deviation frequency, the phase modulation must necessarily be inversely proportional to the modulating frequency. Therefore with this method it becomes necessary to have the amplitude of the phase-modulated signal decrease in proportion to the frequency of the audio input to secure a flat a-f transmitter response. This is usually accomplished by a corrective network in the audio circuits ahead of the modulator.

The amount of frequency multiplication required following the performance of phase modulation to secure the desired f.m. by the indirect method depends upon (1) the amount of phase modulation produced by the modulator, (2) the lowest a.f. transmitted, and (3) the deviation or frequency swing at the

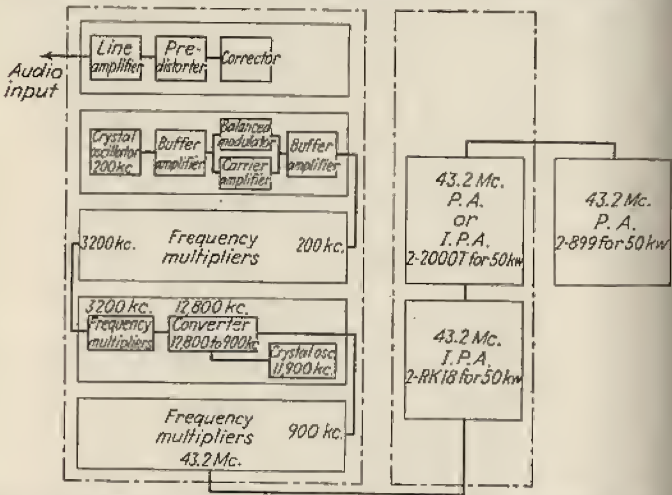


FIG. 66.—Frequency-modulation transmitter utilizing phase-shifting network.

output frequency. For a phase shift of 0.5 radian, frequency deviation of ±60 kc and lowest a.f. 30 cps, the frequency multiplication required is 4,000 times.

To obtain this amount of frequency multiplication, the initial oscillator frequency must be multiplied in several stages, then heterodyned down to a lower frequency, and then again multiplied a number of times more to secure the output frequency.

The 1-kw Western Electric 503A-1 f-m transmitter utilizes reactance tubes directly as frequency modulators in a manner as shown in Fig. 66a. The method used to maintain constant the mean carrier frequency is appropriately called "synchronous f.m." since it operates by comparing the mean frequency (measured in total carrier cps) of the f-m oscillator to a precision-fixed frequency standard. The difference frequency thus derived is then utilized as a control medium for mechanically retuning the oscillator stage and thus keeping the oscillator frequency an exact multiple of the standard. The method used to control the frequency of a turbine-driven generator supplying electric power is similar.

Figure 66b shows the frequency-stabilizing circuit, which functions through a small portion of the 5-Mc f-m oscillator (assume a 40-Mc carrier), being fed back through frequency dividers to obtain a 5-kc frequency equal to that of the precision quartz crystal frequency standard. The 5-kc frequency, a much lower submultiple of the 40-Mc carrier, is necessary to produce a difference-frequency sufficiently low

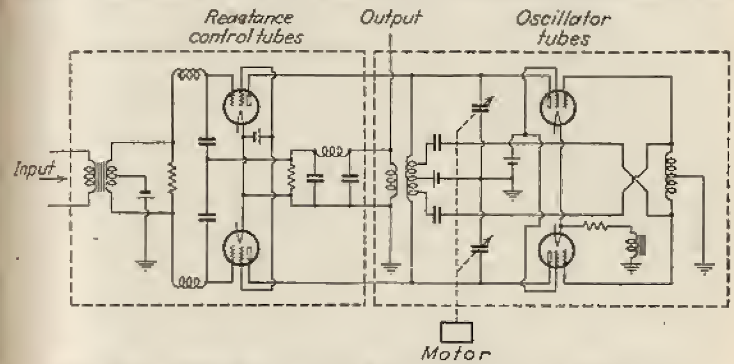


FIG. 66a.—Essential circuits of Western Electric f-m transmitter.

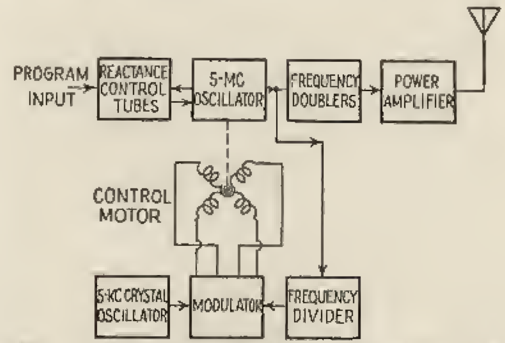


FIG. 66b.—Frequency stabilizing system of Western Electric.

to be within the range of the rotating magnetic field of the electric motor used for the retuning of the oscillator variable capacitors through a suitable speed-reduction mechanism of gear trains. The direction of rotation of the motor depends upon whether the oscillator frequency is higher or lower than the fixed multiple frequency of the standard. Thus automatic readjustment of the oscillator tuning is made in the correct direction so that when exact synchronism occurs between the frequency fed back from the f-m oscillator and the frequency standard the motor is at rest.

Because of the inertia of the motor rotating elements and the high order of frequency division used, the motor is not caused to rotate by frequency deviations produced on the carrier at an audio rate during modulation. The main advantage claimed for the synchronized f-m method over others described for maintaining output frequency stability is that the output frequency is maintained identical in precision to the standard by making all the controlling factors in terms of frequency.

The f-m output signal of the oscillator of the transmitter passes through four pentode stages, three of which are doublers, then through a WE 356-A triode stage and a final WE 357-A triode output stage into the antenna. Operating characteristics of this transmitter are such as are required to transmit faithfully high-quality f-m program signals.

60. Merits of F.M. versus A.M. With the application of f.m. to transmitters operating in the u-h-f band, the relative merit for this system of signal transmission can be evaluated on the basis of an a-m system. The u-h-f signal field intensity[1] at a given distance from a particular transmitting antenna may be determined from theoretical and empirical relationships as published in papers[2] and derived from extensive mathematical and experimental work. Actual experimental tests[3] have shown that an interfering audio signal (output of receiver) will create objectionable interference if its level is about 30 to 40 db below the desired signal. Thus service areas can be defined as zones in which the desired component of the resulting audio signal is at least 35 db above the interference. For very high quality reproduction, this figure runs from 40 to 55 db.

For interfering signals on the same channel as the desired signal, it is evident that, if a.m. is used, a signal input ratio of 35 db is required to secure the desired output ratio. However, in f.m. the ratio of signals at the receiver input needs to be only about 6 db since the receiver for f-m reception responds to frequency variations and limits amplitude variations such as those caused by noise and undesired signals.

On this basis there are claimed advantages of f.m. over a.m. because of (1) improved signal-plus-noise to noise ratio. Experimental results have shown this difference to be as much as 25 db as influenced by intensities of automobile ignition, X rays, and other man-made interference. Atmospheric interference being small at ultrahigh frequencies, it becomes negligible in comparison with man-made interference. (2) A uniform and definite service area from a given transmitter since f-m signal-plus-noise to noise ratio remains high until field intensity reaches a low value. (3) A smaller geographical interference area obtained when two f-m transmitters are operated simultaneously on the same frequency as compared to similar operation of two a-m transmitters. (4) A r-f amplifier used to increase a f-m signal is more efficient than one used for a.m. because f.m. can be accomplished at low level followed by a class C r-f power amplifier. (5) For a given service area, less radiated power is required for f.m. because of the improvement in signal-plus-noise to noise ratio obtained with f.m. (6) For a given power output

[1] TREVOR and CARTER, Notes on Propagation of Waves below Ten Meters, *Proc. I.R.E.*, March, 1933.

[2] DEVINO and HUNT, Ultra Short Wave Propagation over Land Burrows, *Proc. I.R.E.*, December, 1935.

[3] WEIR, I. R., Field Tests of Frequency and Amplitude Modulation with U-h-f Waves, *Gen. Elec. Rev.*, May, 1939; CROSBY, M. G., The Service Range of Frequency Modulation, *RCA Rev.*, January, 1940.

power-tube operating costs are less because smaller tubes can be used for f.m. for a given power output.

The FCC has assigned 40 channels 200 kc wide for f.m. between 42 to 50 Mc.

61. Frequency-modulated Transmitter Measurements. The measuring equipment is considerably different than is required for an a-m station since there is a variation in frequency of the emitted wave with modulation while the amplitude is kept more or less constant. This is exactly the reverse of a.m. where the carrier is varied in amplitude but maintained at a constant frequency. The frequency swing or deviation can be measured by applying sustained tone to the transmitter and then measuring the relative intensities of the carrier and the side frequencies present, the relative amplitudes of which correspond to the Bessel functions involved.[1]

62. Air- and Water-cooled Tubes. For tubes of low power, artificial cooling during operation is usually not necessary, radiation into the air being sufficient. For the larger tubes, however, artificial cooling is usually accomplished by means of a circulating water system which causes a sheet of water to pass over the anode surface at very high velocity.

To restrict leakage of current from the anodes to the grounded pipes of the water system, connection is made between the anodes and the water system through a long length of coiled hose or porcelain tubing

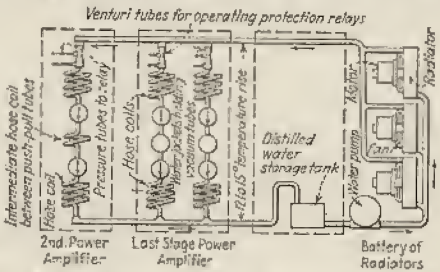


FIG. 67.—Water-cooling and circulation system.

This interposes, between the anode and ground, columns of water long enough to make the electrical resistance to ground very high; as much as 100 ft. of coiled hose may be used, giving resistances of 0.5 up to several megohms.

In many cases distilled water is used, the water being maintained at a satisfactory temperature by an artificial cooler, since for economical reasons it is desirable that the same water be used indefinitely.

The water-cooling and circulating system is automatically started when the transmitter is turned on, and the transmitter is automatically turned off in the event of any failure in the water-cooling system. One method of doing this is shown in Fig. 67, where the water system contains a Venturi tube whose inlet and output orifices are connected to a device operated by the difference in pressure established between the two orifices by the flow of water. If the flow is interrupted or falls below its normal

[1] CROSBY, M. G., A Method of Measuring Frequency Deviation, *RCA Rev.*, April, 1940.

value, a contactor through additional relays causes the power supply t be disconnected.

Sometimes a milliammeter is provided on the transmitter panel wh indicates the magnitude of the current leaking through one of the clos coils, the amount of current serving to indicate the relative purity of tl water and indicating when it is advisable to change the water supply.

In place of water cooling, forced air cooling is also used on some larg tubes. For the large dissipation required, a large number of radiat fins are made a part of a copper radiator attached to the copper anod Sufficient air is forced upward and between the cooling fins to carry awa the heat developed on the anode. Because of the high electrostat capacity created by these anodes, they are not used on the very hig frequencies.

63. Power Supply. Plate-voltage supply for transmitters may b obtained from d-c generators, high-vacuum tube rectifiers, mercury-ar rectifiers, or hot-cathode mercury-vapor rectifiers.

The hot-cathode mercury-vapor rectifier is considered the best metho of supplying high voltages to transmitter plate circuits. The mo striking difference between mercury-vapor tubes and high-vacuum tube is the internal voltage drop between plate and cathode. In the high vacuum tube the voltage drop may vary from a few volts to severa thousand volts, depending upon the current, element spacing, etc. In the mercury-vapor tube the space charge is limited by the arc drop of th vapor which is practically constant at values between 12 and 17 volt regardless of the current.

Table II gives a direct comparison of the relative efficiency of a high vacuum tube and two types of mercury-vapor tube. Note that the mercury-vapor tubes give very low internal voltage drop and have con siderably higher efficiencies.

There are two fundamental limits which determine the power outpu that can be obtained from any number of tubes operated in any type o circuit. These ratings are (1) the maximum peak inverse voltage a which the tube can operate without flashing back and (2) the maximum peak plate current which the cathode can supply with a reasonably long life.

The maximum peak inverse voltage which can exist across a tube in any of the usual types of circuits is equal to the line-to-line peak or cre

TABLE II.—COMPARISON OF HIGH-VACUUM AND MERCURY-VAPOR TYP RECTIFIERS*

No. of tubes	Tube type	Circuit	D-c output			Tube drop		Losses, kilowatts		Efficiency, per cent
			Volts	Amperes	Kilowatts	Volts	At amperes	Filament	Tube-drop	
6	UV-214	3φdouble Y	15,000	12	180	1,560	6	6.9	18.7	87.[illegible]
6	UV-857	3φfull wave	15,000	12	180	15	12	1.5	0.36	98.[illegible]
†6	UV-857	3φfull wave	21,000	30	630	15	30	1.5	0.9	99.4

* I. R. E., Vol. 18, No. 1, January, 1930.
† Maximum rating.

voltage of the power transformer less the voltage drop of the conducting tube.

The peak plate current depends upon the type of circuit, tube, filter, and load. In a single-phase full-wave circuit each tube must carry the full-load current for half the time. In the three-phase half- and full-wave circuit each tube carries the load current for one-third of the time. If the rectifier feeds into an inductance, square blocks of current are drawn from the rectifier and the peak plate current approaches the d-c value. If the rectifier feeds into a capacity load plate current is drawn for only a part of each half cycle and the peak current may reach values of from three to five times that of the d-c load current.

Table III gives data on several typical hot-cathode mercury-vapor tubes designed for radio power supply purposes. The circuits most commonly used with these types of tubes are shown in Fig. 68. The single-phase full-wave and the three-phase and half-wave circuits are quite generally used. The three-phase full-wave circuit is particularly applicable to the half-wave mercury-vapor tube, since it gives a peak inverse voltage whose magnitude is only 4.5 per cent greater than the average output voltage; the wave form is that of a six-phase rectifier.

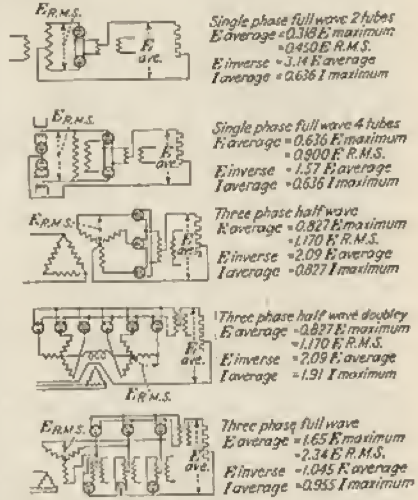


FIG. 68.—Hot-cathode mercury-vapor power circuits.

TABLE III.—HOT-CATHODE MERCURY-VAPOR TUBE RATINGS

Tube type	Filament		Peak inverse voltage	Peak anode current, amperes
	Volts	Amperes		
UX-866	2.5	5	7,500	1.0
UV-872	5	10	7,500	5.0
UV-869A	5	18	20,000	10.0
UV-857B	5	30	22,000	40.0

64. Parasitic Oscillations. One of the most important design features of a transmitter is to provide for adequate suppression of parasitic oscillations. Such spurious oscillations are usually caused by regeneration in an amplifier stage. They have frequencies different from the fundamental or its harmonics.

All classes of amplifiers are subject to these oscillations. Suppressing them in a class C amplifier is not usually so difficult as in the class B type where the grids of the tubes are driven positive for a considerable portion of the cycle. Before reliable and economical service can be realized from a transmitter of any type, all tendencies for parasitic oscillation must be suppressed to prevent serious lessening in the life of vacuum tubes or program interruptions because of arc-overs in the transmitter. Such oscillations may exist in an otherwise normal amplifier stage and may not be evident to casual inspection owing to their disappearance entirely when grid excitation is removed.

A typical class B power amplifier stage of the push-pull type is shown in Fig. 69. This amplifier contains inherent design features which have a tendency to suppress spurious oscillations. C_6 and C_7 assist by acting as a very low reactance path for all parasitics of a frequency higher than the

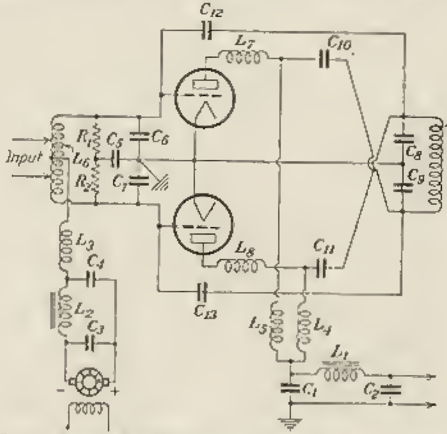


FIG. 69.—Class B amplifier with characteristics to suppress parasitics.

fundamental with a result that they effectively load the parasitic circuit. Connections between these capacitors and the tube grids are kept at an absolute minimum. The grid loading resistors R_1 and R_2, whose real purpose is to improve the regulation of the grid circuit as the grids swing positive, also act as a resistor load to damp out oscillations. C_8 and C_9, with their mid-point grounded, act as a low reactance path to ground for frequencies above the fundamental.

The frequency of parasitic oscillations may be anything from the very low end of the frequency spectrum to the u-h-f region. Parasitics of very low frequencies, in the neighborhood of less than 1 to 10 cycles, are sometimes set up by the dynatron action of the tubes at the natural period of the power-supply filter circuit C_1, C_2, and L_1.

The existence of these parasitics of very low frequencies usually becomes apparent in the form of a severe irregularity in the saturation curve of the linear amplifier. Such a curve is shown in Fig. 70. The point X shows the beginning of this parasitic condition and Y the point where it ceases. It is caused by the dynatron characteristics of the amplifier tube grids and occurs at a point on their operating characteristic just before they are driven positive.

A solution for such a condition is to use tubes whose amplification factor is such that the region XY falls below the carrier operating point. For this reason high-mu tubes have on some occasions been found to be more satisfactory than low-mu tubes.

Low-frequency oscillations of approximately one-third to one-fifth of the fundamental frequency are sometimes caused by tuned-grid tuned-plate

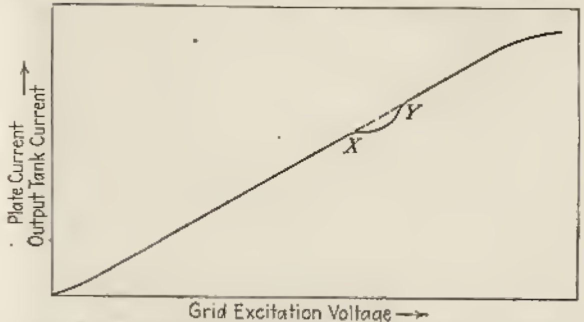


FIG. 70.—Typical saturation curve of class B r-f linear amplifier showing dynatron effect of power tube grids with E_b and E_c constant.

regeneration with the plate chokes L_5 and L_4 in combination with the blocking condensers C_{10} and C_{11} forming an output tank circuit. A similar grid tank circuit is formed by C_6, C_7, and L_3. Inasmuch as all tubes are effectively in parallel for this combination, the neutralizing capacitors tend to aggravate the condition rather than to prevent it. In Fig. 71 is shown an equivalent parasitic circuit of the combination as formed from the circuit in Fig. 69. The remedy is to change the values of inductance and capacity in either the parasitic grid or plate circuits so as to cause their natural periods to depart substantially from a near resonance condition. It is usually possible to suppress such oscillation by tuning the parasitic grid circuit to a higher frequency than the corresponding plate circuit.

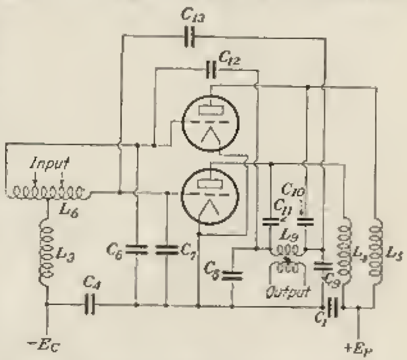


FIG. 71.—Equivalent parasitic circuit of Fig. 69.

The existence of these oscillations may usually be detected by applying excitation at the fundamental frequency to a stage with reduced plate voltage and grid-bias voltage until the tubes draw plate current. If oscillation of the stage continues after fundamental grid excitation is removed, as indicated by neon lamps attached to the tube plates, the frequency of the parasitic may be determined by means of a wavemeter, and thus steps can be taken to eliminate it.

Oscillations within an amplifier stage at frequencies near the fundamental are usually caused by regeneration within an amplifier stage due to improper neutralization causing tuned-grid tuned-plate circuit oscillations. Improper circuit design or too close coupling between the inductances of the input and output circuits or chokes is also liable to cause this condition.

Parasitics of frequencies in the neighborhood of from five to twenty times the fundamental result in cases where the leads from the tube grids and C_7 and C_7 form a grid tank circuit, the resonance frequency of which is determined by various distributed capacities and the inductance of the leads. Oscillations are made possible by the existence of a similar plate tank circuit formed by leads from the tube plates to C_3 and C_4 together with various stray capacities. This form of parasitic is seldom sustained but shows itself most prominently when the stage is subject to high peaks of modulation. The trouble may usually be corrected by insertion in the plate leads at a point adjacent to the tube plates choke coils L_7 and L_8.

These parasitic choke coils L_7 and L_8 together with a shortening of grid leads to an absolute minimum may also assist in suppressing oscillations at ultrahigh frequencies in amplifier stages employing two tubes in parallel. The grid leads of the two tubes, although connected, may combine with stray capacities, thus forming a push-pull oscillation of a very high frequency. Such oscillations in some cases cause high r-f voltages to build up which may result in serious arc-overs from various parts of the tube output circuits.

65. Suppression of R-f Harmonics. It is the inherent characteristic of a vacuum tube, while functioning at a reasonably high efficiency in an amplifier circuit, to generate harmonic frequencies of the fundamental. A station broadcasting on 600 kc, if second and third harmonics were not suppressed, would produce interference with other stations operating on 1,200 and 1,800 kc. Field intensity measurements about a station are necessary to determine how much harmonic energy is radiated and to show the progress of work done toward reducing radiation.

In specifying the allowable harmonic radiation from a broadcasting station the IRE Committee on Broadcasting as of January, 1930, recommended that the maximum radio field intensity of a harmonic component measured at a distance of 1 mile from a station should not exceed 0.05 per cent of the field intensity of the fundamental.

A field strength of 500 μv per meter at a distance of 1 mile is recommended as a maximum allowable intensity from a high-powered transmitting station. If in the case of a 50-kw station a circular-field pattern and equal attenuation are assumed for both a harmonic and the fundamental in the immediate vicinity of the station, a field strength of 500 μv at 1 mile would correspond to approximately 7 mw of radiated power at a harmonic frequency. The effect of directivity (illustrated in curve B, Fig. 72) may cause a field intensity of a number of times the value of 500 μv to be projected in a given direction with a very small fraction of 1 watt of harmonic power in the transmission line and antenna circuits. Such a concentration of radiated power may form very objectionable interference. Considering the factors involved, therefore, it is evident that harmonic suppression must be attacked from a number of angles. These may be briefly outlined as follows:

1. Design of the transmitter circuits to reduce the harmonic content of the power delivered to the antenna circuits to a minimum.
2. Thorough and effective shielding of the entire transmitter or building.
3. Effectively grounding all harmonic drain circuits and elimination of loose conductors near the transmitter coupled to it inductively or capacitatively.
4. Reduction of directivity of harmonic radiation to a minimum.
5. Installation of shielded band- or low-pass filters at the input end of the transmission line to the antenna.

Some commonly used triode amplifier circuits are shown in Figs. 73 and 74. The push-pull amplifier is superior to the single-ended circuit, as it is capable of producing a sum plate current of the two tubes which is

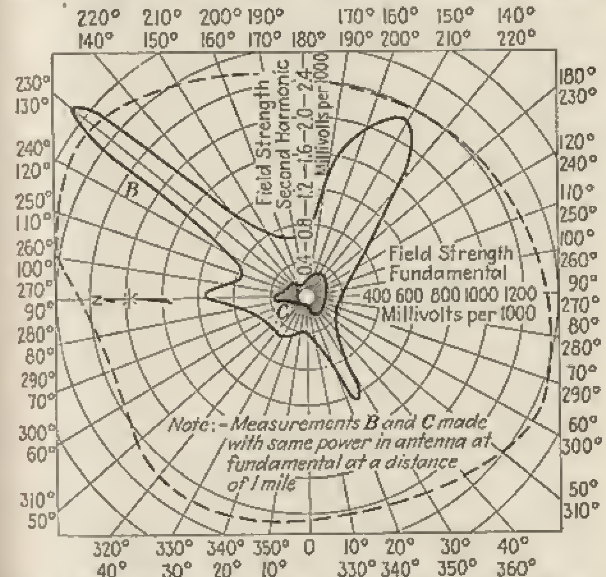


FIG. 72.—Radio field-intensity survey. The dotted curve gives fundamental frequency field strength; B and C are second harmonic intensity before and after reduction.

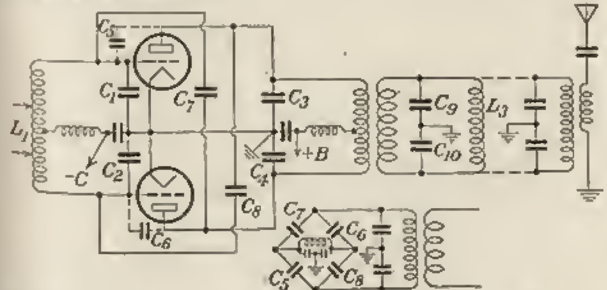


FIG. 73.—Push-pull amplifier with high kilovolt-ampere tank circuit in transmission line.

symmetrical in wave shape and, therefore, it contains no even harmonics. Individual plate currents, of course, contain even harmonics which are drained to ground through C_3 and C_4 resulting in identical instantaneous even harmonic potentials being set up on each side of L_2 but no net even harmonic current through it. Under these conditions an electrostatically shielded inductive coupling is provided to permit transfer only fundamental and odd harmonic frequencies to the coupled circ. For a condition of symmetrical plate current it is evident that the t characteristics must match closely, $C_1 = C_2$ and $C_3 = C_4$. The neutralizing bridge must be balanced not only for the fundamental freque but for even harmonics. This requires that the internal capacities of tubes should match. As will be shown later, a high ratio of circulati kilovolt-ampere in the tank circuit to the kilowatt delivered from amplifier reduces the output of harmonics from a single-ended ampl to a very low value. This is also true in the push-pull circuit.

The circuit shown in Fig. 74 will give a very small amount of harmo output by proper design of the circuit constants. The curves in Fig.

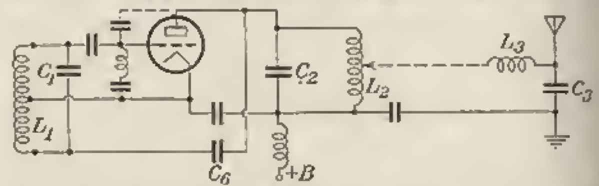


FIG. 74.—Line termination effecting reduced harmonic radiation.

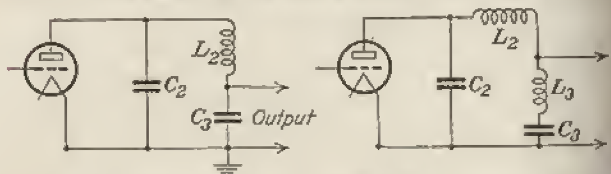


FIG. 75.—Improved tank circuits for suppressing harmonic radiation

show the filtering effect of a high kilovolt-ampere tank circuit in s pressing harmonic components of current generated in the tube. Th curves show actual harmonic transferred to a given load circuit Z_L w a constant output at the fundamental and various kilovolt-ampere kilowatt ratios of L_2 and C_2. Figure 75 shows improvement in ta circuits so as to increase the normal filtering action of an ordinary ta circuit. A high kilovolt-ampere to kilowatt ratio applied to these circu is capable of reducing harmonic output to an extremely small amou There are some limitations in the amount of filtering which can be secur by a high kilovolt-ampere tank circuit, however, since the I^2R losses the circuit increase in proportion to the circulating kilovolt-ampe and the cost of apparatus for increasing kilovolt-amperes in a cir without increasing losses is considerable. In broadcasting transmitt there is the limitation of too low a decrement in a circuit attenuating greatly the high frequencies of a modulated envelope. In Fig. 75 t trap L_3C_3 is tuned to a particular harmonic to be eliminated. The u antiresonant circuits (parallel traps) in the plate lead of an amplifi while reducing to some extent a single harmonic, has a tendency to all

considerable voltage to build up at others. Most satisfactory results are usually secured by designing a minimum impedance path for harmonics to ground as compared with a given high impedance at the fundamental.

The effectiveness of the shielding of a transmitter may be determined by operating the transmitter with full power output into a shielded phantom antenna. Measurement of the harmonic field strengths produced from the transmitter itself is direct evidence of how well it is shielded. Such radiation can usually be traced to a long conductor near the transmitter, coupled to it capacitively or through a common ground return. Ground conductors serving to drain harmonic frequency power to ground therefore should be as direct as possible and should not be extended so as to have a free end which might attain a high potential at resonant frequencies. This is particularly true of the harmonic drains

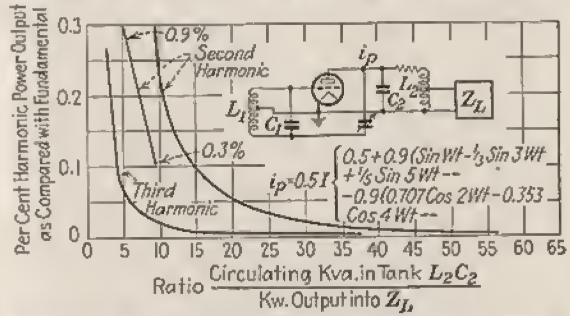


FIG. 76.—Effectiveness of high kilovolt-ampere to kilowatt ratio in reducing harmonic output with constant power output at fundamental.

near the antenna itself. These should have a separate ground to prevent coupling of harmonic frequencies into the antenna.

A sensitive wavemeter is very useful in determining the relative harmonic field intensities near the various circuits of a transmitter. When tuned to the frequencies of various harmonics and coupled to various circuits of the transmitter or placed at positions along near-by open conductors, this instrument will indicate proportionate amounts of the harmonic components of the current flow. By effectively grounding a long open conductor, either directly or through large capacities at a number of distributed points, harmonic radiation can usually be eliminated.

The push-pull amplifier coupled to a long transmission line has often become a source of undesirable even-harmonic radiation because of sufficient electrostatic capacity existing between the coupled circuits to permit a transfer of energy from the amplifier output circuit to the line. Unless this electrostatic capacity is reduced to an extremely low value, i.e., by installation of a well-grounded electrostatic screen between the two coils, even harmonics usually find a path along the transmission line with a ground return to the generating source. An unshielded transmission line serves in this case as an effective directive radiator in the form of a large loop. Its effective height will be dependent upon the height of the transmission line above ground. Parallel flow of even-harmonic currents along the line, therefore, makes it a much more effective radiator in some directions than the push-pull flow of harmonic currents in the line.

A circuit which has been found to be very effective in reducing both the parallel as well as the push-pull flow of harmonic currents in a transmission line is shown in Fig. 73 in the form of a high kilovolt-ampere floating tank circuit $L_3C_9C_{10}$ tuned to the fundamental component of current flowing in the line. This tank circuit, while offering an impedance to the fundamental approaching an infinitely high value, offers a relatively low impedance path to ground for the parallel flow of even harmonics equivalent to

$$Z_{ne} = \frac{-1}{4\pi f_{ne}C_9} = \frac{-1}{4\pi f_{ne}C_{10}}$$

where resistance of circuit is negligible
Z_{ne} = impedance to nth even harmonic
f_{ne} = frequency of nth even harmonic
and for the push-pull flow of odd harmonics between transmission-line conductors

$$Z_{no} = \frac{-2\pi f_{no}L_3}{(2\pi f_{no})^2L_3C - 1}$$

where resistance of circuit is negligible
Z_{no} = impedance to nth odd harmonic
f_{no} = frequency of nth odd harmonic

$$C = \frac{C_9}{2} = \frac{C_{10}}{2}$$

where $C_9 = C_{10}$.

It is evident that as C_9 and C_{10} are increased in capacity the effectiveness of the circuit in reducing harmonics is increased. Since the transmission-line termination impedance is usually made to match the line impedance for the fundamental frequency, it usually happens that the line impedance is matched

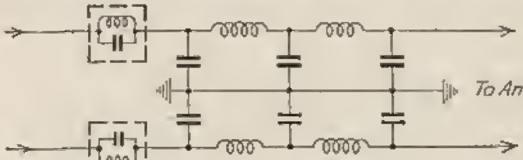


FIG. 77.—Low-pass filter combined with antiresonant circuits in transmission line.

for this frequency only and as a result harmonic components of current and voltage in the line appear as standing waves along the line. In such a case the above tank circuit is most effective for eliminating a particular harmonic if it is placed at a point along the line of maximum voltage. This circuit alone was effective in one case in reducing second-harmonic radiation from a station to one-fifth of its former value.

Antiresonant circuits installed in a transmission line at current antinodes have been found very effective in reducing a single harmonic to which they were tuned. Extreme care should be taken in shielding these antiresonant circuits to secure best results. A combination of antiresonant circuits and a low-pass filter is shown in Fig. 77. This combination has been used successfully in severe cases of harmonic radiation from a very long transmission line and antenna system. The filter matches the surge impedance of the line and has a cutoff frequency between the fundamental and second harmonic. Antiresonant circuits have been found useful to sharpen the cutoff

so as to attenuate sufficiently the second-harmonic frequency. Considerable experience in filter design and adjustment is required to secure optimum results from such an arrangement. For use with concentric lines with the outer sheath grounded, the filter shown in Fig. 77 is simplified to the extent of one-half, *i.e.*, one line to ground.

The methods of line termination shown in Figs. 73 and 74 are effective in reducing the possibility of harmonics reaching the antenna circuit. The termination shown in Fig. 74 may be improved by use of a multisection low-pass filter.

66. Antenna Circuit Terminations for R-f Transmission Lines. Considerable improvement in antenna efficiency can be secured from an antenna located at some distance from the station so as to approach the ideal case of an antenna radiating in free space. The r-f transmission line is used for conveying the energy from the transmitter to the antenna. A simple form of such a transmission line is the parallel two-conductor type, each conductor having a diameter of approximately ½ in. The spacing of the conductors is normally 12 to 15 in

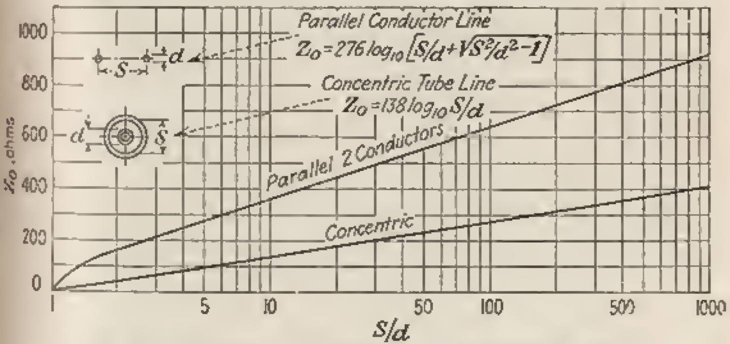


FIG. 78.—Impedance of parallel and concentric-tube lines.

The curves of Fig. 78 show the characteristic impedance values with respect to spacing and conductor size of both the parallel conductor line and the concentric-tube type.

67. Transmission-line Calculations. There are diverse methods of measuring the characteristic impedance of a transmission line. A simple but effective method is illustrated in Fig. 79. With the setup shown and the switch thrown to the line position, a trial value of resistance R_2 is inserted. C is adjusted for maximum I_2. Then with switch thrown in the opposite position and R_1 set to equal R_2, the capacitor C is adjusted for maximum I_1. By trial, a combination may be found where there is a maximum value of I_1 and I_2 for the same setting of C with R_1 equal to R_2. This value of R is the characteristic or surge impedance of the line.

When r-f power is transmitted over a transmission line to an antenna load, the line termination may be adjusted to afford a condition where there are no wave reflections by making the effective resistance of the termination equal to the characteristic impedance of the line. Several

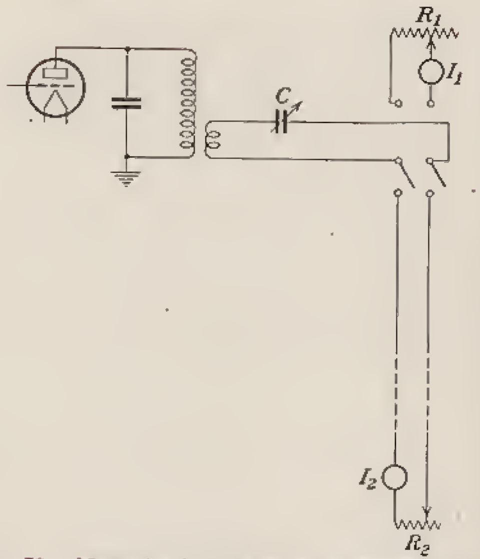


FIG. 79.—Measuring impedance of transmission line.

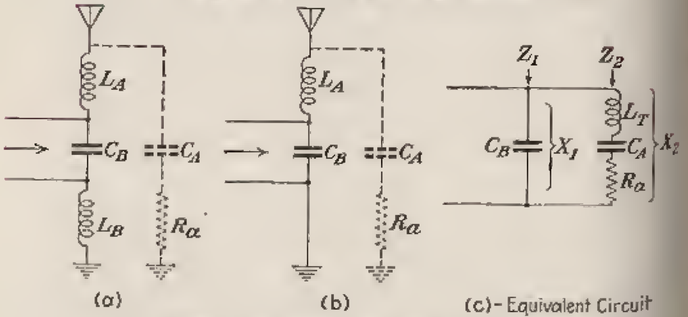


FIG. 80.—Terminations for transmission lines.

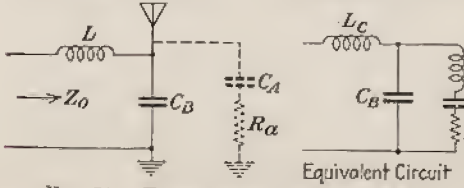


FIG. 81.—Transmission-line termination.

circuits used for terminating transmission lines are shown in Figs. 80 to 82 together with their equivalent circuits.

A formula for calculating the value of capacitor C_B for an effective resistance value Z_0 equal to the characteristic impedance of a two-conductor transmission line balanced to ground as shown in Fig. 82 as well as for a transmission line having one conductor grounded is as follows:

Let Z_0 = effective resistance of transmission-line termination
R_a = antenna resistance consisting of radiation resistance plus equivalent loss resistance
L_T = combined inductance-balance coils plus equivalent antenna inductance
C_A = equivalent antenna capacity
C_B = line-termination capacity
X_1 = reactance of C_B
X_2 = reactance of $X_L - X_{CA}$
Z_1 = impedance branch 1 $= -jX_1$
Z_2 = impedance branch 2 $= R_a + jX_2$

$$Z_0 = \frac{R_a X_1^2 - j(X_2^2 - X_1X_2 + R_a^2)}{R_a^2 + (X_2 - X_1)^2}$$

$$C_B = \sqrt{\frac{Z_0 - R_a}{4\pi^2 f^2 Z_0^2 R_a}}$$

where $Z_0 > R_a$

$$X_1 = \frac{Z_0 R_a}{\pm\sqrt{R_a(Z_0 - R_a)}} = \frac{1}{2\pi f C_B}$$

where, in Fig. 82, C_B is dependent only on values of Z_0 and R_a where Z_0 is equivalent to a pure a-c resistance with the antenna circuit adjusted for resonance. *Unless Z_0 exceeds the value of R_a an effective resistance equivalent to the characteristic impedance of the line cannot be secured.*

When low-impedance lines are used, such as the concentric-tube type, the termination shown in Fig. 81 is useful, since it affords a condition where correct termination may occur in the form of an effective resistance even though R_a equals or exceeds Z_0.

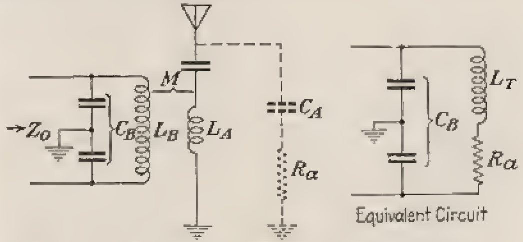


FIG. 82.—Balanced transmission-line termination.

In Fig. 82 is shown a transmission line in the form of a tank circuit. The tank condenser C_B across the line is selected so as to provide a suitable kilovolt-ampere ratio of the tank circuit with respect to the kilowatts transferred to the antenna circuit; this kilovolt-ampere to kilowatt ratio is normally about 10 and should never be less than 2.

$$X_1 = \frac{Z_0 R_e}{\pm\sqrt{R_e(Z_0 - R_e)}}$$

$$X_1^2 = \frac{Z_0^2 R_e}{(Z_0 - R_e)}$$

from which

$$R_e = \frac{X_1^2 Z_0}{Z_0 + X_1^2}$$

where R_e is the effective value of resistance reflected into the tank circuit from the antenna circuit.

The value of R_e can be calculated from

$$R_e = \frac{\omega^2 M^2 R_a}{R_a^2 + X_a^2}$$

where the inherent resistance of the tank circuit is negligible

M = mutual inductance between L_A and L_B

X_a = reactance of antenna circuit

R_a = resistance of antenna circuit.

For a condition of proper termination X_a approaches zero and may be neglected and

$$\frac{\omega^2 M^2}{R_a} = \frac{X_1^2 Z_0}{Z_0^2 + X_1^2}$$

$$M = \sqrt{\frac{X_1^2 Z_0 R_a}{\omega^2(Z_0^2 + X_1^2)}}$$

In Fig. 83 are shown values of M required for a transmission impedance of 400, 500, and 600 ohms and a line-termination capacitor of between 0.001 and 0.004 μf. The transmitter frequency was assumed as 670 kc and the antenna resistance 30, 70, and 140 ohms. In the design of a tank-circuit termination for a given line the value of C_B across the line is selected so as to provide a proper kilovolt-ampere in the tank circuit with respect to the power transferred to the antenna circuit. This kilovolt-ampere to kilowatt ratio is normally about 10.

68. Termination Adjustments. The usual procedure in adjusting a transmission-line termination for a condition of no-wave reflection on the line is as follows:

1. The number of coupling turns is calculated so as to give the proper value of M. With the tank circuit open, the antenna is tuned to exact resonance by means of an external oscillator loosely coupled to it at the fundamental frequency.

2. The tank circuit is now connected into the circuit and tuned to resonance. This is indicated by a condition where the current in the antenna circuit becomes a minimum.

3. The transmission line is then connected across the tank circuit without making any changes in previous adjustments.

4. Correct termination may be checked by measuring the transmission-line currents at the ends and quarter-wave points along the line by means of suitable meters. When proper termination has been effected, the transmission-line currents will be identical at all points along the line.

69. Concentric Line Terminations.[1] The growing use of concentric lines of the low-impedance type has led to cases where the characteristic impedance of the transmission line is *lower* than that of the antenna resistance. In general there are three cases to consider as follows: (1) when the antenna impedance contains a resistance component only; (2) when the antenna impedance contains a resistance component and a reactive component, either (a) capacitive or (b) inductive; and (3) when the antenna impedance contains resistive and reactive components, the latter being partially compensated by the insertion of an extra reactance of opposite sign. These three cases are considered in order.

[1] This and the following article are from *Electronics*, December, 1936.

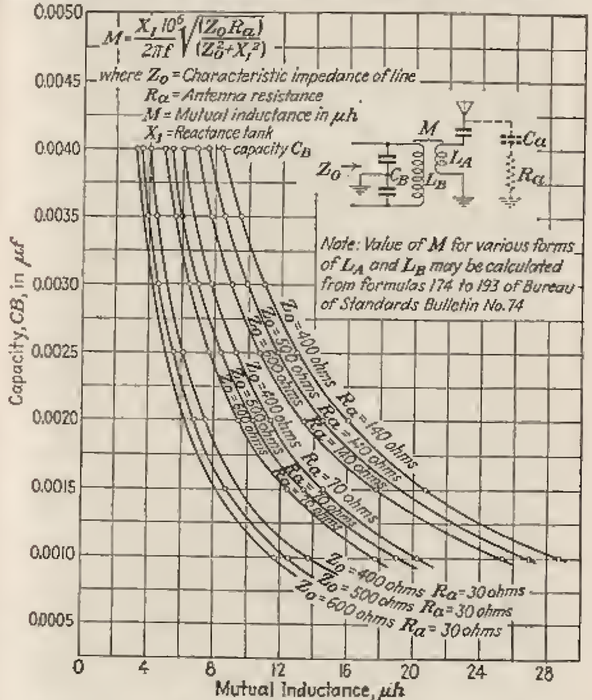


Fig. 83.—Values of M required for proper termination.

Case 1. Antenna Impedance Purely Resistive. From Fig. 84 the concentric line characteristic impedance, Z_0, is terminated by a network consisting of C_B, L_C, and the antenna impedance Z_A. For case 1 the reactance of the antenna impedance is zero, and $Z_0 < Z_A = R_A$. Then the complex impedance Z_L presented to the end of the transmission line is as follows:

$$Z_L = \frac{R_A[X_3 X_1 - X_1(X_3 - X_1)] + j[X_3 X_1^2 + R_A(X_3 - X_1)]}{R_A^2 + X_1^2}$$

where R_A, X_1, and X_3 are as given in Fig. 84. For proper termination Z_0 must equal Z_L. X_1 becomes

$$X_1 = R_A\sqrt{\frac{Z_0}{R_A - Z_0}}$$

$$X_3 = \frac{R_A^2 X_1}{X_1^2 + R_A^2}$$

Since $C_B = 1/(2\pi f X_1)$ and $L_C = X_3/(2\pi f)$, their values in microfarads and microhenrys are then readily calculable from f, the frequency of operation. Figure 85 gives various values of X_1 and X_3 in terms of values of Z_0 and R_A.

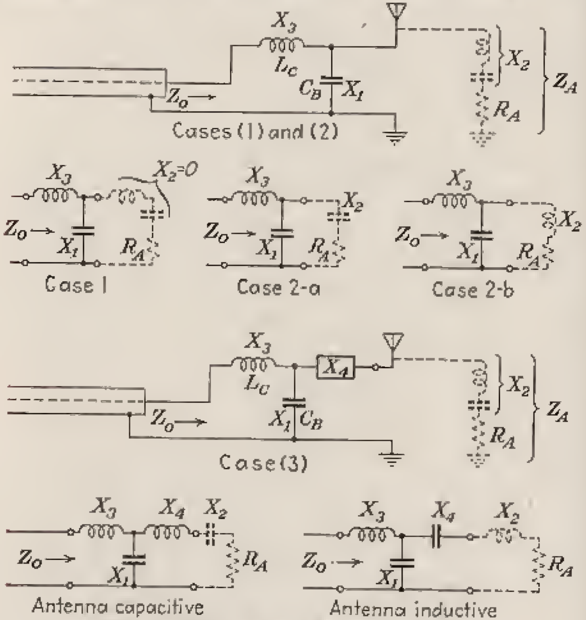


Fig. 84.—Concentric line terminations.

Case 2a. Antenna Impedance with Capacitive Reactance. Refer again to Fig. 84. It will be noted that the equivalent diagram for case 2a is the same as for case 1, except that the antenna impedance is now $Z_A = R_A - jX_2$. Then

$$Z_L = Z_0$$

$$Z_0 = \frac{R_A X_1^2}{R_A^2 + (X_2 + X_1)}$$

from which

$$X_1 = \frac{Z_0}{R_A - Z_0}\left[X_2 \pm \sqrt{\frac{R_A}{Z_0}(R_A^2 + X_2^2 - Z_0 R_A)}\right]$$

$$X_3 = \frac{X_1(R_A^2 + X_2 X_1 + X_2^2)}{R_A^2 + (X_1 + X_2)^2}$$

Hence, with Z_0, R_A, and X_2 given, X_1 and X_3 can be calculated. From the values of X_1 and X_3, L_C and C_B can be calculated, exactly as in case 1. Values of X_1 and X_3 for various values of R_A and values of X_2 for the cases where Z_0 is 80 and 100 ohms are given in Fig. 86.

Case 2b. Antenna Impedance Inductively Reactive. Case 2b is the same as case 1 except that $Z_A = R_A + jX_2$.

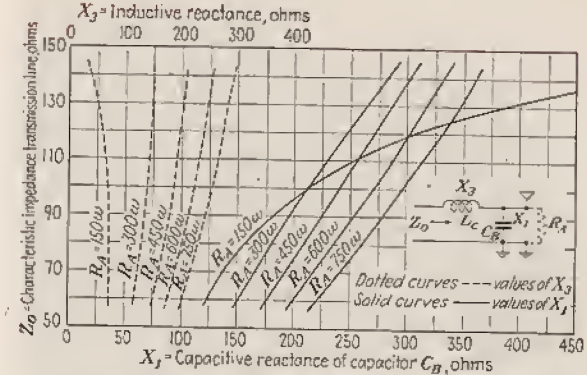


Fig. 85.—Values of reactances for line termination.

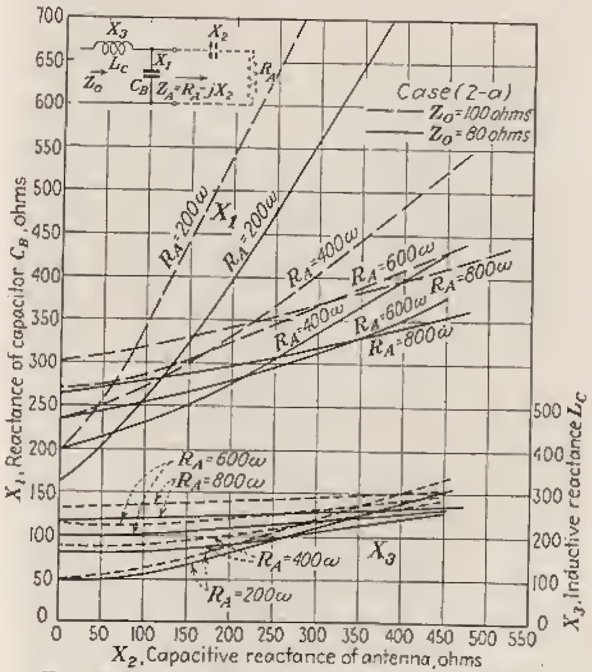


Fig. 86.—Values of terminating reactors for case 2a.

$$X_1 = \frac{Z_0}{R_A - Z_0}\left[-X_2 \pm \sqrt{\frac{R_A}{Z_0}(R_A^2 + X_2^2 - Z_0R_A)}\right]$$

$$X_3 = \frac{X_1(R_A^2 + X_2^2 - X_2X_1)}{R_A^2 + (X_2 - X_1)^2}$$

from which L_C and C_B are calculated. Figure 87 shows various values of X_1 and X_3 in terms of R_A and X_2, for Z_0 values of 80 and 100 ohms.

Case 3. Added Reactance to Antenna Impedance. When the transmission line impedance "looks into" a complex antenna impedance, it is possible to simplify the adjustment of the circuit greatly by adding a reactance X_4 as shown in Fig. 84 for case 3. This reactance X_4 may be either inductive or capacitive, as shown. If the sum of X_3 and X_2 is inductive, then X_4 is made capacitive and vice versa. The value of X_4 is such that the algebraic sum of X_1, X_2, and X_4 is equal to zero. Since X_4 is in series with the antenna impedance, it adds directly with the reactive part of the antenna impedance. The effect of the presence of X_4 can then be taken into account by applying the formulas of case 2a or 2b.

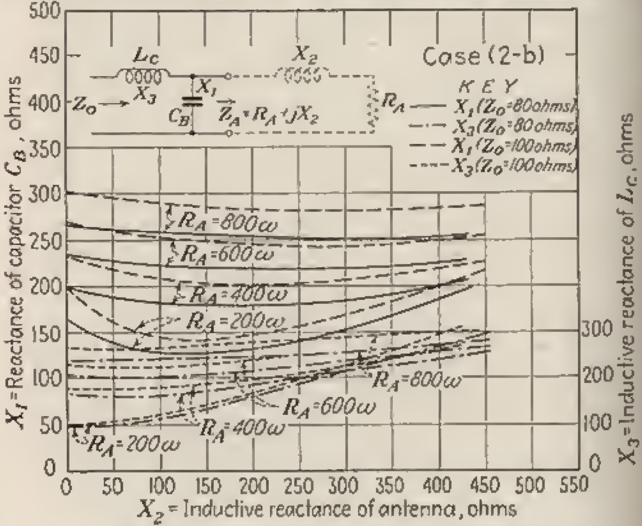


FIG. 87.—Terminating reactors for case 2b.

$$X_1 = X_3 = \sqrt{Z_0R_A}$$

This occurs only, however, if X_4 is so chosen that

$$\pm X_4 - X_1 \mp X_2 = 0$$

The reactance X_4 must always have the opposite sign from X_2, as indicated by the plus-or-minus signs in the equation. When X_4 is so chosen, the reactance X_1 and X_3 may be obtained for various value of Z_0 and R_A by reference to Fig. 88. Note that these values apply regardless of whether R_A is larger than, equal to, or greater than Z_0.

Practical Procedure in Designing Matching Circuits. In making suitable adjustments on the impedance matching circuits to provide a correct

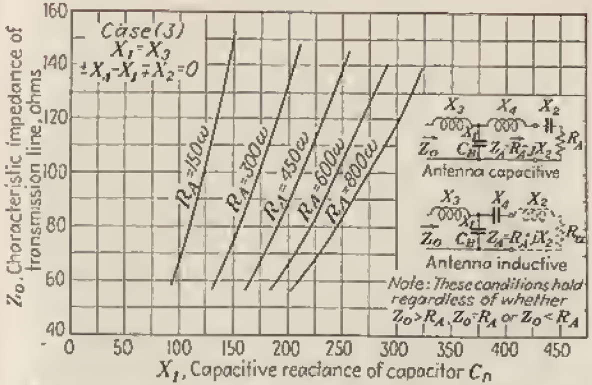
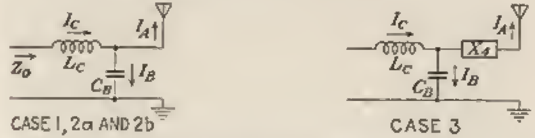


FIG. 88.—Values of terminating reactors for case 3 when antenna reactance is compensated.

CURRENT AND VOLTAGE RELATIONS



CASE	ANTENNA CURRENT I_A	CURRENT THROUGH CAPACITOR C_B = $I_C\left(\frac{Z_A}{Z_A+Z_B}\right)=I_B$	TRANSMISSION LINE CURRENT= $I_A\sqrt{\frac{R_A}{Z_0}}=I_C$	VOLTAGE AT BASE OF ANTENNA $E_A=I_AZ_A$
1	$\sqrt{\frac{W}{R_A}}$	$I_C\left[\frac{R_A^2+jR_AX_1}{R_A^2+X_1^2}\right]$	$\sqrt{\frac{W}{Z_0}}$	I_AR_A
2a	$\sqrt{\frac{W}{R_A}}$	$I_C\left[\frac{R_A^2+X_1X_2+X_2^2+jR_AX_1}{R_A^2+(X_2+X_1)^2}\right]$	$\sqrt{\frac{W}{Z_0}}$	$I_A(R_A-jX_2)$
2b	$\sqrt{\frac{W}{R_A}}$	$I_C\left[\frac{R_A^2-X_2X_1+X_2^2+jR_AX_1}{R_A^2+(X_2-X_1)^2}\right]$	$\sqrt{\frac{W}{Z_0}}$	$I_A(R_A+jX_2)$
3 ANTENNA CAPACITIVE	$\sqrt{\frac{W}{R_A}}$	$I_C\left[\frac{R_A+jX_1}{R_A}\right]$	$\sqrt{\frac{W}{Z_0}}$	$I_A(R_A-jX_2)$
3 ANTENNA INDUCTIVE	$\sqrt{\frac{W}{R_A}}$	$I_C\left[\frac{R_A+jX_1}{R_A}\right]$	$\sqrt{\frac{W}{Z_0}}$	$I_A(R_A+jX_2)$

W is power in watts

FIG. 89.—Current and voltage relations in terminating circuits.

termination for a given transmission-line characteristic impedance, under cases 2a and 2b above, where $R_A > Z_0$, the following procedure is recommended:

1. The transmission-line characteristic impedance should be calculated and the results checked by actual measurements if possible, either by means of a r-f impedance bridge or by the methods described in the literature.

2. The antenna base resistance should be measured over a frequency band width covering at least 100 kc each side of the operating frequency. A curve should then be constructed with values of antenna resistance as a function of frequency. A smooth curve drawn through the points of measurements will assist in checking their accuracy.

3. Together with antenna resistance measurements, the antenna reactance should be measured, either by means of a r-f impedance bridge or in a manner shown in Fig. 90 over a wide frequency range and a curve constructed with antenna reactance as a function of frequency.

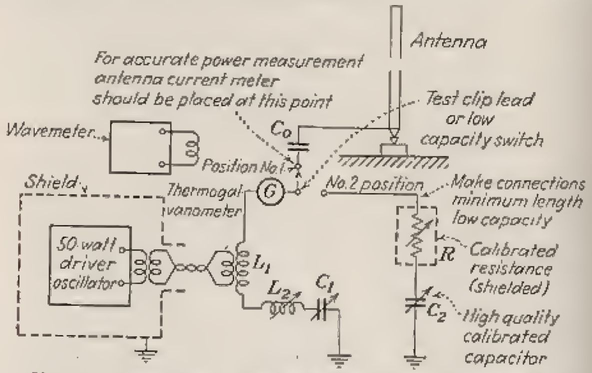


FIG. 90.—Setup for measuring antenna characteristics.

4. With the values of antenna resistance and reactance known, values of capacitance C_B and inductance L_C may be calculated for case 2a or 2b, as may be required, and connected into the circuits as shown in Fig. 84.

5. With the transmission line connected, correct termination may be checked by measuring the transmission-line currents at the ends, if its length is equal to a quarter wave length or odd multiplies thereof. For a very long line it is good practice to make these measurements at a number of points along the line. The existence of stationary waves of current or voltage of the fundamental frequency along the line is an indication of incorrect termination. In such a case slight adjustments may be necessary in L_C and C_B to correct for stray capacity of leads and tuning equipment or slight errors in measurements. If a r-f impedance bridge is available, its measuring terminals may be connected across the input to the matching circuit in place of the transmission line and the termination circuit checked for an effective resistance equivalent to the characteristic impedance of the line without the line attached.

Although case 3 requires the addition of another piece of apparatus in the form of an inductance or capacity in the antenna lead, which may be rather expensive, the adjustment procedure is less difficult and is as follows:

1. With values of the line characteristic impedance, antenna resistance and reactance obtained by measurement, the value of C_B is calculated, which gives the reactance X_1 necessary.

2. With L_C disconnected from C_B, reactance X_4 (inductive or capacitive) is added in the antenna circuit in series with X_1. By means of X_4 the antenna circuit is tuned to resonance, as indicated by maximum current through a thermogalvanometer, when the antenna circuit is excited by means of an external oscillator loosely coupled to it.

3. A sufficient value of inductance L_C having a value X_3 equal to X_1 is then connected into the circuit as shown in Fig. 84.

4. The line is then checked for stationary waves, the absence of which indicates a condition of correct termination.

The mechanical properties of long concentric-tube transmission lines makes the measurement of current in the center conductor rather difficult. In some cases removable plugs are placed in the outside tube at various intervals along the line. These plugs, which, when inserted, make the outer tube airtight, permit connections from an antiresonant circuit across the line. Such an antiresonant circuit, when tuned to the fundamental frequency, presents a very high impedance to the line, when bridged across it, and therefore does not effect its characteristic impedance at the fundamental frequency. With about 10 watts flowing through the line, the galvanometer reading is an indication of the voltage at the points measured along the line.

70. Method Used in Measuring Antenna Characteristics.

Refer to Fig. 90.

Value of C_0 (usually about 0.0005 μf) is selected to provide sufficient series capacitance reactance to make the antenna capacitive over the frequency range measured. Then, with the antenna excited by the driver oscillator at the frequency indicated by the wavemeter and the switch at position 1, adjust C_1 and L_2 for resonance, as indicated by the maximum reading of G. R is then adjusted until G reading is the same as before. Then R is the antenna resistance.

For antenna reactance measurement, the circuit is first calibrated for stray capacity in the shielded resistance box by resonating circuit (switch in position 2) first with box in the circuit and then entirely removed. Difference in reading of capacitor C_2 between the two conditions equals capacity of box. This value should be added to each reading of C_2, when circuits are resonated, which is done as above for resistance measurement. The antenna reactance X_2 is equal to the reactance of C_2 minus that of C_0. When reactance of C_0 is greater than that of C_2 the antenna reactance is positive.

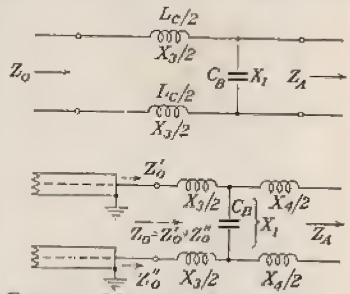


FIG. 91.—Matching circuits for balanced transmission lines.

When it is found desirable to apply the matching circuits described above (Fig. 84) to balanced lines (open wire or double concentric types), the value of X_3 derived by the particular formula for cases 1 and 2 is halved and placed on each side of the circuit (see Fig. 91), while the value of X_4 is halved and placed on each side of the circuit for case 3. Under these conditions the formulas given above apply to the respective cases mentioned. The systems become quite useful in matching a given balanced transmission line or r-f circuit into another having entirely different input impedance characteristics.

In the foregoing analysis of antenna matching circuits, they were considered as providing for a given transmission line, a termination impedance equivalent to an ohmic resistance at the fundamental frequency. An analysis of the input impedance that such a line "looks into" at various harmonic frequencies discloses that it may assume an infinite number of different impedances containing resistance and positive or negative reactance components, the values of which depend upon the termination circuit constants as well as those of the antenna. The values of antenna resistance and reactance may vary widely with frequency. For harmonic frequencies, stationary waves of current and voltage will form on the transmission line as well as in the antenna circuit, unless suitable harmonic filtering is provided either within the vacuum-tube transmitter or at the input to the transmission line.

The effectiveness of a given filter design for various harmonics depends upon its position in the line with respect to the positions of current and voltage antinodes of the harmonic frequencies along the line.

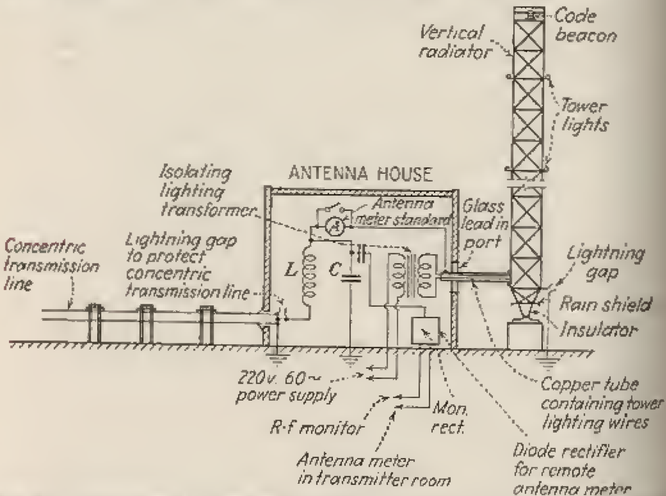


FIG. 92.—Equipment in antenna house of modern broadcast transmitter.

71a. Loss in R-f Transmission Lines. By reference to Fig. 93, it is evident that the most prominent factors contributing to power loss in open wire transmission lines are as follows:

1. Power loss due to conductor thermal resistance

$$R_T = \frac{.1262}{d}\sqrt{\rho\mu f} \text{ ohms per centimeter length}$$

where $S >> d$ (see Fig. 78)

ρ = resistivity of conductors in microhm-centimeters
μ = permeability of conductors
f = frequency in megacycles
d = diameter of conductor in centimeters.

2. Power radiated from balanced and unbalanced line currents.
3. Power component of mutual inductance due to secondary currents induced in near-by conductors.
4. Power loss due to leakage or conductance of the insulating medium or

$$G = \frac{\gamma A}{l} \text{ mho per centimeter length}$$

5. Power loss due to dielectric hysteresis.

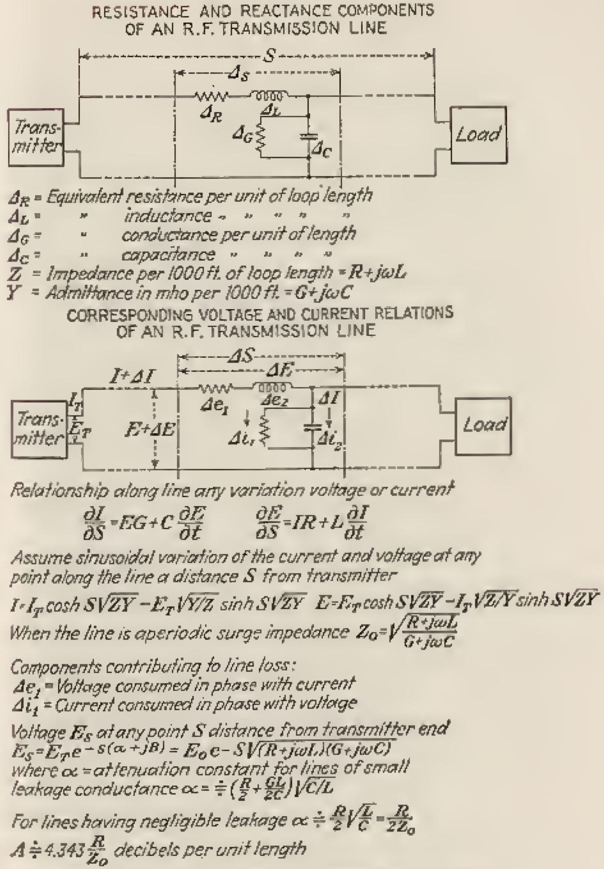


FIG. 93.—Losses in transmission lines.

For coaxial lines the major factors contributing to power loss in transmission lines are as follows:

1. Power loss due to conductor thermal resistance

$$R_T = 0.0631\sqrt{\rho\mu f}\left(\frac{1}{d_1} + \frac{1}{d_2}\right) \text{ ohms per centimeter of line}$$

where d_1 = outside diameter of inner conductor in centimeters
d_2 = inside diameter of outer conductor in centimeters
ρ = resistivity of conductors in microhm-cms
μ = permeability of conductors
f = frequency in megocycles

2. Power loss due to leakage conductance of insulating medium.
3. Power loss due to dielectric hysteresis.

By reference to the loss curves shown in Fig. 94, it is evident that the greater part of the power loss in both the open wire and coaxial types when operating with negligible reflection is due to the conductor thermal resistance. Owing to the low-loss insulation materials now available, the losses due to dielectric leakage and hysteresis can be reduced to a negligible quantity especially at standard broadcast frequencies.

The curves of Fig. 94 are the results of extensive r-f measurements which confirmed mathematical formulas given for calculation of losses

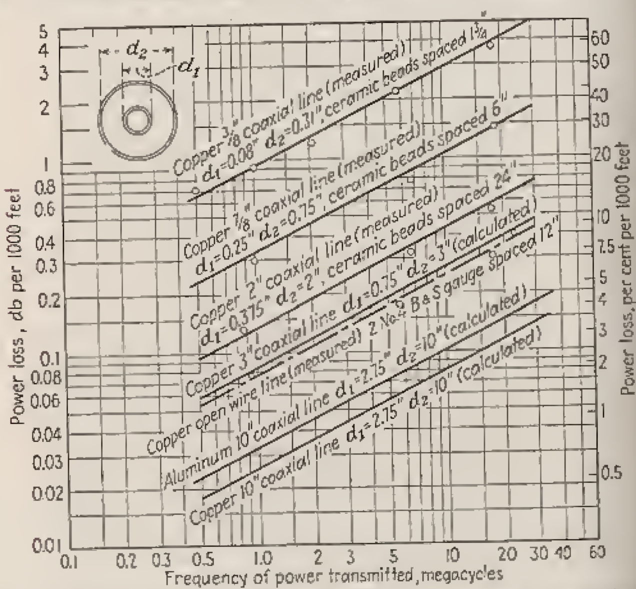


FIG. 94.—Power loss in lines operating at 20°C with negligible reflection in open wire and coaxial lines of various standard sizes.

Worthy of mention is the rather high efficiency of the open wire line consisting of two No. 4 B. & S. gage solid copper conductors spaced 12 in. center to center. The losses in this line are less than those in a 3-in.-diameter copper coaxial line. The resistivity ρ of aluminum being greater than copper, the losses in an aluminum coaxial line are somewhat greater than those in a copper line of the same dimensions.

71. Broadcast-station Signal Coverage. The reception of satisfactory signals from a given broadcasting station by a particular listener at a given point depends upon the following: (1) the intensity of the signal radiated from the antenna system of the station as influenced by the radiated r-f carrier power, antenna directivity, and percentage of modulation on the carrier; (2) distance between the broadcasting-station antenna and the point of reception and the attenuation characteristics of the intervening space or terrain; (3) intensity of objectionable interference at the receiving point; (4) fading as produced by the rays of direct and indirect signals; (5) the quality of the broadcast receiver and its ability to discriminate against local noise or interference and against adjacent channel interference and to convert the received r-f signals into sound without appreciable distortion or inherent receiver noise. The surrounding area about a given broadcasting station wherein satisfactory program signals can be received determines the service area of the station.

The service area of a standard broadcasting station of the 550- to 1,600-kc band consists essentially of two distinct regions. That region in close proximity to the station is served by the direct ray or ground wave called the *primary coverage area* of a broadcasting station, while the region at some distance from the station and served by virtue of indirect ray or sky-wave reflections is called the *secondary coverage area*. During daylight hours of broadcast transmission on frequencies of the standard broadcast band (between 550 and 1,600 kc), a broadcast listener is concerned with the primary coverage area signals of near-by stations for programs since there is very little sky-wave energy reflected during this period under normal conditions. The daylight service area of such a broadcasting station therefore consists almost entirely of that region served by the direct ray.

During the hours of twilight and darkness, the secondary coverage area of stations in the standard broadcast band becomes apparent. The secondary coverage area of a particular station begins at a considerable distance from a given station and is served by the predominant sky wave. The primary and secondary coverage areas of a broadcasting station are separated by a region known as the *fading area* of the station. In this area the signal intensities of the direct and indirect rays approach an equality with a result that violent fluctuations in signal intensities are apparent. The fading areas of stations are dependent upon a number of factors, such as frequency of transmission, antenna radiation characteristics, conductivity of intervening terrain, and time of day and season, and are independent of the transmitter powers of the stations. The fading area is normally in the form of a band about the broadcasting station normally contained within radii of between 20 and several hundred miles, depending upon the factors mentioned. The fading band may be as much as 50 miles in width.

Considering a broadcasting station radiating equally in all directions over surrounding terrain, and assuming equal ground attenuation, the service area would consist of a primary coverage area or circular area near the station and served by a steady ground-wave signal. Outside of this would exist the fading area consisting of a ring about the primary area. Beyond the fading ring the secondary coverage area would exist. Inasmuch as broadcast reception is rather uncertain in the fading region and in the secondary coverage area, the real value of a given station is dependent normally upon its primary coverage area.

The primary service area of a particular station can be most accurately determined by means of a field-intensity survey. A survey[1] of this kind is accomplished through the use of mobile field-intensity measuring equipment. This consists essentially of a field-intensity meter or

[1] KIRBY, S. S. and K. A. NORTON, Field Intensity Measurements, *Bur. Standards Jour. Research*, April, 1932.

carefully shielded receiver equipped with an indicating meter at its output terminals to read carrier-signal intensity as induced in the loop antenna. The field-intensity meter together with the loop antenna are carefully calibrated in their position in the measuring car to give accurate readings in microvolts per meter over a wide range of carrier-signal intensity.

72. Field-intensity Measurements. The procedure of making a field intensity survey consists usually of making frequent measurements at satisfactory positions (in free space) along radials progressing to and from the station. Eight or more radials at equal angular spacing are generally made about a point established on the field survey map by the broadcast station antenna system and extending to a signal intensity of 500 mv or beyond. Each radial is then plotted on loglog coordinate graph paper and a smooth curve drawn through these points to show directly the signal intensity along one ordinate, with distance along the other. Later the values required are transferred to a map in the form of signal contour lines representing positions about the station where field intensities of 100, 50, 10, 2, and 0.5 mv per meter exist. The contour map for reference purposes also contains information such as (1) station call letters, (2) frequency, (3) antenna power and its directivity and other characteristics, (4) scale of map, (5) date, etc.

Since fading occurs after sunset, these measurements are an indication of satisfactory daytime coverage only from the particular station. As recommended by reports of the I.R.E.,[1] FCC,[2] and the National Association of Broadcasters, values of standard broadcast field intensity considered necessary for reliable broadcast service are given for three areas as follows: (1) a business city area where a field intensity of from 10 to 25 mv per meter is required to override high interfering electrical noise and overshadowing effects of large buildings, (2) a residential district of a city where a field intensity of 2 to 5 mv per meter is required, (3) a rural area where 0.1 to 0.5 mv per meter signal intensity is sufficient. In addition it is stated that for fair service a signal intensity of one-half the above values is needed and for poor service one-fourth of these values. These figures are based upon the average signal intensity necessary to override the noise levels of these districts. In large cities, where large, tall buildings are numerous, a free space field intensity of as much as 50 mv per meter over the city may be necessary to provide a signal intensity at a particular receiving antenna between buildings of one-fifth of that amount.

Since the primary service area includes nighttime reception as well as daytime, fading measurements are necessarily a part of the field-intensity survey in determination of this area. Fading measurements are made with the same field-intensity measuring equipment used for the survey except that the field-intensity meter is equipped with a recording milliammeter (usually of 0 to 5 ma range) attached to the output of the field-intensity meter. A d-c amplifier sometimes is necessary to secure sufficient signal level to actuate the recording meter from the field-intensity measuring set. The equipment is set up for periods of time at a given distance and location from the station, and fluctuations in carrier-

[1] Report of Committee on Radio Propagation Data, *Proc. I.R.E.*, **21**, No. 10, October, 1933.

[2] Fifth annual report to the Congress of the United States, by Federal Communications Committee, gives tabulated values of field strength.

signal intensity are noted on the continuously moving recording chart. Amplitude fluctuations as recorded on the chart indicate the amount of fading. Fading measurements of considerable periods of time and over a wide area are necessary to determine the fading region about a given station and to evaluate the secondary coverage area about the station, particularly those designated as class I stations.

73. Calculations of Station Coverage. A mathematical investigation of the attenuation of radio waves propagating over plane earth has led to mathematical expressions which follow very nearly the characteristics of waves as indicated by actual measurements. A simplified form of this expression requires the following information for a solution: (1) the frequency of the transmitted wave, (2) the distance from the station, (3) conductivity of the soil in electromagnetic units, and (4) the inductivity of the soil in electrostatic units. Since the inductivity can be generally assumed to be 14 to 15 e.s.u., then, with a measured value of conductivity σ, the field intensity at a given distance from a station may be calculated. With further assumptions concerning the irregularities in general characteristics of the terrain about the station, it is possible to calculate the contours. The value of σ (the soil conductivity) is usually secured from a measured radial or taken from available field-intensity measurements of some other station in the vicinity. The FCC has published charts showing soil conductivity over the United States. Provided measured values are not available, these may be used.

For convenience the chart shown in Fig. 95 is given. It may be used to calculate signal attenuation of standard broadcast frequencies. The attenuation curves shown are derived from a simplified form of Sommerfeld's attenuation formula.[1] With a single set of Sommerfeld curves to cover all the standard broadcast frequencies and soil conductivities, the conductivity of a given soil can be rather easily computed from the attenuation of a particular signal. This is accomplished by first converting a given radial to an inverse field strength of 1,000 mv at 1 mile and then determining the frequency of the ground-wave curve with which it coincides, *i.e.*, the conversion frequency. The conductivity is secured from the soil-constant curve passing through the intersection of the operating and conversion frequencies on the conversion chart in the upper right corner of the figure. The conversion chart has been prepared from the following relationships:

$$f_1 = f\sqrt{\frac{\delta}{\delta_1}}$$

where f = operating frequency
δ = standard conductivity of chart (100×10^{-15} e.m.u.)
δ_1 = actual soil conductivity
f_1 = conversion frequency.

[1] SOMMERFELD, ARNOLD, Ausbreitung der Wellen in der drahtlosen Telegraphie Einfluss der Bodenbeschaffenheit, und gerichtete und ungerichtete Wellenzüge. *Jahrb. drahtlosen Tele, Tele.*, **4**, December, 1910. ROLF, Numerical Discussion of Sommerfeld's Attenuation Formula, *Proc. I.R.E.*, **18**, No. 3, March, 1930. ECKERSLEY, P. P., The Calculation of the Service Area of Broadcast Stations, *Proc. I.R.E.*, **18**, No. 7, July, 1930. ECKERSLEY, T. L., Direct Ray Broadcast Transmission, *Proc. I.R.E.*, **20**, No. 10, October, 1932. NORTON, K. H., Propagation of Radio Waves over a Plane Earth, *Nature*, June 8, 1935; Propagation of Radio Waves over the Surface of the Earth and in the Upper Atmosphere, Part I, *Proc. I.R.E.*, **24**, October, 1936; Part II, *Proc. I.R.E.*, **25**, September, 1937. FITCH, W. A., The Sommerfeld Formula, *Electronics*, **9**, No. 9, September, 1936.

For example, assume a station operating on an assigned frequency of 660 kc where the field strength radial, as plotted from measurements, follows the 1,500-kc curve. Then, from the soil conversion chart,

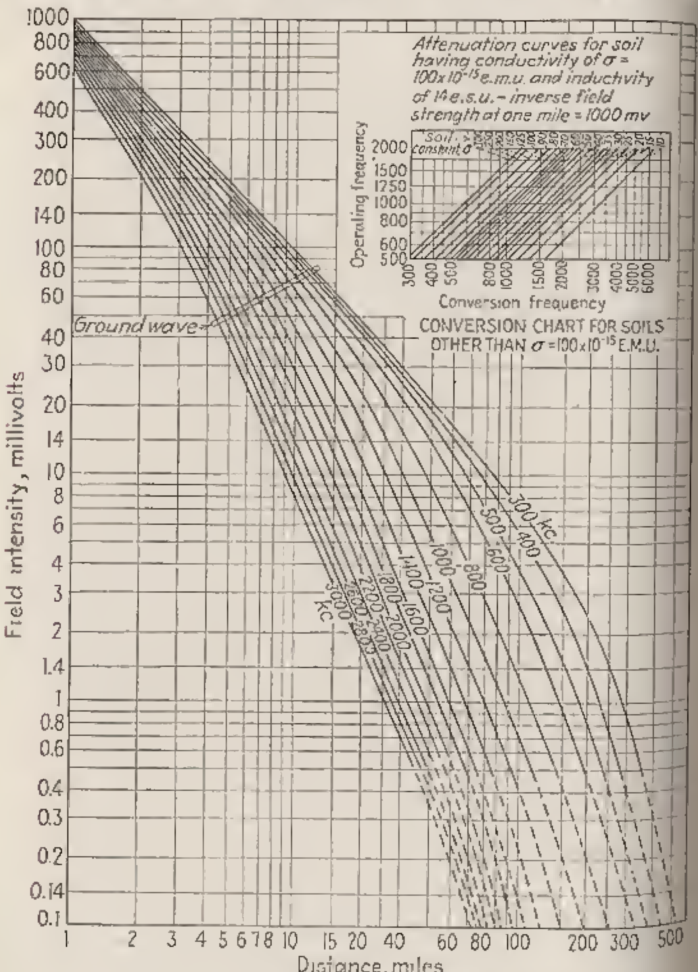


FIG. 95.—Ground-wave field-intensity curves.

the conductivity is very nearly 20×10^{-15} e.m.u. On the other hand, if the soil conductivity is known, the signal attenuation can be determined from the conversion chart and the attenuation curves

of various frequencies. Since these curves are based on a field strength of 1,000 mv at 1 mile, the actual signal at a given distance from a station is, of course, derived from the ratio of the actual signal intensity in millivolts at 1 mile from the particular station divided by 1,000. At considerable distances from the transmitter these curves are subject to corrections for the effects of curvature of the earth.

The curves in Fig. 96 refer to sky-wave intensities under various conditions of propagation. These are most useful in the determination of the fading regions about a particular station and are plotted to give intensities of reflected sky-wave intensity for different antenna electrical heights based on a signal intensity of 1,000 mv along the ground at 1 mile from a given antenna. In this case the electrical height of the antenna in degrees equals $3.85 \times 10^{-4} Hf$, where H is the physical height

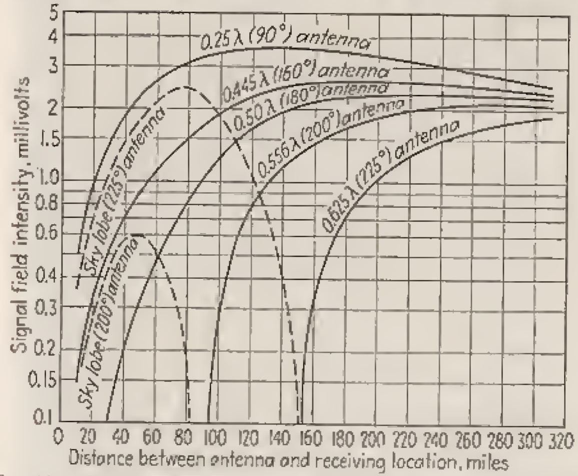


FIG. 96.—Sky-wave propagation curves for antennas of various heights. Layer height = 100 km, reflection coefficient = 1.

of the antenna in feet and f is the operating frequency in kilocycles. This is based on a velocity of propagation equivalent to 0.95 that of light.

Inasmuch as the attenuation curves of ground-wave intensity (Fig. 95) are also based on 1,000 mv at 1 mile, then the particular distance from a given antenna where the sky-wave intensity, shown on curves of Fig. 96, equals the ground-wave signal intensity of Fig. 95 is the distance from the antenna where one would expect to observe greatest fading or is an estimate of the center of the fading band. Owing to the height of the Heaviside layer being other than 100 km and reflection being less than unity, on which these curves are based, calculated distances given by these curves are approximate. Measurements are required for more exact determination of the fading region.

The service rendered by a standard broadcast station depends also on interference caused by other stations on the same and near-by channels. This interference is greatly increased at night because signals from undesired distant stations are reflected by the Heaviside layer and may be received with varying intensities within the service area of a desired station. Following extensive survey work covering nighttime signal propagation over the period February to May, 1936, the FCC issued a report[1] wherein a great amount of information concerning sky-wave propagation is given. In Fig. 97 are illustrated curves representing the average sky-wave field intensity (second hour after sunset) at the recording station. An interfering or undesired signal existing for 10 per cent of the time has been standardized as an interfering signal.

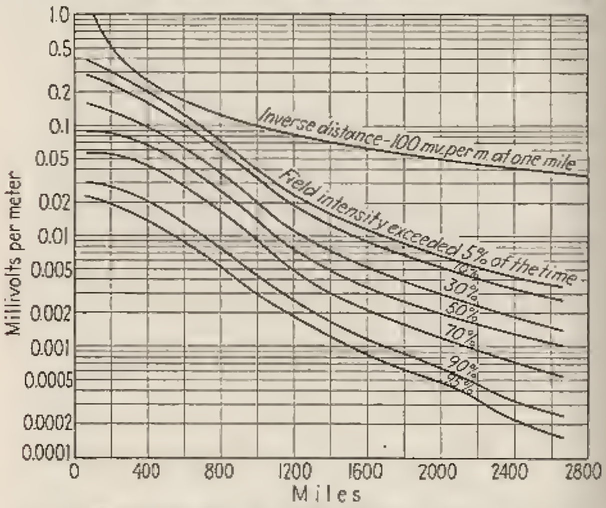


FIG. 97.—Average sky-wave field intensity, 640 to 1190 kc.

Thus, with the antenna sky-radiation characteristics of a given station known,[2] it becomes possible to estimate the amount of undesirable interference it is liable to cause to another distant station. In the determination of interference problems, the FCC has indicated its approval of ratios as follows in the protection of a desired standard broadcast station signal against an undesired one, using an average receiver:

Desired and Undesired Signal	Ratio of Intensities Desired to Undesired Signal
Same frequency	20
±10 kc	2
±20 kc	0.1
±30 kc	0.02

These apply to ground-wave signals, whereas the ratio for sky-wave signals are 0.2 for 10-kc channel separation and 0.04 for 20-kc channel

[1] *FCC Report* 18108, September, 1936.

[2] Standards of Good Engineering Practice concerning Standard Broadcast Stations, *FCC Report* 41831, June 29, 1940.

separation. The FCC has classified standard broadcast stations with respect to protected service contours and permissible interference signals in accordance with Table IV.

TABLE IV.—PROTECTED SERVICE CONTOURS AND PERMISSIBLE INTERFERENCE SIGNALS FOR BROADCAST STATIONS

Class of station	Class of channel used	Permissible power, kilowatts	Signal intensity contour of area protected from objectionable interference: Day,† microvolts per meter	Night, microvolts per meter	Permissible interfering signal on same channel: Day,† microvolts per meter	Night,‡ microvolts per meter
I-A	Clear	50	SC 100 AC 500	Not duplicated	5	Not duplicated
I-B	Clear	10–50	SC 100 AC 500	500 (50% sky wave)	5	25
II	Clear	0.25–50	500	2,500§ (ground wave)	25	125§
III-A	Regional	1–5	500	2,500 (ground wave)	25	125
III-B	Regional	0.5-1 per night and 5 per day	500	4,000 (ground wave)	25	200
IV	Local	0.1–0.25	500	4,000 (ground wave)	25	200

† Ground wave. SC = same channel; AC = adjacent channel.
‡ Sky-wave field intensity for 10 per cent or more of the time.
§ These values are with respect to interference from all stations except class I-B, which stations may cause interference to a field-intensity contour of higher value.

74. **High-frequency broadcast-station coverage** concerns the stations licensed by the FCC primarily for the transmission of radio telephone emissions in the h-f broadcast band for reception by the general public. The h-f broadcast band contains the band of frequencies extending from 43 to 50 Mc, inclusive. In accordance with Sec. 3.225(d) of the FCC rules, the stations in this band must use a system of modulation of the radio signal in which the frequency of the carrier wave is varied with the program signal; this being commonly termed *frequency modulation* or f.m. The assigned operating frequency or "center frequency" is that of the r-f carrier without modulation. It must be maintained within 2,000 cycles of the assigned center frequency assigned. Channels for h-f broadcast stations begin at 43.1 Mc and continue in successive steps of 200 kc to and including the assigned frequency of 49.9 Mc.

According to Sec. 3.222 of the FCC rules, h-f broadcast stations shall be licensed on the basis of an area in square miles within the service area. The contour bounding the service area and the radii of same are determined in accordance with the FCC standards. On this basis, a h-f broadcast station has a single service; that corresponding to the primary service of a standard broadcast station. Secondary, sky-wave or intermittent, service is not recognized in h-f broadcast coverage.

In *FCC Report* 41831[1] the standard of field intensity necessary for satisfactory service is given as follows:

TABLE V.—SERVICE

Area	Median Field Intensity, Millivolts per Meter
City areas near factories, car lines, or busy streets	1
Rural areas away from highways	0.050

These figures are based on the absence of objectionable fading and usual noise levels encountered in these areas and are not dependent upon interference from other h-f broadcast stations. The chart of Fig. 98

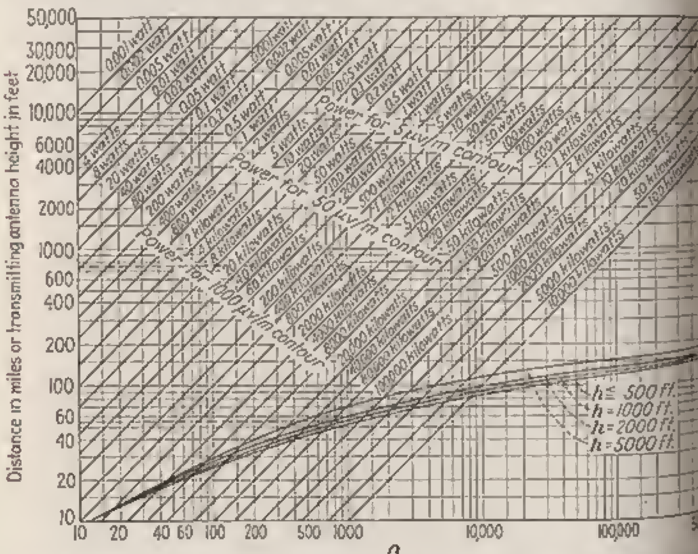


FIG. 98. Signal range for high-frequency broadcast stations. (f = 46 Mc; $\sigma = 5 \times 10^{-14}$ e.m.u.; dielectric constant ϵ = 15.)

may be found useful in the determination of the signal intensity rendered from a given h-f broadcast station. The results obtainable are based on a signal intensity at a receiving antenna with an elevation of 30 ft. The distance to the 50-μv per meter contour about a given station is dependent upon values of the transmitting antenna height, the antenna power, and the antenna field gain.

The procedure in using the chart is as follows: Assume that there is a h-f station with an antenna height of 750 ft., an antenna power of 500 watts, and an antenna field gain of 2. To determine the distance to the 50-μv per meter contour, refer to the dashed horizontal line extending from the 750-ft. antenna elevation over to the 45-deg. line marked 2 kw. Then proceed vertically downward to a point midway between the curved lines represented as 1,000 and 500 ft. Finally, proceed horizontally again to the left to find that the expected range is 54.5 miles or the radius to the 50-μv per meter contour. The procedure may be reversed to determine the power required for a given antenna height to produce a signal intensity of 50 μv per meter for a certain distance.

[1] Standards of Good Engineering Practice Concerning High Frequency Broadcast Stations, *FCC Report* 41831, June 29, 1940.

The additional power scales are useful in estimating the distance to the 5- and 1,000-μv per meter contours. The scale indicated by θ at the bottom of the chart is used for the purpose of finding the distance to any desired contour. In this instance

$\theta = h \times P^{1/2} \times G \times (50/F)$

h = transmitting antenna height in feet

$P^{1/2}$ = square root of the antenna power in kilowatts

G = antenna field gain

F = desired field intensity in microvolts per meter.

By means of the above equation, θ may be determined. Then the corresponding distance can be determined from the chart by proceeding vertically at that value of θ to the proper curved line and then in a horizontal direction to the the left, where the distance is given.

In the consideration of objectionable interference from other stations on the same and adjacent channels, the FCC Sec. 3.225(f), requires that the proposed station shall not have interference to such an extent that its service may be reduced to an unsatisfactory amount. For this reason objectionable interference is considered to exist when the signal for 50 per cent of the distance in any sector on a radial exceeds 0.005 mv per meter at the 0.050 contour of the desired station. If it is considered that a station is protected to the 1-mv per meter contour, objectionable interference occurs when the signal for 50 per cent of the distance in any sector exceeds 0.1 mv per meter. For other field intensities the ratios in Table VI govern allowable ratios of the desired to undesired signals.

TABLE VI.—ALLOWABLE SIGNAL RATIOS

Channel Separation	Ratio of Desired to Undesired Signals
Same channel	10:1 median field intensity
Adjacent channel	2:1 median field intensity

The service contours in the cases above are determined by actual measurements or by means of Fig. 98.

High-frequency broadcast transmitters are normally located as near to the center of the proposed service area as possible. A high elevation of the transmitting antenna is necessary to reduce the shadowing effects upon propogation due to hills, buildings, and other obstructions in the city area. The position of the transmitter site is also dependent upon the purpose of the station, *i.e.*, whether it is intended to serve a small city, a metropolitan area, or a large region. A suitable transmitter site may be made available by the use of a directive antenna. Where a directive antenna is used, a centrally located station site may not be a desirable one. As one may understand by studying the chart in Fig. 98, the transmitter antenna height above the average elevation of the service area is a consideration of greatest importance to secure optimum coverage with a high-frequency broadcast station.

SECTION 22

LOUD-SPEAKERS AND ROOM ACOUSTICS

By Hugh S. Knowles[1]

In the design and operation of electroacoustic devices, consideration must be given both to the physical or "objective" properties of the sounds that are to be reproduced and to the psychophysiological or "subjective" processes involved in hearing.

a. Sound is an alteration in pressure, particle displacement, or particle velocity propagated in an elastic material or the superposition of such propagated alterations.
b. Sound is also the sensation produced through the ear by the alterations described above. In case of possible confusion the term "sound wave" may be used for concept (*a*), and the term "sound sensation" for concept (*b*).[2]

In the case of a sound wave in air the pressure is alternately above and below atmospheric.

The velocity of propagation, c, of a sound wave of small amplitude is

$$c = 33{,}060 + 61\theta \text{ cm per sec.} \quad (1)$$

where θ is the temperature in degrees centigrade. The wave length λ is given by the relation $\lambda = c/f$, where f is the frequency in cycles per second. The density ρ of dry air at 20°C. and at a pressure of 760 mm is 0.001205 g per cubic centimeter.

The intensity of a plane or spherical "free" sound wave (no reflection) in the direction of propagation is

$$I = \frac{P^2}{\rho c} = 2.42 \times 10^{-9} P^2 \text{ watt per sq cm} \quad (2)$$

where P is the effective sound pressure (dynes per square centimeter).

The standard reference intensity is 10^{-16} watt per square centimeter. The intensity level in decibels of a plane or spherical free sound wave in the direction of propagation is

$$I_L = 10 \log_{10} 2.42 \times 10^7 P^2 \quad (3)$$

The standard reference pressure is 0.0002 dyne per square centimeter. In a plane or spherical free wave the intensity is proportional to the square of the pressure. In this case the pressure level in decibels of a sound wave is defined as

[1] Jensen Radio Mfg. Co.
[2] American Tentative Standard Z24.1, Acoustical Terminology.

$$P_L = 20 \log_{10} 5{,}000P \quad (4)$$

Two pressures are said to differ by x db if 20 times the logarithm to the base 10 of their ratio is x. As in the analogous electrical case involving a field strength or voltage ratio, this is valid only if the impedances are identical so that the energy is proportional in both instances to the square of the respective pressures. For this reason Eqs. (2) and (4), in general, do not hold in more complicated fields.

1. Speech. The variation in conversational speech power with frequency is shown in Fig. 1 (after Sivian and Fletcher). The ratio of ⅛ sec. peak to averaged power in 15-sec. intervals is roughly 20 db. In overloaded amplifiers such as are frequently used in public address systems, the ratio may be 10 db or less. This ratio is important in

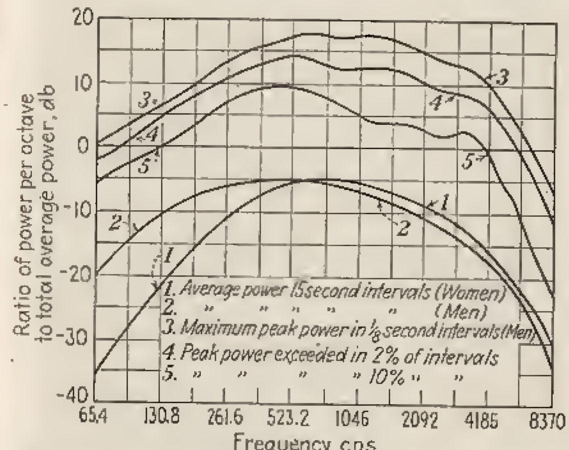


Fig. 1.—Variation of conversational speech power with frequency. (*After Sivian and Fletcher.*)

temperature-limited loud-speakers (see Tests). The distribution of energy with frequency is brought out differently in Fig. 2 (after Fletcher).

Articulation curves which give a measure of the "recognizability" of speech are shown in Fig. 3 (after Fletcher). The percentage of called sounds correctly recognized is the per cent articulation. Tests of syllable, sound, vowel, individual sound, and other types of articulation are now widely used in the laboratory and to an increasing extent in the field to determine the suitability of a system for the transmission of speech. "Intelligibility" tests, in which the content of a simple sentence is to be understood, are also used. On the average 30 per cent syllable articulation corresponds to nearly 90 per cent "discrete sentence" intelligibility, indicating the relative ease of understanding connected speech. From Figs. 2 and 3 we note that reproducing only the frequencies above 400 cycles halves the system power requirement and yet reduces the articulation by a negligible amount. In a power-limited system in which speech articulation is important, the transmission band is sometimes limited to from 600 or 800 to 4,000 cycles, corresponding roughly to the 90 per cent articulation points at each end. This reduced band roughly quarters the power requirement.

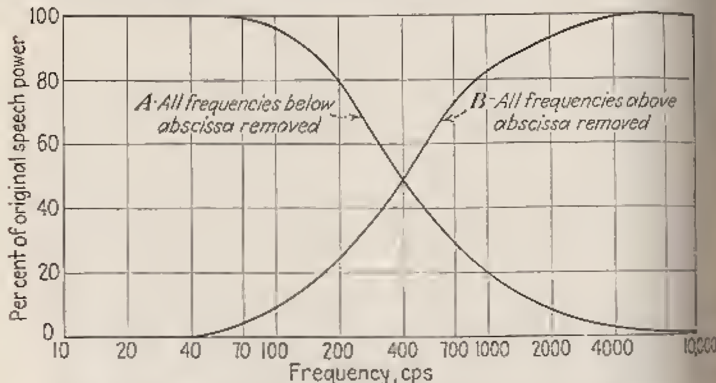


FIG. 2.—Speech power variation with frequency. (*After Fletcher.*)

Articulation and naturalness are not to be confused. By successively raising the cutoff of high-pass filters and lowering the cutoff frequencies of low-pass filters, each by a barely perceptible amount, Schäfer has shown

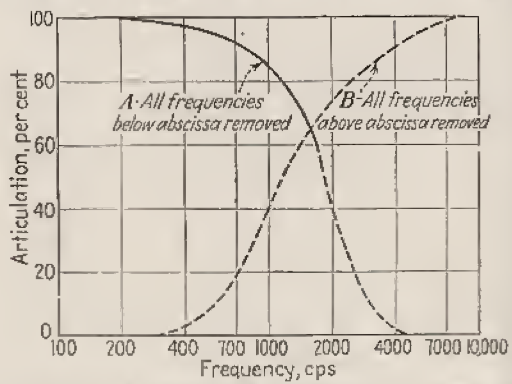


FIG. 3.—Variation of articulation with transmitted frequency range. (*After Fletcher.*)

that the required transmission band for natural speech reproduction includes some 32 to 36 minimum perceptible changes in band width.

[1] SCHÄFER, E., The Audibility of Variations in Frequency Band in Speech Transmission, *Elek. Nach.-tech.*, **15**, No. 8, 237–240, August, 1938.

The steps are roughly logarithmic. Some change in quality could be detected when frequencies above 8,000 were attenuated. The transmission of natural sounding speech and noises which accompany it therefore appears to require the transmission of all frequencies from 100 to about 10,000 cycles.

2. Music. The frequency distribution of the maximum and most probable peak powers for a 75-piece orchestra is shown in Fig. 4 (after Fletcher). The curves are based on average measurements of four selections which gave whole "spectrum" peak powers from 8 to 66 watts and average powers of 0.08 to 0.13 watt. Zero level corresponds to an average power of about 0.1 watt. As in the case of speech the average power over 15-sec. intervals is about 1 per cent of the peak power in ⅛-sec. intervals.

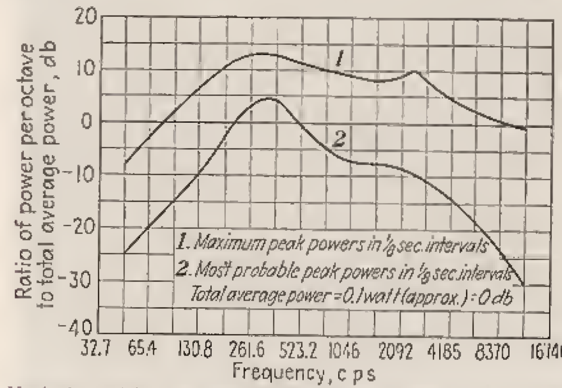


FIG. 4.—Variation with frequency of the power output of a 75-piece orchestra. (*After Fletcher.*)

The power output of various musical instruments is shown in Table I. The bass drum may radiate over a third of the peak power of a 75-piece orchestra. The large peaks in the 20- to 62.5-cycle range of the organ are well known to recorders and electronic organ people who find it desirable to use l-f stops which are "rich in harmonic development" and therefore sound much louder without badly overloading the record, amplifier, and speaker. The 15-in. cymbals follow the drums and organ closely in peak power output with 9.5 watts. Their maximum peaks occur in the 8,000- to 11,300-cycle range. Transmission systems having a "predistorted" frequency characteristic which includes a marked rise in h-f response in some part of the system (such as f-m and television transmitters) are frequently overloaded by this instrument. The same problem occurs in recordings recorded with a similar characteristic.

The high output of the trombone in the 2,000- to 2,800-cycle band near the frequency of maximum ear sensitivity gives the trombone (and other brass instruments) their piercing "bite." It has been found that the ear critically appraises the response of a system in this range and that surprisingly small changes can be detected. This suggests that the balance of the brasses in a studio pickup merits special attention.

TABLE I. PEAK POWERS IN MUSIC

Instrument	Microphone position and assumption in converting to total sound power	Field pressure, dynes per square centimeters: Average in 15-sec. interval	Field pressure, dynes per square centimeters: Peak in ⅛-sec. interval	Total peak power, watts	Percentage of intervals	Frequency band containing maximum peaks cps.
Bass drum, 36 × 15 in.	3 ft. in front, on axis. Radiation confined to a cylinder having drum diameter	99.0	1,260.0	24.6	6.0	250– 50
Bass drum, 30 × 12 in.	Same as above	35.0	980.0	13.4	1.0	125– 250
Snare drum....	4 ft. in front, 90 deg. off axis. Peak pressure increased 8.5 db for 1-ft. distance. Radiation confined to hemisphere	14.6	365.0	11.9	2.5	250– 500
15-in. cymbals.	3-ft. distance. Peak pressure increased 7.2 db for 1 ft. Radiation confined to hemisphere	18.0	360.0	9.5	7.5	8,000–11,300
Triangle.......	3-ft. distance. Conversion as for cymbals	2.3	25.8	0.05	1.0	5,600– 8,000
Bass viol.......	3-ft. distance. Radiation confined to hemisphere	4.2	37.8	0.156	2.0	62– 250
Bass saxophone	3-ft. distance. Radiation confined to hemisphere	4.1	58.2	0.288	25.0	250– 500
BB♭ tuba.....	3-ft. distance. Conversion made from measurements with a complex sound source attached to a horn of similar size	5.4	43.2	0.206	17.0	250– 5
Trombone.....	3-ft. distance. Conversion as for tuba	6.5	228.0	6.4	5.0	500– 2,000– 2.
Trumpet.......	3-ft. distance. Conversion as for tuba	8.6	54.2	0.314	18.0	250– 7
French horn...	As for trumpet	3.8	27.0	0.053	6.0	250–
Clarinet......	As for trumpet	3.3	26.4	0.050	5.5	250– 5

TABLE I. PEAK POWERS IN MUSIC.—(*Continued*)

Instrument	Microphone position and assumption in converting to total sound power	Field pressure, dynes per square centimeters: Average in 15-sec. interval	Field pressure, dynes per square centimeters: Peak in ⅛-sec. interval	Total peak power, watts	Percentage of intervals	Frequency band containing maximum peaks cps.
Flute..........	As for trumpet	1.6	25.6	0.055	1.0	700– 1,000 1,400– 2,000
Piccolo........	As for trumpet	2.2	30.8	0.084	0.5	2,000– 2,800
Piano.........	10-ft. distance. Room 29 × 29 × 13 ft. Reverberation time 1 sec., 60–1,000 ~, average of 3 methods	2.6	23.4	0.267	16.0	250– 500
[illegible]-piece orchestra	6 ft. from nearest instruments, in same room as piano. Average of 2 methods	7.9	126.0	9.0	1.5	250– 500 2,000– 2,800
[illegible]-piece orchestra	15 ft. from nearest instrument in theater	4.6	129.0	66.5	1.0	250– 500 8,000–11,300
[illegible] organ.....	Effective distance 15 ft. Radiation assumed uniform over ¼ sphere	20.0	90.0	12.6	36.0	23– 62.5

The audible frequency ranges of many musical instruments are shown in Fig. 5 (after Snow). The vertical ruled portions indicate the frequency range in which noises accompanying the playing of the instrument occur. While the elimination of these frequencies permits the fact that the frequency range is restricted to be detected, it does not mean that the quality is judged to be best with the unrestricted range. In many cases the quality of the reproduced music from instruments which radiate extraneous noises (reed, bowing, key, and others) is improved by eliminating the noise range.

In restricting the transmitted frequency range of reproduced music, we have to be primarily concerned with the degradation in quality as judged by a good "sound jury" rather than with recognition of the selection played or the power distribution with frequency or "spectral composition" of the music. The average results of a test of this kind, using a jury of 10 and an 18-piece orchestra, are shown in Fig. 6 (after Snow). Considering the many variables involved, the maximum and mean deviations from the curve were surprisingly small. It was the judgment of the observers that the quality improved rapidly as the lower range was extended to 80 cycles and the upper to 8,000 cycles.

It has been found experimentally that, if the transmitted frequency range is to be restricted, good balance between low and high frequencies may be obtained by so choosing the range that approximately eq degradation in quality occurs because of loss of low and high frequen For reasonable degradation the product of these two frequencies roughly 640,000. The square root of this product or the geometric m of these frequencies is therefore roughly 800 cycles. A system tra mitting more octaves below 800 cycles than above usually soun

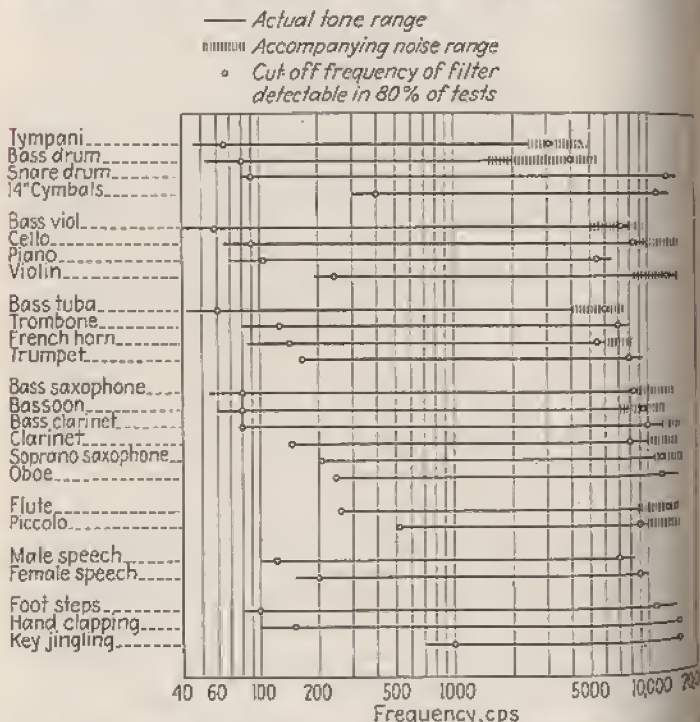


FIG. 5.—Audible frequency ranges of some musical instruments and soun (*After Snow.*)

"heavy," "thick," or "drummy." Likewise a system transmitti more octaves above 800 cycles than below will sound "thin" or "tin This assumes flat response in the range and similar cutoff characteris A sharp cutoff at one end will increase the apparent output at that because of the transient response which accompanies such a cutoff. peak in either range will increase the steady-state and transient respo in that region. This can be only partly balanced by added response the other range.

In considering the problem of reproducing sounds in a compl system including the effect of the room at the source of sound and at

source of the reproduced sound, it is desirable to know the energy distribution with frequency of a typical sound. The importance of this

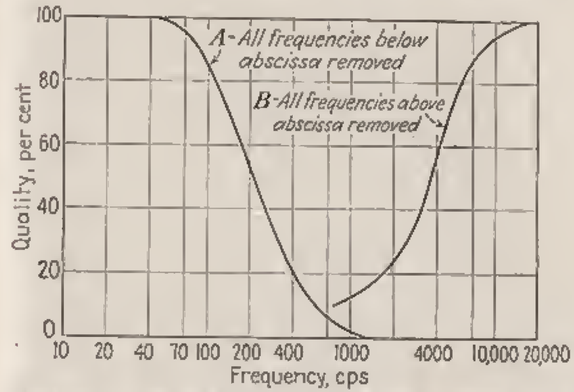


FIG. 6.—Variation in quality of reproduced orchestral music with transmitted frequency range. (*After Snow.*)

will be discussed under Room Acoustics. Since a common type of sound in music is a damped sinusoid, corresponding, for example, to the sound output of a plucked string instrument, the spectral analysis

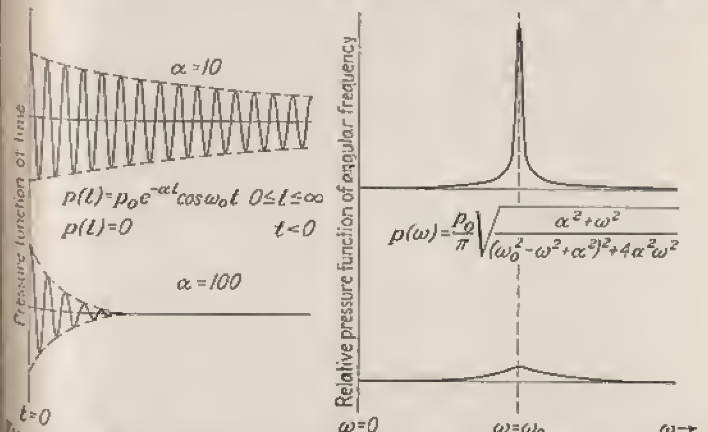


FIG. 7.—Pressure variation with frequency for two isolated damped sinusoids with different rates of decay, values of α the same in the two sets of curves.

of two waves with different rates of decay is shown in Fig. 7. Any isolated wave train of this type contains energy which covers an infinite frequency interval. By analogy with the optical case the spectrum is said to be a continuous or band spectrum.

The highly damped wave contains appreciable energy at frequencies differing up to from 20 to 30 per cent from the frequency of a corresponding undamped wave. As the rate of decay is decreased, the wave train contains more energy, and an increasing amount of this is concentrated near the undamped frequency of the wave. In the limit when the rate of decay is zero and the wave has existed for an infinite length of time, i.e., when we have a steady state, the band spectrum degenerates into a line spectrum with all the energy concentrated at the undamped frequency of the wave.

The fact that music and speech are not of a steady-state character but vary from instant to instant (and therefore have a continuous distribution with frequency of their energy) substantially aids their satisfactory transmission in a room (see Room Acoustics).

3. Noise. Noise is an "unpitched" sound composed of a large number of discontinuous, non-periodic sounds. Therefore the energy in noise is distributed in a continuous manner with frequency. A click, for example, closely approaches the hypothetical pulse which lasts for an infinitesimal length of time and the energy of which is continuously and uniformly distributed with frequency. A noise may have one or more broad peaks in its band spectrum, but a sharp peak indicates a nearly periodic disturbance which will give the noise a definite pitch.

The properties of noises are of some importance because (1) the proper reproduction of intended noises may enhance the dramatic value of the reproduction, (2) the ambient noise levels in studios or halls and in rooms frequently limit the dynamic range at the "pickup" and "playback" points, and (3) they influence the response of the ear by producing masking or artificial deafening.

The reproduction of most noises requires the transmission of substantially the entire audible frequency range. For this reason noises are frequently used as test material in high-quality systems. The frequency ranges of footsteps, hand clapping, and key jingling are shown in Fig. 5. These indicate that it is particularly important that all the upper audible frequencies be transmitted.

The intensity level of various representative noises is listed in Table II. In urban locations, particularly in large buildings, the ambient noise level in moderately quiet rooms is of the order of 45 to 60 db. This noise level is high enough so that even in specially treated broadcasting studios it frequently limits the dynamic range of the transmitter.

Even in relatively quiet residential sections the ambient noise level in a typical listening room is high enough so that it, too, places a lower limit on the intensity of the sound required to overide the noise.

4. Hearing. All the data contained under Speech, Music, and Noise which were obtained by the use of a sound jury or listener involve the sensation produced in the listener by the designated source of sound. All tests of this type depend to some extent on the techniques employed and, of course, on the observers. All similar tests are of principal value when the jury is composed of a large selected sample with known hearing characteristics.

One of the most important properties of sound is its loudness. This has been found to vary with both the frequency and intensity of the sound. To a rough approximation it has been found that in the middle

TABLE II. NOISE LEVELS

Pressure, dynes per square centimeter	Noise level, decibels above reference threshold	Type of noise
630	130	Pain threshold
250	122	Airplane—1,600 rpm, 18 ft.
45	109	Boiler factory
25	102	Subway train passing station
13	96	Elevated train—15 ft.
4.0	86	Heavy traffic—15 ft.
2.0	80	Average truck—15 ft.
1.3	75	Average factory location
6.3×10^{-1}	70	Average automobile—15 ft.
3.2×10^{-1}	64	Department store
1.1×10^{-1}	55	Average office
2.8×10^{-2}	43	Quiet office
6.3×10^{-3}	30	Very quite residence
1.4×10^{-3}	17	Gentle whisper—5 ft.
4.5×10^{-4}	7	Threshold (for street noise)

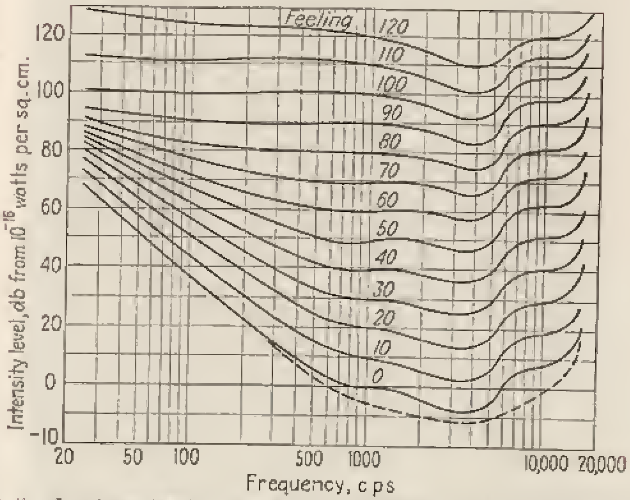


FIG. 8.—Loudness level curves showing variation in sound intensity with frequency required to produce a sound judged to be as loud as the 1,000-cycle reference sound intensity given on the curves. (*After Fletcher and Munson.*) Solid curves obtained with listener facing sound source. Dashed curve indicates threshold (corresponding to solid curve 0) but for sound of random incidence. (*After Sivian.*)

frequency range equal percentage increases in intensity produce equal increases in loudness. The *loudness* is the magnitude of the hearing sensation and is assumed proportional to the number of nerve impulses reaching the brain per second.

The *loudness level* contours for a sample of 200 ears are shown in Fig. 8 (after Fletcher and Munson). These curves were obtained by alternately listening to a sound of arbitrary frequency and intensity and comparing it with a 1,000-cycle tone the intensity of which was adjusted until the two were judged to be equally loud. At 1,000 cycles therefore the loudness level of the sound corresponds to the intensity level because

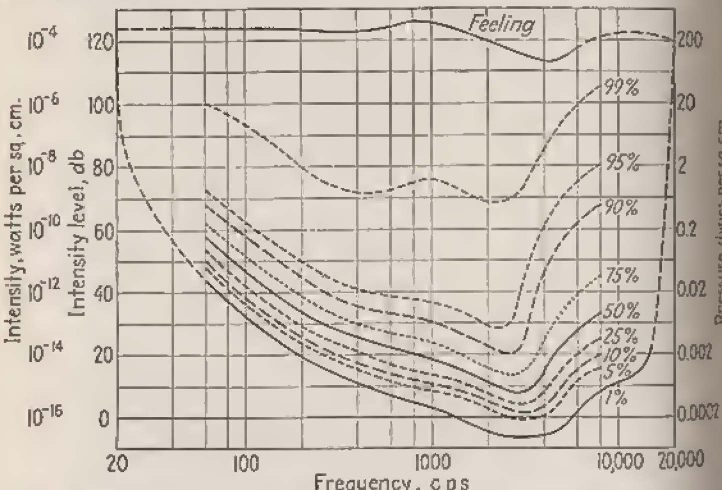


FIG. 9.—Threshold of hearing curves for large population sample. Percentage figures indicate percentage of sample tested having a hearing threshold lower than the corresponding curve. (*After Beasley.*)

this is the reference test frequency. The intensity is that which exists in an undisturbed sound field before the listener is immersed in it. The observer faces the source and listens to the sound binaurally. By plotting the differences in minimum audible field intensities for sound of normal and random incidence found by Sivian, we obtain the dotted curve in Fig. 8. This indicates that the other contours for sound of random incidence would also be more regular.

Recently reports have been made by Beasley on a sample of 16,000 ears. Some of the results are shown in Fig. 9. The curves show the percentage of the sample tested which had lower thresholds of hearing than the indicated value. For example, the solid curve marked 50 per cent indicates that 50 per cent of the ears tested had thresholds of hearing equal to or better than that indicated by this curve. From these data we see that the Fletcher and Munson threshold curves are for ears in the upper 1 per cent of the 16,000-ear sample, and that hearing deficiencies are prevalent enough to justify their consideration in equipment design.

The loudness or apparent response or transmission characteristic of a system emitting a plane free sound wave of three constant intensity levels is shown in Fig. 10. A loud sound (constant 100-db intensity level) seems almost equally loud from 30 to 6,000 cycles. A sound of moderate intensity (constant 60-db intensity level) is inaudible below 60 cycles and increases in loudness rapidly up to 400 cycles. In the presence of noise, masking would substantially reduce the loudness at low intensities. The l-f characteristic varies much more rapidly with intensity level than the high and for this reason compensated volume controls are designed to have their maximum effect at low frequencies.

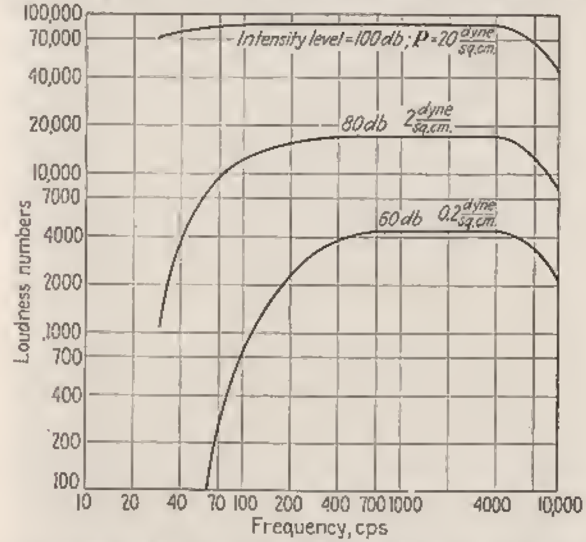


FIG. 10.—Loudness variation with frequency for three pure tones of the indicated intensity showing reduced loudness of low intensity l-f sounds.

It should be noted that the intensity-level compensated characteristic is a purely arbitrary thing and, although it is considered superior by some, it does not correspond to our normal experience. In practice, when we go some distance from the source, the low frequencies "drop out." When an orchestra plays at low intensity, we get the same effect.

The effect of noise on hearing is to produce artificial deafness or "masking." The nature of the apparent deafness which results depends on the spectral composition of the noise. Many noises produce fairly uniform deafening or masking. The effect of moderate noise levels is to decrease articulation. This may be largely compensated by raising the intensity level of the sound.

LOUD-SPEAKERS

A loud-speaker is a device which is actuated by electrical signal energy and radiates acoustical energy into a room or open air. The shorter term *speaker* is used when no confusion with a person addressing a microphone results.

The selection and installation of a speaker as well as its design should be guided by the problem of coupling an electrical signal source as efficiently as possible to an acoustical load. This involves the determination of the acoustical load or radiation impedance and selection of a diaphragm, motor, and means for coupling the loaded loud-speaker to an electrical signal source. The performance of the speaker is inti-

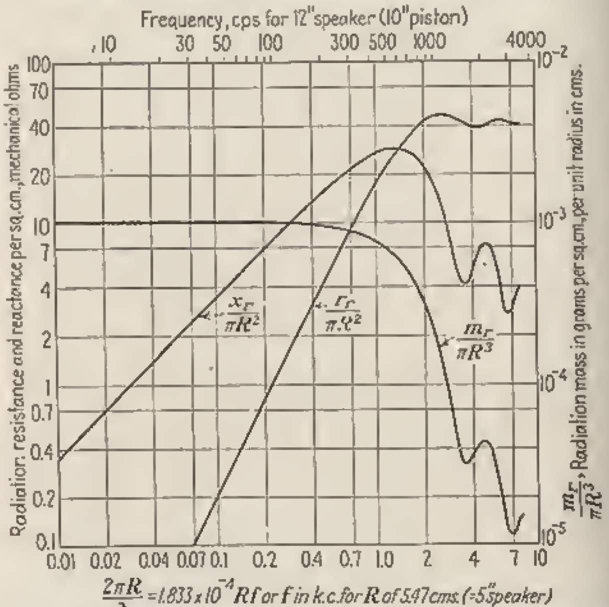


FIG. 11.—Radiation resistance, reactance, and mass per square centimeter of a flat, rigid piston vibrating in an infinite, rigid non-absorbing baffle. Piston radiates into a solid angle, $\Omega = 2\pi$ steradians (hemisphere).

mately connected with the nature of its acoustic load and should not be considered apart from it. The nature of the radiating system, and therefore the acoustic load impedance it sees, is primarily determined by space, acoustical environment, and cost factors.

5. Radiation Impedance. When a vibrating diaphragm is placed in contact with air, its impedance to motion is altered. The added impedance seen by the surfaces which emit useful sound energy may be called the *radiation impedance.* By analogy with antenna systems the resistive part is called the *radiation resistance.* The radiation reactance or reactive part is usually positive, and the corresponding apparent mass may be called the *radiation mass.* The radiation impedance seen by a

diaphragm depends on its size, shape, the frequency, the acoustical environment, and the medium into which it radiates.

6. Single Piston. The average radiation impedance per unit area seen by a flat circular piston vibrating in a thin, rigid, non-absorbent, infinite plane or baffle in air is shown in Fig. 11. When the length of the radiated sound wave λ exceeds the circumference of the piston, $2\pi R$,

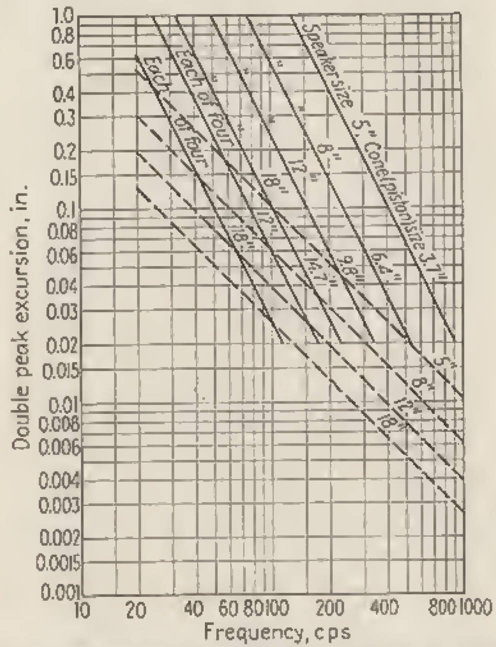


FIG. 12.—Total displacement required of diaphragm of indicated size to radiate one watt. Solid curves for pistons radiating into a hemisphere at low frequencies where the radiation resistance is proportional to the square of the frequency ($2\pi R/\lambda$ less than about 1.4 in Fig. 11). Dashed curves for constant radiation resistance of 41.5 mechanical ohms per square centimeter (exponential horn value well above horn cutoff frequency).

the radiation resistance is nearly proportional to the square of the frequency. In this frequency range the piston *velocity* should vary inversely with frequency to radiate constant power since this is equal to the square of the r-m-s velocity and the radiation resistance [see Eq. (5)]. This variation in velocity with frequency is usually obtained by placing the fundamental resonant frequency of the diaphragm and motor near the lowest frequency to be transmitted so the system has mass reactance or is "mass-controlled" in this frequency range. When the wave length is less than half this value, the resistance is very nearly

41.3 mechanical ohms per square centimeter and the diaphragm (real or virtual) is efficiently coupled to the air (see Diaphragms, Size).

When the length of the radiated wave exceeds the circumference of the piston, the air increases the apparent mass of each side of the diaphragm by approximately the mass of air contained in a cylinder whose base is the piston and whose height is 0.85 times the piston radius. At high frequencies the radiation mass ("accession to inertia") and the mass reactance decrease and approach zero for infinite frequency.

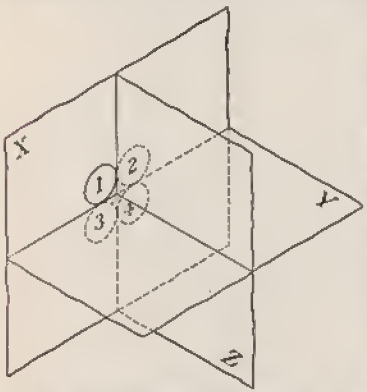


FIG. 13.—Primary images 2, 3, and 4 of piston 1 introduced by planes Y and Z.

7. Mutual Radiation Impedance. When a sound wave radiated from one surface of a diaphragm has access to another surface of the same diaphragm or to a surface of another diaphragm, there is said to be coupling between the surfaces. Consideration of this mutual radiation impedance is simplified by fixing attention on what occurs at each diaphragm. The motion of the diaphragm is opposed by the ("self-") radiation impedance. It is also opposed or aided by the force exerted on it by the waves generated by any other diaphragms which are coupled to it. The (complex) ratio of the force due to all other diaphragms to the velocity of the diaphragm itself is the mechanical impedance seen by the diaphragm due to the other diaphragms. This we will call the total mutual radiation impedance.

The total radiation impedance seen by a diaphragm is the sum of the self- and the "mutual" radiation impedances. The acoustic power W_a radiated by a diaphragm is

$$W_a = (r_S + r_M)V^2 \times 10^{-7} \text{ watt} \tag{5}$$

where r_S = real part of self-radiation impedance (total)
r_M = real part of mutual-radiation impedance (total)
V = r-m-s diaphragm velocity in centimeters per second.

Note that the velocity appears as current would in the corresponding electrical equation. The diaphragm displacement is $V/2\pi f$, where f is the frequency. The total displacement of various diaphragms required to radiate 1 watt is shown in Fig. 12. These curves clearly show the need for large diaphragms if appreciable low-frequency power is to be radiated.

By knowing the self- and mutual radiation impedances of diaphragms mounted in a single infinite baffle, we can determine the impedance seen when other baffles are added. In Fig. 13 assume four pistons (1, 2, 3, 4) mounted in the X-plane. Because of symmetry there is no net sound flux through the plane Y since for every positive vertical (z) component from pistons 3, 4 there is a negative component downward from pistons 1, 2. We may therefore introduce the rigid, thin, non-

absorbent plane or baffle Y without altering the impedance seen by any of the pistons. With Y in place we may remove pistons 3, 4, and pistons 1, 2 will continue to see the same impedance. The sound wave reflected by the plane corresponds exactly to the wave which would come from diaphragms 3, 4 and therefore the plane is said to have created "primary images" (by analogy with the optical case) of diaphragms 1, 2 which 1, 2 cannot distinguish from the real diaphragms 3, 4. Similarly the plane Z

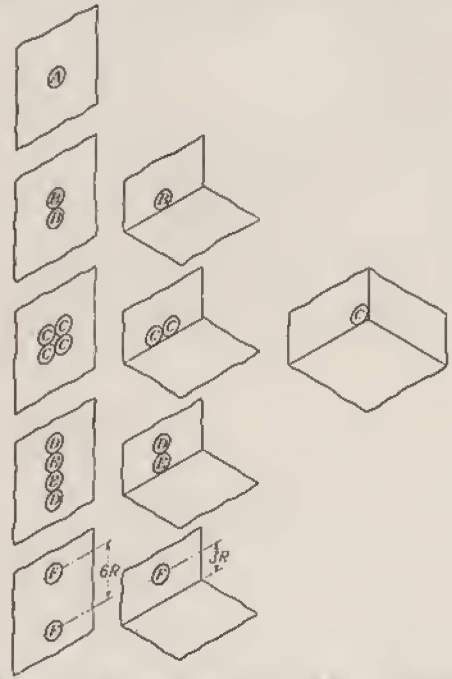


FIG. 14.—Effect of adding pistons and reflecting planes on radiation impedance. All pistons marked with the same letter see the same radiation impedance.

may be introduced and pistons 2, 3, 4 removed, leaving 1 looking into its original impedance. In all cases pistons of equal size, vibrating in phase and with the same amplitude in infinite, rigid non-absorbing baffles, are assumed. The relations hold approximately when the baffles are a wave length or more long. Finite impedance of a baffle may be treated by assuming reduced amplitude of the image to account for absorption and a change in phase to account for the reactive part of the impedance. The principle is readily extended to multiple sources of arbitrary size, phase, and displacement such as occur in vented enclosures, labyrinths, and the like.

Several piston combinations are shown in Fig. 14. All pistons marked with the same letter see the same radiation impedance. The ratio of the radiation resistance and reactance seen in each case to that seen by a single piston A is shown in Figs. 15 and 16 (after Klapman). The actual impedance is therefore obtained by multiplying the ordinates of Fig. 15 or 16 by the corresponding ordinate of Fig. 11. The letters on the curves correspond to those on the pistons in Fig. 14.

Values of the ordinate less than 1 indicate the piston sees less resistance or reactance than it would if alone in a single infinite plane. This

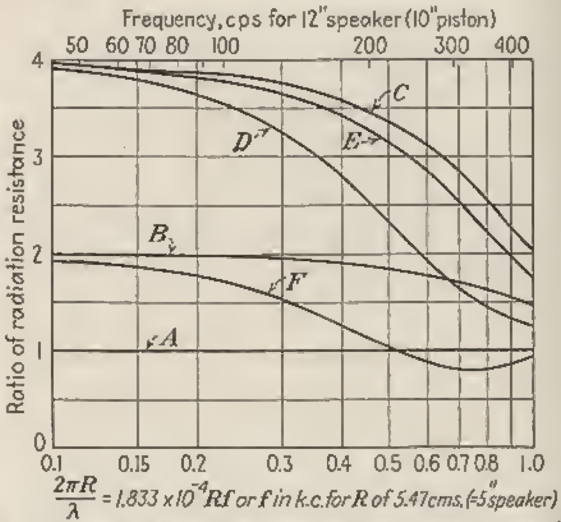


FIG. 15.—Ratio of radiation resistance seen by a piston in the presence of one or more others (real or images) vibrating with equal amplitude and phase to radiation resistance it would see alone radiating into a hemisphere. Designating letter of curve is same as that of corresponding piston in Fig. 14. (*After Klapman.*)

occurs when the time delay in the wave from one diaphragm and the frequency are such as to give out-of-phase components at the other.

Figure 15 shows that the radiation resistance is increased by the largest factor and over the greatest frequency range when each diaphragm is as close to all others as possible. At low frequencies the group then behaves as a single large piston. Figures 15 and 16 show that the radiation resistance increases much more rapidly than the reactance as speakers (or their images) are added. The exact improvement in efficiency which results depends on the impedance seen looking back into the speaker diaphragm but a typical speaker efficiency is increased by a factor of nearly 2 (3 db) for case B and 3.2 (about 5 db) for case C. This indicates qualitatively the improvement gained by operating a speaker at the intersection of the floor and wall and in a corner, respectively.

For a given diaphragm amplitude one speaker in locations B and C will radiate two and four times as much l-f energy, respectively, as one in location A. The radiated power for constant amplitude is therefore proportional to the square of the number of actual diaphragms. The radiation resistance at high frequencies is not improved by the use of additional speakers. A group of speakers therefore has better low but no better high response than a single one, and they therefore sound as though they had relatively less high response. Except for cost reasons multiple speakers are usually preferred to a single speaker with the same size motor because (1) the small diaphragms are lighter per unit area than a large one of adequate rigidity, giving better efficiency and high response; (2) the angle of individual speakers may be adjusted a moderate

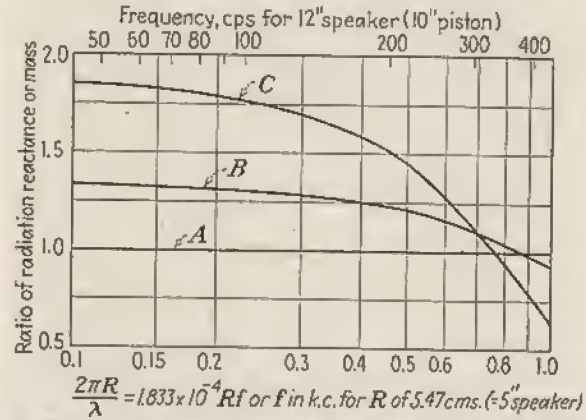


FIG. 16.—Ratio of radiation reactance (or mass) seen by a piston in the presence of one or more others (real or images) vibrating with equal amplitude and phase to reactance (or mass) it would see alone radiating into a hemisphere. Designating letter of curve is same as that of corresponding piston in Fig. 14. (*After Klapman.*)

amount to give a good h-f directional pattern without injuring the l-f response; (3) improved reliability, since failure of a single unit usually does not seriously affect the performance of the group; and (4) the temperature rise of each voice coil is reduced.

MULTIPLE LOUD-SPEAKERS

Some of the numerous advantages of multiple direct-radiator speakers where these all cover the same frequency range are discussed in Art. 7, Mutual Radiation Impedance, above. Multiple-speaker systems in which the speakers cover complementary frequency ranges also have certain advantages and are widely used. The more important advantages are (1) improved frequency response, since each type of unit covers a moderate range; (2) higher system efficiency, for the same reason; (3) improved directivity characteristic, since the diaphragm (or horn mouth) for the highest frequency range may be made relatively small (see Figs.

17, 17*a*, and 17*b*); (4) improved transient response, since many of the artifices used to obtain extended frequency ranges in single units make the transient response worse, particularly at high frequencies; (5) reduced intermodulation, since large amplitudes are confined to speaker reproducing low frequencies; and (6) reduced frequency modulation which occurs when a single diaphragm moves with large amplitude with respect to the listener, thereby altering the frequency (due to the Doppler effect).

8. Piston Directivity. With rising frequency the radiation from a rigid piston becomes increasingly concentrated on the axis, as shown in

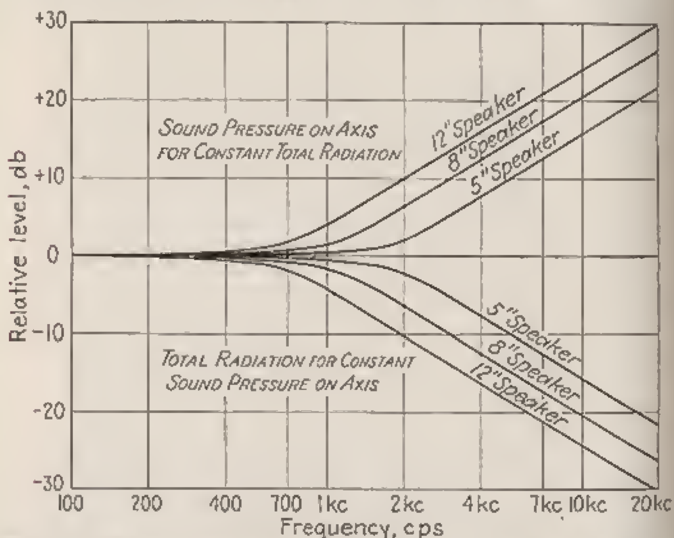


Fig. 17.—Effect of concentration of radiation on piston axis at high frequencies for the case of pistons radiating constant *total* power (top curves) and the case of constant pressure response on the axis (lower curves). Because of cone flexing the concentration is less marked in actual diaphragms where the equivalent piston diameter at high frequencies approaches about six-tenths the actual cone diameter.

Figs. 17 and 17*a*. Figure 17 shows that, if the flat axial pressure response curve so often sought after is obtained, the total radiation and therefore the efficiency are actually falling rapidly at high frequencies. Conversely, if a speaker is to have constant efficiency its axial pressure response must rise appreciably at high frequencies. The variation in relative response with angle up to the angle for which the first minimum occurs is shown in Fig. 17*a*. The response on the axis has been arbitrarily adjusted to the same reference level in all curves. At high frequencies the effective area of an actual cone is reduced by flexing, so that the directivity of actual cones is somewhat less than that shown for the piston.

Typical directional curves for 6- and 10-in. (designating size) speakers are shown in Fig. 17*b*. The axial response is assumed equalized to give

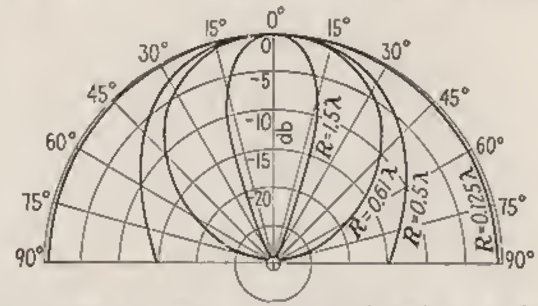


Fig. 17*a*.—Variation in relative response with angle up to the angle for which the first minimum occurs. The response on the axis has been arbitrarily adjusted to the same reference level in all curves.

flat response and the relative response for other angles is shown. Typical directional curves for a 6- by 9-in. (designating size) elliptical speaker

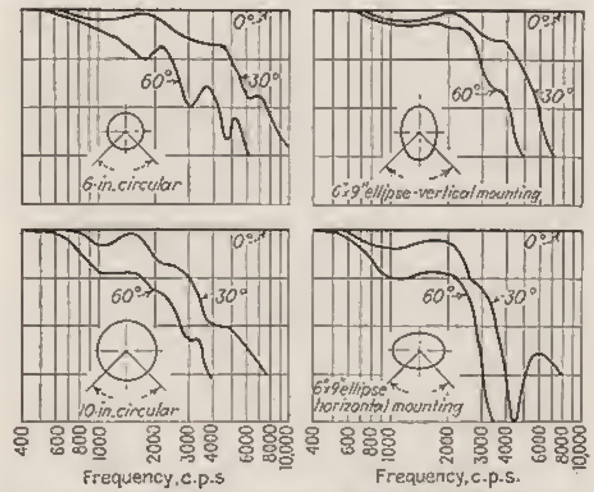


Fig. 17*b*.—Experimentally determined directional characteristics for two circular and one elliptical diaphragms. The elliptical diaphragm has the *broadest* characteristic in the horizontal plane when its major (long) axis is *vertical*.

are also shown. These show that the directional response of this elliptical speaker in the plane of the *minor* or short axis is roughly comparable to that of a circular speaker with a diameter equal to the minor axis. The directional response in the plane of the major or long axis is worse than that of a 10-in. circular speaker up to about 6,000 cycles. Above this frequency it is better. Contrary to popular belief the elliptical speaker should be mounted with its major axis *vertical* to get broadest distribution in the *horizontal* plane. This corresponds to the theoretical predictions of Stenzel. This same effect is present in rectangular mouth-shaped horns where the middle and middle h-f response is *most directional* in the plane of the *broadest* horn mouth dimension unless partitions, or separate cells are used. Even then the effect is present when the wave length is comparable to the smaller dimension of the entire mouth.

9. Horns. A horn is a tapered acoustical transmission line used to couple the impedance it sees, looking back into the diaphragm, as efficiently as possible to the load it sees looking out of its mouth (see Motors). The small end of a horn is called its *throat*, the large end its *mouth*. If its mouth has an infinite flange or baffle, the radiation impedance it sees is approximately the impedance given in Fig. 11. If there is no flange, the radiation resistance is half and the reactance approximately seven-tenths this value at low frequencies. At high frequencies the flange does not alter the impedance.

Exponential horns are usually employed because they provide more efficient coupling at low frequencies. Their cross-sectional area varies exponentially with length and is defined by the following relation:

$$S = S_0 e^{mx} \tag{6}$$

where S and S_0 are the areas of plane section normal to the horn axis at a distance x from the throat and at the throat, respectively; m is a constant which determines the rate of flare and theoretical cutoff frequency; and $e = 2.71828$. Curves showing the axial length of the horn for different area ratios and cutoff frequencies are given in Fig. 18.

The impedance per unit area seen at the throat of an infinitely long non-absorbing horn is

$$z_A = r_A + jx_A = \rho c\left(\sqrt{1 - \frac{m^2c^2}{4\omega^2}} + j\frac{mc}{2\omega}\right) \tag{7}$$

where ρ, c, and m have been defined and $\omega = 2\pi$ times the frequency. The total mechanical impedance seen by the diaphragm is $A_d z_A$, where A_d is the diaphragm area, assumed equal to the horn throat area. The radiated acoustic power W_a, assuming no absorption in the horn, is

$$W_a = V_d^2 A_d r_A \times 10^{-7} = V_d^2 \rho c A_d \times 10^{-7} \sqrt{1 - \frac{m^2c^2}{4\omega^2}} \text{ watt} \tag{8}$$

where V_d = r-m-s diaphragm velocity. The exponential horn behaves as a high-pass filter, since its input resistance is zero when ω is less than $mc/2$ and rapidly approaches a constant at higher frequencies. The theoretical cutoff frequency is $f_c = mc/4\pi$.

In a horn of finite length the outgoing wave does not see a radiation impedance, at the horn mouth, equal to the characteristic impedance of the infinite horn unless the length of the wave is approximately two-thirds the mouth diameter, or less. The wave is therefore partially

reflected at low frequencies. Partial reflection of the return wave also occurs at the diaphragm unless the impedance looking into the diaphragm (or sound chamber) equals the characteristic impedance of the infinite horn. This requirement can be met over a wide frequency range only with very efficient motors. These reflections result in maxima and minima in the throat impedance of the horn which become more severe as the horn mouth is made smaller. The actual cutoff frequency of most exponential horns is about 20 per cent above the theoretical. The horn-

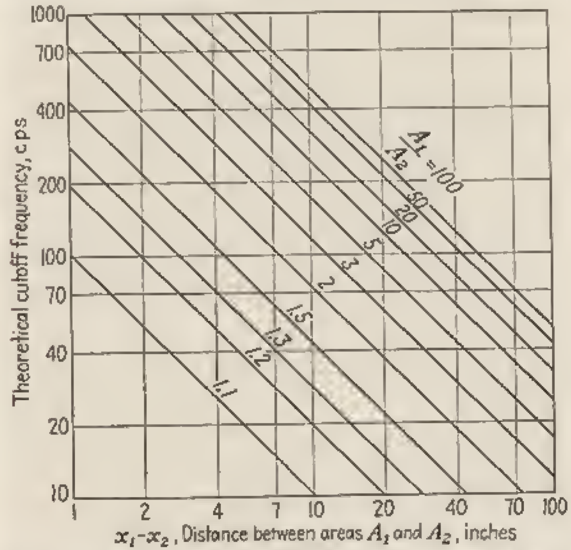


FIG. 18.—Distance along exponential horn axis between area ratios indicated on curves for theoretical cutoff frequency given by the ordinate. For example, the area of a 70-cycle horn doubles every 11 in. and the area of a 100-cycle horn increases 20 per cent every 2 in.

mouth diameter is usually made about one-third the length of the wave corresponding to this actual cutoff frequency.

If the impedance seen by the diaphragm and given by Eq. (7) is not high enough, an air "transformer" or sound chamber is used to increase it. The horn throat is then made smaller than the diaphragm. At low frequencies the impedance seen by the diaphragm is increased by a factor $(A_d/A_t)^2$, where A_t is the throat area. To maintain the radiation resistance seen by the diaphragm up to high frequencies, the sound chamber is usually divided in some manner or made narrow to avoid cancellation effects. Two of the more recent sound chamber constructions to achieve this are shown in Figs. 19 and 20. Figure 19 shows a dome-shaped diaphragm and a series of concentric circular slots. Figure 20 shows an annular trough-shaped diaphragm which minimizes the distance from any part of the diaphragm to the circular exit slot.

From Fig. 11 we note that a diaphragm looks into a radiation resistance of about 42 ohms per square centimeter when the peripheral length of the piston is more than twice the length of the radiated sound wave ($2\pi R$ greater than 2λ). This corresponds to frequencies higher than 1,000 and 2,000 cps for 10- and 5-in. speakers, respectively. From Eq. (7) we note that this is also the maximum resistance seen looking into an exponential horn unless a sound chamber is used. This accounts for the

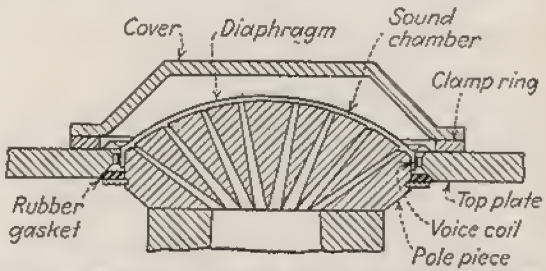


FIG. 19.—Cross section of a diaphragm and sound chamber in which concentric annular slots are used to reduce destructive sound chamber interference at high frequencies.

fact that the addition of a horn to a conventional large-diaphragm speaker helps the efficiency only below about 1,000 or 2,000 cycles, depending on the diaphragm size. This added efficiency is obtained down to a frequency from 10 to 30 per cent above the theoretical cutoff frequency of the horn. Below this the efficiency may be lower than it would be if the diaphragm were on a large baffle. The net effect is to make the unit sound more efficient but relatively deficient in high-frequency response.

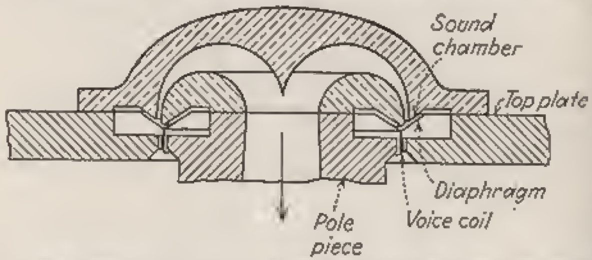


FIG. 20.—Cross section of annular V-shaped diaphragm and sound chamber used to reduce destructive interference in sound chamber at high frequencies.

The idea is prevalent that a long horn with a slow rate of flare is much more directional than a short one with a rapid rate of flare when both horns have the same mouth diameter. Theoretically and experimentally it has been found however that the directivity depends almost altogether on the mouth size and shape since the mouth becomes a "virtual" sound source. Near the l-f cutoff all horns become relatively non-directional

and they are made more directional by increasing the mouth diameter, not by making the horn longer. Minor effects occur at high frequencies but may usually be neglected compared with the l-f effect.

DIAPHRAGMS

10. Principle of Operation. The diaphragm is the part of the speaker which couples the radiation impedance to the speaker motor. In the usual hornless or direct radiator speaker the force exerted by the motor is localized, and this must be transmitted to the acoustic load which is spread over a large area. To do this effectively and to add as little as possible to the impedance, the cone is made as rigid and light as possible.

The usual conical diaphragm may be thought of as a continuous mechanical transmission line radiating acoustic energy from each element of area. Radial waves which travel from the driving point to the edge and are reflected and circumferential waves which travel around the cone, both occur in various combinations depending on the "mode of vibration."

The lowest frequency mode and the simplest one is the one in which the effective radial wave length of the cone, including the edge termination, is one-quarter wave. (This must not be confused with a quarter wave length in *air* at the same frequency.) At this frequency, which ranges from 700 cycles in large to 1,400 cycles or more in small cones, no circumferential wave is present, and all parts of the cone move in phase. The displacement is a maximum at the apex and a minimum at the flexible annulus which supports the outer edge and terminates the transmission line. The impedance of this termination plays an important part in the diaphragm behavior, especially at frequencies near the fundamental resonance of the diaphragm and motor and in the 1,000- to 2,000-cycle range.

At frequencies below the lowest mode of the cone itself all parts of the cone move in phase, and the cone behaves approximately as a piston unless the annulus stiffness increases rapidly with displacement, in which case the cone may flex at even very low frequencies. The annulus is frequently made this way deliberately in inexpensive speakers to produce distortion of low frequencies and substantially increase their loudness by radiating most of the energy at harmonic frequencies. Unfortunately, intermodulation of low and high frequencies then also occurs, which makes the high end sound rough or garbled when a strong low note is reproduced.

11. Size. It has been found experimentally that at low frequencies the effective area of the cone is its projected or base area. This is approximately the "cone" size where this is defined for a circular cone as "the diameter to the nearest quarter-inch of the minimum circle determined by the tangency of the cone and a plane touching its base."[1] This is not to be confused with the *designating* size of a loud-speaker which is commonly used in describing a speaker.

"The designating size of a loudspeaker employing a circular radiator shall be twice the maximum radial dimension, measured to the nearest eighth-inch, of the front of the speaker except that the designating size shall not exceed the maximum diameter of the unsupported portion of the vibrating system by more than 25%."[2]

[1] Radio Manufacturers Association, definition M5-111.
[2] Radio Manufacturers Association, definition M5-110.

This definition is intended to limit the amount of functionally useless cone housing included in the designating size. Representative cone sizes for various speaker-designating sizes are shown in Fig. 12.

In direct radiator speakers and at low frequencies the radiation resistance is proportional to the fourth power of the radius (square of the area) and the reactance to the cube of the radius. The resistance-reactance ratio or power factor of the radiation impedance is therefore proportional to the piston radius. For constant radiated power the piston displacement varies inversely with area. With fixed amplitude the radiated power is proportional to the square of the area at a given frequency, or a frequency one octave lower may be reproduced if the area is increased by a factor of four. The upper limit to diaphragm size is set by the increased weight *per unit area* required to get a sufficiently rigid structure. The nature of the acoustic load (horn, enclosure,

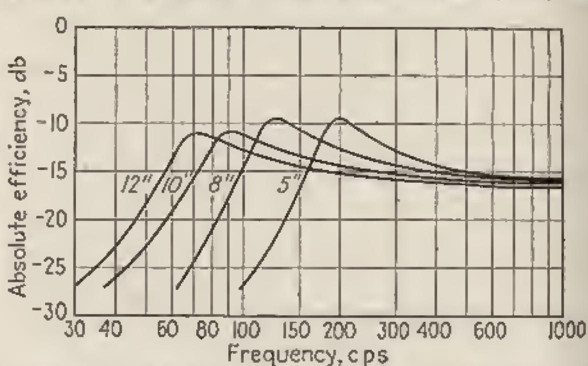


FIG. 21.—System efficiency of four speakers using identical moving-coil motors but different cone sizes (calculated). Designating size of speaker is given.

cabinet, etc.), space limitations, cost, and the motor employed also control size.

It is customary to increase the size of the motor as the diaphragm size is increased, since the mechanical impedance looking back into the motor (voice coil, say) should go up as the impedance looking into the driving point of the cone rises to maintain good energy transfer. If a large cone is put on a small motor, the displacement and distortion for a given acoustic output drop and lower frequencies may be reproduced with the same distortion, but the efficiency in the mid-range may actually drop. These effects are illustrated in Fig. 21, in which the calculated system efficiency of four speakers using different size cones but the same motor are shown. Minimum cone weights, found to be satisfactory experimentally, and average mechanical resistance and resonant frequencies were assumed. The motor is an intermediate size normally employed on 8-in. speakers but frequently used on all four diaphragm sizes. Speaker efficiency, even at low frequencies, is therefore not limited by cone size. The cone size must be large, however, if appreciable power is to be radiated with reasonable cone excursions at low frequencies.

12. Shape. The most efficient shape at low frequencies is circular. This is also the most satisfactory structurally. Theoretical and experimental investigations have shown that an ellipse with a major to minor axis ratio of two, and a two-to-one rectangle have an average of 5 and 7 per cent lower radiation resistance in the useful l-f range than a circle of the same area. The loss is progressively greater as the shape departs still further from circular. In spite of the appeal of elliptical and other diaphragm shapes, which were used in early magnetic speakers and even in more recent European moving-coil speakers to a limited extent, their disadvantages have prevented their general adoption.

At high frequencies all pistons have the same radiation resistance per unit area, but most cones cannot be considered pistons, both because they are not flat and because their radial length exceeds a quarter wave and flexing is therefore important.

The shape of the cross section or profile of the cone depends on the application and response desired. Straight side cones are usually employed when good 2,000- to 5,000-cycle response is required and when reproduction above 6,000 to 7,000 cycles may actually be undesirable. This is frequently the requirement of public address and phonograph systems where noise and distortion are otherwise objectionable. Curved cones improve the response above 6,000 to 7,000 cycles by providing a diaphragm impedance, viewed from the voice coil, which has a more uniformly high negative reactance and therefore absorbs more power from the high positive reactance (due to the voice-coil mass) seen looking back into the voice coil. This improvement is obtained at the expense of 2,000- to 5,000-cycle response and with a weaker cone structure, with the result that straight cones predominate by ten or more to one in actual use.

13. Material. Hard, impregnated or filled, and pressed or calendered papers are used when loudness efficiency and apparent h-f response are important. Radiation resistance provides very little dissipation in direct radiator cones; hence, by using a paper having low internal flexural losses, the conical transmission line is made to have strong resonances. Nearly all speakers now use material of this type. The transient response of diaphragms of this type is necessarily poor since non-center moving modes of the cone are inappreciably damped by the motor. Soft, loosely packed, or felted blotterlike cones are used when some loss in h-f response can be tolerated and a smoother response curve with reduced transient distortion is required. The loudness efficiency of high-loss cones of this type is several decibels lower than that of low-loss cones.

Felt, leather, rubber, and similar materials are used as the annulus to terminate the conical transmission line in a low-stiffness high-resistance material. Their effect is to add considerable dissipation to the cone at the termination, resulting in reduced reflection of the flexural wave. The effect is similar to that obtained in soft cone materials where, however, the dissipation is distributed along the line. The objection to leather is that it is very sensitive to changes in humidity, resulting in mechanical cone alignment problems unless adequate air-gap clearances are provided.

14. Breakup Subharmonics. The term *cone breakup* is sometimes applied to the flexing or wave-transmission process in a cone. Since there is nothing discontinuous in the process to suggest the word "break" and to avoid ambiguity, it is suggested that this term be applied only in the other sense in which it is used, to name the process which results in the generation of subharmonics.

If the apex of a cone is driven with an adequate sinusoidal force at certain critical frequencies, the radiated wave contains not only the

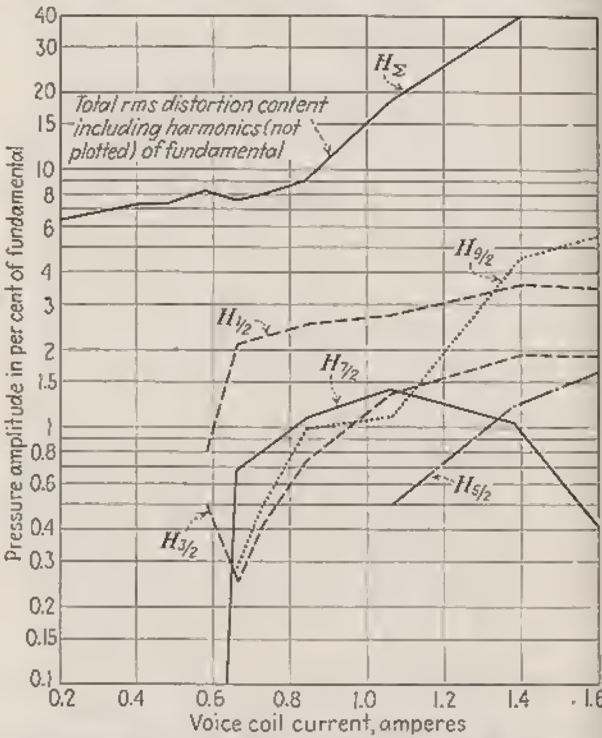


FIG. 22.—Variation in radiated distortion products with voice-coil current showing abrupt start of half frequency (subharmonic) and odd multiples of it. Subscript indicates factor by which fundamental frequency is multiplied to obtain frequency of indicated distortion product.

fundamental and integral multiples (harmonics) of it but also a frequency corresponding to half (and rarely to a quarter) that of the fundamental and integral multiples of this subharmonic. While distortion in the motor may contribute to this, some unpublished research has shown that the cone is the important source.

The half frequency appears very suddenly at a critical input as shown in Fig. 22. To simplify the graph, the fundamental and usual harmonics which would include even multiples of the subharmonics, are not included. From the total (r-m-s) harmonic curve H_Σ which includes these unplotted

terms, we see that negligible rise in total distortion occurs when the subharmonic begins. The ear, however, reports a large increase because the pitch sense of the output has dropped an octave and the distortion has a high annoyance or objection ability factor. This type of distortion is not so important as is frequently supposed, however, because (1) it occurs only in limited frequency regions; (2) it does not occur below a moderate, critical level; (3) the time required to start it is large unless the force is large; (4) the spectral composition of speech and music are such that the probability of its production is small. Because of the statistical improbability of its frequent occurrence, it would be uneconomical to design most systems to avoid completely this occasional distortion.

MOTORS

A loud-speaker motor converts electrical into mechanical energy and couples the electrical signal source as efficiently as possible to the mechanical impedance seen looking into the diaphragm which it drives.

15. Force Factor. The mechanical circuit of a speaker motor experiences a force when a current is applied to the electrical terminals. The (complex) ratio of this force when the mechanical circuit is blocked (infinite impedance) to the current which produces it is the force factor. Since force is analogous to voltage, the force factor is analogous to mutual impedance between two electrical circuits. It differs from the conventional electrical mutual impedance in that it makes no contribution to the electrical impedance when the mechanical circuit is blocked (secondary open-circuited) because its counter e.m.f. is due only to motion of the mechanical circuit, and in that the force factor has opposite signs when viewed from the electrical and mechanical circuits. In usual circuit notation $z_{12} = -z_{21}$ (not $z_{12} = z_{21}$ as in the electrical case). Since only the product of the force factors looking in both directions is involved in the following equations this will be called M^2.

The normal impedance of a speaker is defined as the impedance measured or seen at its signal terminals when operating normally with its proper acoustic load. The normal impedance z_N of moving coil and magnetic armature speakers is

$$z_N = z_e + \frac{M^2}{z_m} \tag{9}$$

where z_e = blocked impedance of the speaker

z_m = the total mechanical impedance seen by the mechanical circuit including diaphragm and acoustic load.

In moving-coil speakers $M^2 = B^2l^2$, where B is the average radial flux density which the coil embraces and l is the conductor length. In balanced magnetic armature speakers $M^2 = 4B_0^2N^2/R^2$, where B_0 is the steady flux density in the gaps, N is the number of turns on the voice (armature) coil, and R is the effective reluctance of the alternating flux path (see Magnetic Armature, Art. 19).

A two-terminal load impedance absorbs maximum power from a two-terminal source when the impedance of the load is the conjugate of the impedance measured or "seen" at the source terminals. The conjugate impedance is one having the same resistive or real part and a reactive or imaginary part equal in magnitude but opposite in sign. This holds for acoustical and mechanical circuits as well, but in these the terminals are not always so readily determined.

The speaker motor therefore absorbs maximum energy from the source, regardless of the complexity of the source network, when its normal impedance is the conjugate of the source impedance. The usual

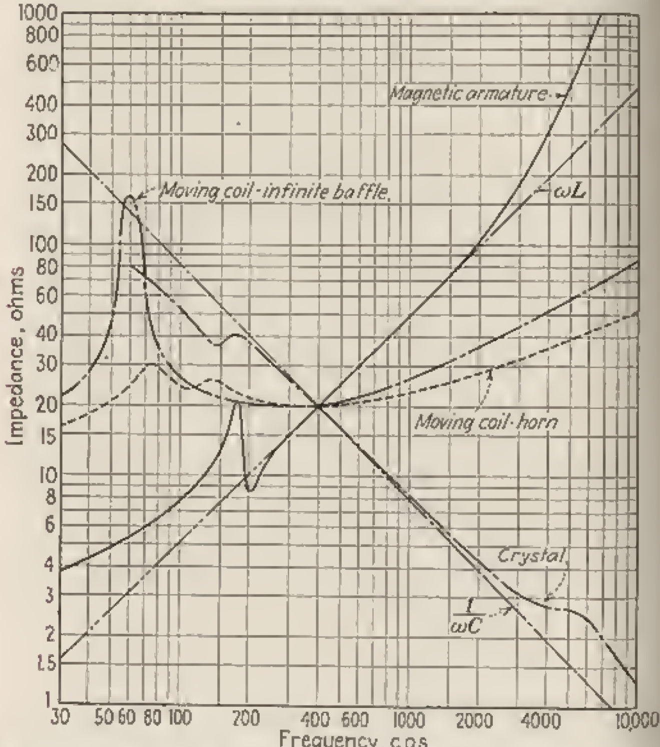


FIG. 23.—Magnitude of normal input impedance of various speakers all adjusted to same value at 400 cycles to simplify comparison. Magnitude of reactance of pure capacitance and inductance shown for comparison.

source is a vacuum tube, and its associated loud-speaker should ideally have a normal impedance which is a constant resistance. From Fig. 23 we see that this requirement is most closely met by moving coil or dynamic speakers.

16. Magnetic Motors. By the I.R.E. definition, "A magnetic speaker is a loud speaker in which the mechanical forces result from magnetic reactions." This includes both moving-conductor or moving-coil (dynamic) and magnetic-armature speakers.

17. Moving Coil. A moving-coil motor is one in which the mechanical forces result from magnetic reactions between the field of the moving coil and the applied steady radial field in the air gap. A section of half of a moving-coil speaker is shown in Fig. 24.

Moving-coil motors are now used almost exclusively because (1) their electrical impedance permits good energy transfer from the source, (2) the large amplitudes required by the popular direct radiator diaphragms are obtained conveniently with minimum non-linear distortion, (3) the mechanical impedance of the moving element may be made low, (4) the structure is simple and rugged mechanically, and (5) the cost is low.

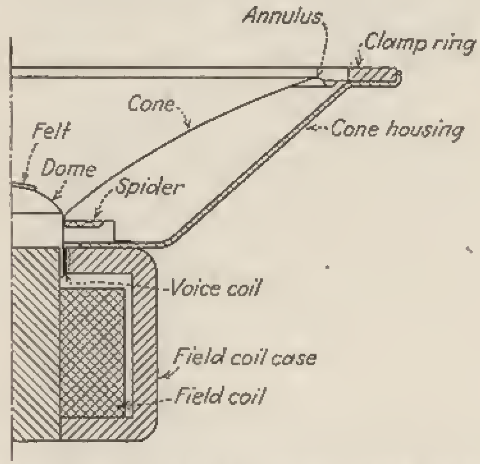


FIG. 24.—Sectional view of small moving-coil speaker showing structural simplicity.

The impedance seen at its electrical terminals when the coil is rigidly blocked is called the *blocked impedance* of the motor or speaker. This is approximately a high resistance and low inductance in series (R_e and L_e in Fig. 25) and is therefore easily coupled to a vacuum tube. Near the fundamental resonance of the speaker the impedance rises, and, if a low impedance source is used, the mismatch reduces the energy absorbed.

Moving-coil speakers are sometimes called *electrodynamic* or briefly *dynamic* speakers. Both terms have been applied for many years to speakers having either electromagnet (or "energized") or permanent magnet fields. The prefix "electro" in electrodynamic has nothing to do with the source of steady flux in the gap.

18. Permanent and Electromagnets. Magnetic speakers require a source of magnetomotive force to provide steady flux. If the current source is hum free and therefore the flux absolutely steady, the voice coil cannot distinguish between a given flux density due to permanent and electromagnets. The efficiency of any electromagnet speaker can be equaled or excelled by a permanent magnet if cost is neglected. In

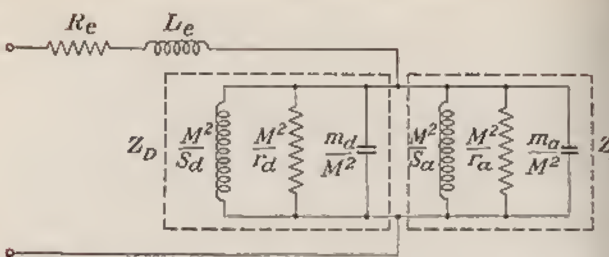


FIG. 25.—Equivalent l-f electrical circuit of moving-coil or magnetic armature speaker in a total enclosure or in an infinite baffle. In the latter case the enclosure stiffness S_a is zero and its equivalent electrical inductance infinite.

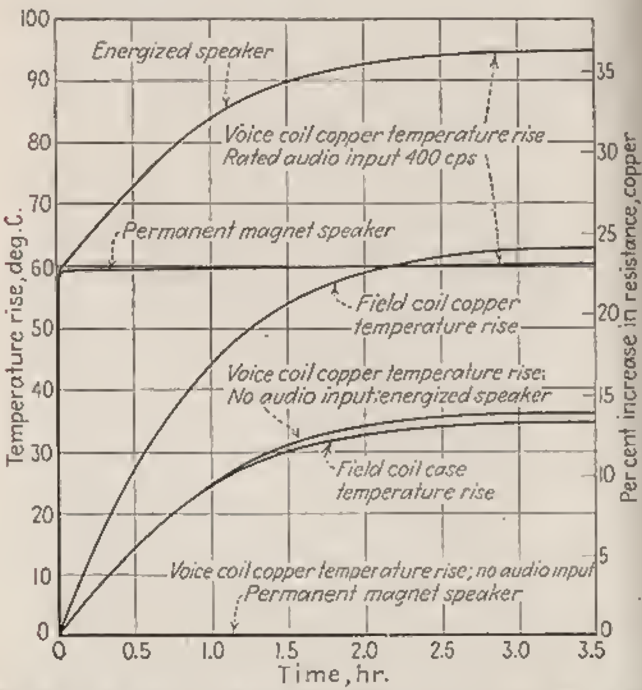


FIG. 25a.—Reduced voice-coil temperature rise in permanent-magnet speakers compared to energized types. Abnormal voice-coil temperature rise when rated "complex-wave" input is applied at 400 cycles also shown.

small speakers the differential in cost between permanent and electromagnet types is small even when no current supply source cost is added to the electromagnet type. In intermediate sizes the cost of the two types is comparable if the cost of a source of field power is included. In larger speakers the permanent magnet type is more expensive. The installed cost, however, of permanent magnet types is frequently lower because of simplified wiring.

The trend is toward the use of permanent-magnet speakers particularly when a special field current supply must otherwise be provided. The temperature of the electromagnet and consequently that of the voice coil rises with time as shown in Fig. 25a. The field coil resistance rises, lowering the field current and flux density. The higher voice-coil impedance and reduced flux reduce the speaker efficiency. The higher voice-coil temperature reduces the permissible signal input power in voice-coil temperature-limited speakers.

The temperature rise when the rated complex-wave input is applied at a single frequency (400 cycles) in a typical intermediate size radio speaker is also shown. The single-frequency rating is normally much less than the "complex-wave" (speech and music) rating since in the latter case advantage is taken of the high ratio of peak to average power (see Arts. 1, 2, 35).

19. Magnetic Armature. "A magnetic armature speaker (or motor) is a magnetic speaker (or motor) whose operation involves the vibration of the ferromagnetic circuit." The shorter term "magnetic" may be used where no confusion will result with moving conductor or moving-coil speakers, which are also by definition magnetic speakers. A cross-sectional view of a balanced armature motor of this type is shown in Fig. 26. Flux increases in one pair of pole faces and decreases in the other pair, when current flows through the voice coil and when the armature moves, resulting in operation analogous to a push-pull tube circuit. The voice coil does not move and therefore is made relatively large. The resulting high inductance plus distributed capacitance in high impedance types accounts for the large rise in impedance at high frequencies (see Fig. 23). This makes it difficult to couple it to a tube properly. To get high efficiency the armature pole piece clearance must be small, and this leads to instability of the armature and a limitation on its displacement. These factors plus mechanical difficulties in construction and maintenance have reduced the acceptance of the magnetic-armature type.

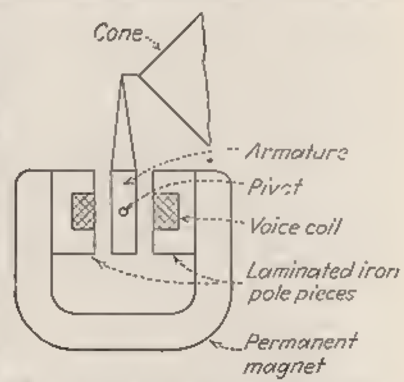


FIG. 26.—Sectional view of balanced armature magnetic speaker.

20. Condenser. "A condenser speaker (or motor) is a speaker (or motor) in which the mechanical forces result from electrostatic reactions." They are really large condensers in which one flexible electrode is free to move and to act as a diaphragm. In push-pull types the flexible electrode is mounted between two perforated fixed electrodes resulting in cancellation of the even harmonics which occur in the two-plate type. Its blocked impedance is that of a condenser, and it is therefore difficult to couple to a vacuum tube. The electrode clearance must be small or the steady polarizing potential which is applied must be large to get high efficiency. The former limits the diaphragm amplitude, and the latter causes rapid disintegration of any flexible dielectric used to support the electrode.

21. Crystal. "A crystal speaker (or motor) is a speaker (or motor) in which the mechanical forces result from the deformation of a crystal having converse piezoelectric properties." The crystal has a high mechanical impedance viewed from the driving point. Only a small displacement is possible without distortion or crystal fracture, so a mechanical transformer or lever arm is used when moderate excursions are required. This leads to mechanical complications, particularly at high frequencies where the transformer is not ideal, and to added cost. The application of this type has therefore been largely limited to h-f speakers in which the diaphragm amplitude is small. The blocked impedance is that of a leaky condenser. The normal impedance of an 8-in. unit is given in Fig. 23.

COMPLETE LOUD-SPEAKERS

The more important characteristics of a complete speaker system, which includes an electrical source of known impedance, motor, diaphragm, and known acoustic load, are its efficiency-frequency (including response frequency and impedance), directional, and distortion characteristics (see Tests).

22. Efficiency-frequency Characteristic. The energy efficiency, or simply efficiency, of a loud-speaker is the ratio of the useful acoustic-energy output to the signal-energy input. The "absolute" or system efficiency is the ratio of the useful acoustic energy output to the signal energy an ideal load would absorb from the signal source. The latter definition is a practical one in that it penalizes the speaker for its inability to absorb maximum power from the source. At a resonant frequency of a speaker the two efficiencies frequently differ by a factor of 10 or more.

If the effective internal resistance of the source and its ideal resistance load (both seen from the voice coil) are r_s, then the absolute efficiency is given by

$$\text{Absolute efficiency} = \left|\frac{4r_s M^2 r_r}{[z_E + (M^2/z_m)]^2 z_m^2}\right| \qquad (10)$$

where z_E = blocked voice coil impedance plus r_s
z_m = total mechanical impedance of the mechanical mesh including diaphragm radiation and air load
r_r = total radiation resistance seen by the diaphragm
M is defined under Force Factor. (See Art. 15.) The vertical lines indicate that the absolute value is to be taken.

The 400-cycle system efficiency of the speakers commonly used in radio receivers ranges from 1 to 4 per cent. The corresponding efficiency of direct-radiator speakers with very large motors ranges from 10 to

30 per cent. Efficiencies of this order are more readily obtained in horn speakers, but 30 per cent is rarely exceeded over any appreciable frequency range. Although higher values are frequently claimed, these values, if based on any measurements, are usually based on motional impedance measurements in which all horn, diaphragm, air, eddy-current, and hysteresis losses have been assumed to be useful acoustic radiation.

23. Response-frequency Characteristic. If a loud-speaker is to be used indoors, a graph showing the efficiency-frequency characteristic is probably the most useful single curve. If a loud-speaker is to be used outdoors, then we are primarily interested in its pressure response-frequency characteristic (see Tests).

24. Baffles, Enclosures, and Cabinets. "A baffle is a partition which may be used with an acoustic radiator to increase the effective length of the acoustic transmission path between front and back of the radiator." This term is usually reserved for a relatively flat baffle in which both sides of the diaphragm look into substantially a hemisphere (solid angle of 2π steradians). The term *directional baffle* is sometimes applied when one side of the diaphragm looks into a smaller solid angle. The baffle then begins to take on the properties of a horn. There is no sharp line of demarcation, but there appears to be little reason for calling any structure which restricts the solid angle to less than $\pi/2$ (an octant of a sphere) anything but a horn.

If a baffle is used outdoors, appreciable destructive interference or pressure cancellation between the front and back waves of the speaker may occur at the listener's position at some frequency above the cutoff frequency. The frequency at which this occurs depends on the baffle size and listener location. Destructive interference at the cone itself is usually unimportant except near the cutoff frequency of the baffle. To distribute this effect and make it cover a broad band, baffles shaped as shown in Fig. 27 are sometimes used. Since the effect depends primarily on the listener's location, no such simple result occurs indoors and a space average of the pressure in a moderate-size listening room shows no such effect. Conventional rectangular baffles may therefore be used indoors unless the room approaches free field or outdoor characteristics.

FIG. 27.—Irregular baffle shapes used outdoors to broaden frequency band of destructive interference between speaker front and back waves at listener's position.

The equivalent l-f electrical circuit of a moving-coil or magnetic-armature speaker in an infinite baffle is shown in Fig. 25. Here R_e and L_e are the blocked voice-coil resistance and inductance. Z_D is the electrical equivalent of the diaphragm less air load. Z_A is the equivalent of the air load, except that in this case there is no stiffness S_a provided by the air load, so its equivalent inductance M^2/S_a is infinite. M is defined under force factor S_d, m_d and r_d are the effective diaphragm stiffness, mass, and resistance, respectively, and r_a and m_a are the radiation resistance and mass which may be determined from Fig. 11. Note that both sides of the diaphragm have radiation resistance and mass in this case and the values *per unit area* given in Fig. 11 must be multiplied by twice the piston area to give the r_a and m_a used in Fig. 25.

The magnitude of the impedance of a moving-coil speaker in an infinite baffle is shown in Fig. 23. The antiresonant impedance of the parallel circuit corresponds to the resonant frequency of the diaphragm and air load and is limited by the parallel value of the two resistances. These resistances are proportional to the square of the flux density and inversely proportional to the diaphragm and air (radiation plus dissipation) resistances. A high resonant impedance is therefore not necessarily undesirable, as is generally supposed, since it may be due to a high flux density and therefore mean a high efficiency over a wide frequency band.

The effect of the source impedance, connected to the input terminals, on response may be noted from this circuit. If the source resistance r_s is low, the speaker will absorb very little power at resonance and the acoustic output may not rise appreciably. The voice-coil resistance R_c and the source resistance r_s in series are effectively across the antiresonant circuit at low frequencies since the reactance of L_c can be neglected. When the flux density is high, the Q of the antiresonant circuit alone is high but the source and voice-coil resistances then provide substantial shunt resistance and "electromagnetic damping." The effect of this on the speaker response to a pulse is shown in the experimentally determined curves of Fig. 40. The minor irregularities in these damped sinusoids are due to h-f modes of vibration of the diaphragm. As the source resistance is raised, more power is supplied the speaker at antiresonance and at high frequencies where the effect of the voice-coil inductance is important and less is supplied in the mid-frequency range. The steady-state response of a vented enclosure (see Figs. 28, 29, 29a) as the source resistance or "impedance match" is changed is shown in Fig. 39. Here the response at 400 cycles has arbitrarily been adjusted to the same value as that to which the source resistance was raised.

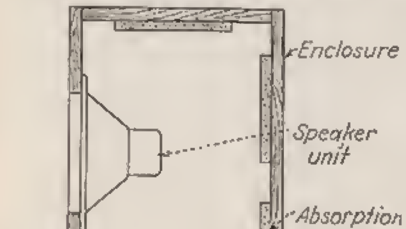


FIG. 28.—Bass reflex type of vented enclosure in which port area is large and placed near diaphragm to obtain maximum aid from mutual radiation resistance between diaphragm and port. Phase shift of back-side radiation obtained by choice of circuit constants.

25. Total Enclosure. A total enclosure which prevents radiation from the back side of a diaphragm may be used to prevent destructive interference between the front and back waves from a diaphragm. This might be obtained by closing the vent or port in Fig. 28. When the wave length exceeds four times the maximum enclosure dimension, the enclosure adds a *total* stiffness S_a viewed from the diaphragm. $S_a = \rho c^2 A_d{}^2/V_0$ cm per dyne, where A_d is the effective piston area of the cone and V_0 is the equilibrium volume of the box. The "capacitance" is the reciprocal of this value. This stiffness raises the natural frequency of the speaker. If the enclosure includes absorbing material, this stiffness will be altered by the reactance seen at the surface of the material. Each square centimeter will dissipate $P^2 \times 10^{-7}/r_B$ watt, where P is the sound pressure in the box and r_B is the resistance per unit area for sound of normal incidence on the absorbing material. From this the equivalent resistance in parallel with the box stiffness may be obtained.

If the volume is small enough or the natural frequency of the speaker out of the enclosure low enough, the enclosure and not the diaphragm stiffness will control the natural period.

The l-f equivalent electrical circuit of such an enclosure is shown in Fig. 25. Here z_D is the electrical equivalent of the diaphragm alone; S_d, r_d and m_d are the stiffness, resistance, and mass of the diaphragm measured *in vacuo*. The electrical equivalent of the air load including radiation impedance is z_A; S_a is the effective enclosure stiffness, r_a the total air or fluid resistance (enclosure dissipation if any, and radiation resistance), and m_a is the effective air (radiation-plus-enclosure) mass. Normally the parallel value of z_D and z_A or a single parallel "antiresonant" circuit is shown but the contributions of individual elements are then not as clear. Since the electrical circuit elements are inverses with respect to M^2 [see Eq. (9)], the stiffness appears as an inductance and the mass as a capacitance.

A total enclosure is sometimes called an *infinite* baffle. While it resembles one in preventing front and back wave interference, it has two important differences which make this designating term undesirable. The enclosure adds an air impedance to the rear of the diaphragm, which may be very different from that seen in an infinite baffle. An infinite baffle restricts the radiation to a hemisphere, and the radiation impedance seen by the diaphragm is given by Fig. 11. If the enclosure is used outdoors, the radiation resistance which the outside of the diaphragm sees at low frequencies is only half this value and the reactive part approximately seven-tenths this value. In practice the useful efficiency is almost halved at low frequencies. Indoors the impedance seen will depend on the environment as described in Art. 5, Radiation Impedance; also in Room Acoustics, below.

26. Vented Enclosures. The idea of putting a vent or "port" in an enclosure is very old. It was first done to provide "pressure relief." In more recent types, known as bass reflex enclosures, detailed consideration has been given to the very important effect of the mutual impedance between the port and diaphragm. The port area is large and the port is near the diaphragm to increase the mutual radiation resistance and extend the frequency range over which it is effective (see Art. 7, Mutual Radiation Impedance). Such an enclosure is shown in Fig. 28. The effective or virtual diaphragm in the opening is coupled through the stiffness of the air in the enclosure to the diaphragm. The equivalent l-f circuit is shown in Fig. 29. z_D corresponds to Fig. 25 and r_a and m_a correspond except that the mutual-radiation impedance must be added.

The vent and enclosure have therefore added one LRC circuit. The effect of this is to shift the back-side-cone radiation by nearly 180 deg. above the frequency at which the port mass m_p and box stiffness *viewed from the port* are resonant when the cone is blocked. For about one-third of an octave above and below this frequency most energy is radiated by

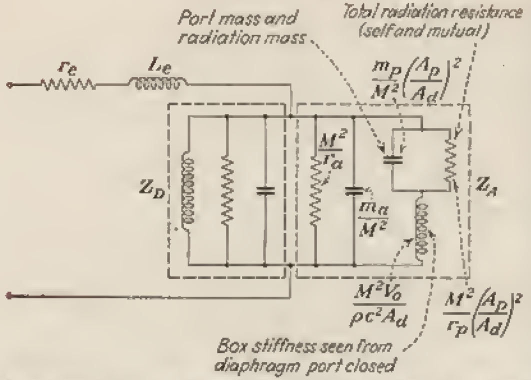


FIG. 29.—Equivalent l-f electrical circuit of moving-coil or magnetic-armature speaker in vented enclosure.

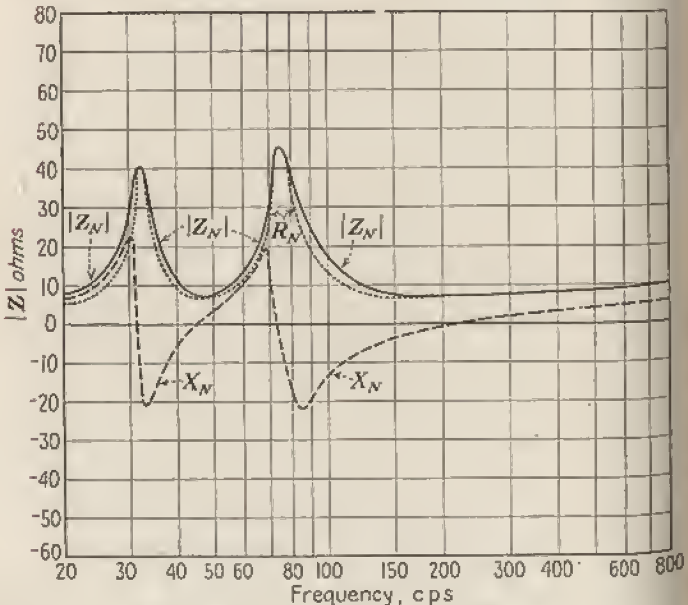


FIG. 29a.—Impedance of circuit of Fig 29.

port. Although the diaphragm and port radiation are out of phase low this frequency, the port radiation greatly exceeds the diaphragm diation near this frequency.

The enclosure is made as compact as possible. The port can be ced near the diaphragm to increase the mutual-radiation resistance ce the phase shift is not due to transmission time delay but occurs cause the acoustic circuit goes through antiresonance, the phase ift occurring suddenly at this frequency. In properly designed closures, advantage is taken of a large mutual-radiation resistance to prove the l-f efficiency. Very little absorption in the enclosure is

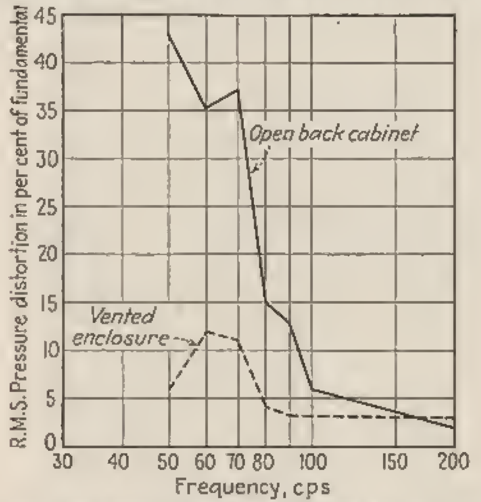


FIG. 30.—Total distortion of open-back cabinet and of the same cabinet closed as shown in Fig. 28, same speaker and electrical input in both ses. L-f distortion much reduced because diaphragm sees high antiresonant pedance of enclosure and therefore has only small displacement whereas in port (which lacks the non-linear edge stiffness and non-uniform flux the speaker) moves with large displacement.

anted at low frequencies to take maximum advantage of back-side diation. At frequencies of several hundred cycles or more where the rt radiates negligible sound the enclosure is made absorbent to avoid box" resonance. The advantages of vented enclosures are (1) backle radiation is used to substantially increase the l-f output; (2) most this output comes from the port which has no non-linear cone susnsion stiffness to produce non-linear distortion; (3) antiresonance of e enclosure occurs near the lower frequency of maximum radiation the diaphragm amplitude is much less than it would be otherwise. he result of these factors on non-linear distortion reduction is shown in g. 30 in which the effect of converting an open-back cabinet to a vented rt enclosure of the same internal volume is shown. The change in sponse is shown in Fig. 31.

27. Transmission-line Speaker. The phase and amplitude of the back-side radiation of a cone may be altered by coupling a conduit or

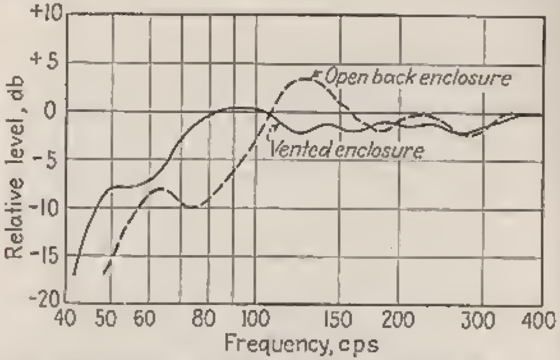


FIG. 31.—Total sound power radiated by cabinet for which distortion is shown in Fig. 30. Shape of l-f response may be varied between wide limits depending on enclosure volume, port area, and speaker used.

acoustic transmission line to it. In early types the multiple-resonant properties of such a line were used to influence the response. In a more

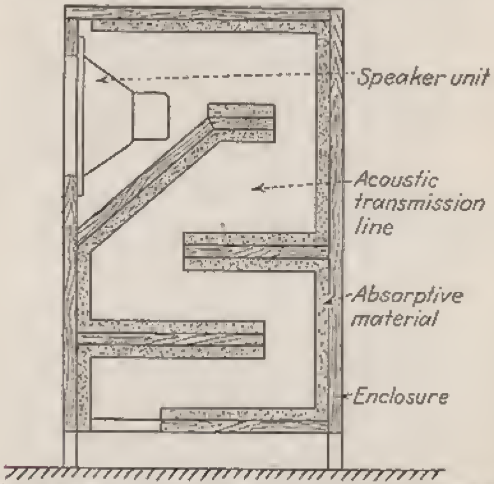


FIG. 32.—Labyrinth type of transmission-line speaker. Phase shift of back-side radiation obtained by time of transmission delay in line.

recent type, known as the labyrinth, the line is folded to conserve space and made highly dissipative (see Fig. 32). Phase shift between the

diaphragm and port or open end is due to time of transmission in the line. At very low frequencies the line is a small fraction of a wave length long, the phase shift is negligible, and the port and diaphragm radiation are out of phase. When the line is a quarter wave long, it acts as an impedance inverter (as in the electrical case); the cone sees a high impedance, and the radiation from the port is a maximum. Non-linear distortion is therefore reduced at and near this frequency. The resonant frequency of the diaphragm may be placed at this frequency to aid damping. Between this frequency and the one for which the line is a half wave length long, the port phase shifts gradually but maintains some component of its radiation in phase with the diaphragm (neglecting separation between the port and diaphragm) outside the line. Because of the infinite series of resonant and antiresonant frequencies of the line high

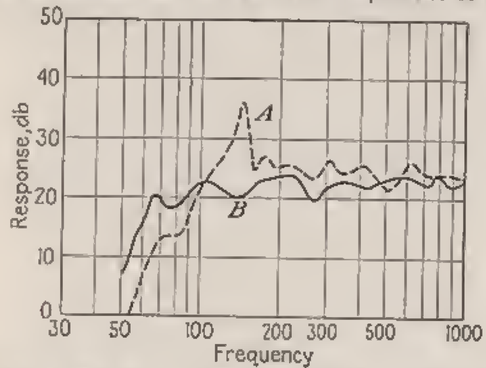


FIG. 33.—Relative response of open-back cabinet (A) and labyrinth (B). (*After Olney.*)

absorption must be introduced to prevent the production of objectionable resonances and radiated out-of-phase components of the port. Most of the rear-side radiation is therefore absorbed. The comparison of the response of an open-back cabinet and labyrinth is shown in Fig. 33 (after Olney).

ROOM ACOUSTICS

28. Room Characteristics. The trend in the theory of room acoustics is toward considering the source of sound, the room, and the sound receiver or "sink," all as part of a unified dynamical system. This is required to bring out the interaction between source, sink, and room and their effects on the steady-state and transient aspects of sound transmission in the room.

In this theory the room is considered as an assemblage of resonators and the walls of the room as terminal impedances determining absorption and reflection. A rectangular room has a triple infinity of resonant frequencies. If the wall impedances are pure resistances, these frequencies are given by

$$f = 17{,}140\left[\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2\right]^{1/2} \qquad (11)$$

where $n_x, n_y, n_z = 0, 1, 2, \cdots$

l_x, l_y, and l_z = dimensions of the rectangular room in centimeters.

The distribution of these "allowed" frequencies (at which resonance occurs) may be graphically shown as in Fig. 34 by a three-dimensional plot in "frequency space" (after Morse). Each vector to a lattice point is associated with a "natural frequency" or "normal mode" of the room. The shortest vector, corresponding to the lowest frequency, is determined by the longest dimension of the room. The direction of the vector from the origin to a lattice point indicates the direction of excitation of that frequency in the room, and the length of the vector is proportional to its frequency

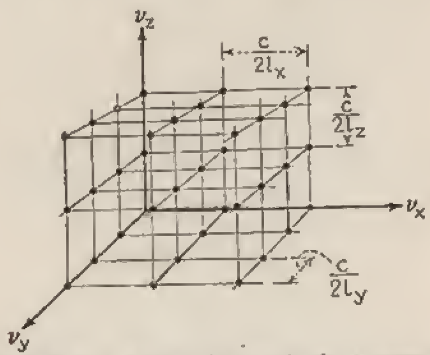


Fig. 34.—Distribution of resonant frequencies for a rectangular room with side lengths l_x, l_y, and l_z. The length of a vector from the origin to each lattice point indicates the frequency and the direction of the vector indicates the direction of the corresponding standing wave. The velocity of the sound wave c is given by Eq. (1).

At low frequencies there may be an appreciable frequency interval between the natural frequencies if the room is small. At high frequencies the number of natural frequencies in a given frequency interval is proportional to the square of the frequency.

28a. Reverberation. Using this concept of multiple natural frequencies, the decay of sound in a room may be described as follows: Assume energy has been supplied the room until the energy level is constant, *i.e.*, the rate of absorption at the boundaries equals the rate of supply to the room. The resulting standing wave system depends not only on the room and frequency but on the location and orientation of the source. When the source of energy is stopped, each individual mode of vibration of the room will decay exponentially, and the combined effect of these is called *reverberation*. Only the modes having allowed frequencies near the frequency of the steady-state excitation will contain appreciable energy.

By definition the reverberation time is the time required for the mean energy density in the room to drop 60 db. While this mean may be the result of a large number of rates of decay each of which is individually exponential, the combined value in general is not given by a single exponential term. This accounts for the fact that the slope of the mean-energy-density time-decay curves for the average room are not uniform and therefore for the fact that the apparent reverberation time depends on the time interval over which the decay is averaged. If the absorption is moderate the approximate reverberation time in seconds is given by

$$T = 0.00161\frac{V}{a} \tag{12}$$

where V = room volume in cubic centimeters
$a = (A_1\alpha_1 + A_2\alpha_2 + \ldots)$ total room absorption
and A_1, A_2, etc., are areas in square centimeters having absorption coefficients α_1, α_2, etc., respectively.

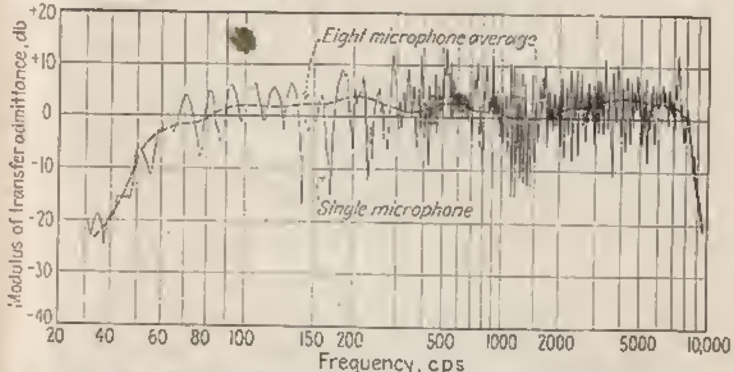


Fig. 35.—Transfer admittance or "response curve" of a speaker, room, and microphone or electro-acousto-electrical transducer. Speaker and microphone are in diagonally opposite corners of an 18- by 20- by 11-ft. room.

Corresponding to this type of energy decay in the room, there is a growth curve. When a source suddenly emits energy, each of the excited modes absorbs energy in an exponential manner. This occurs until the asymptotic or steady-state value is reached after an infinite length of time. When this is reached the acoustic power supplied the room must equal that absorbed at the room boundaries. By definition the absorption coefficient α of the boundary is the fraction of the incident energy absorbed for a specified angle of incidence. The intensity I of a sound wave [Eq. (2)] is a measure of the energy per square centimeter per second. The energy absorbed by the boundary per square centimeter per second is therefore $I\alpha$ watt. The total power absorbed by the room will be Ia watt where a is as defined in Eq. (12). This assumes, of course, that I is uniform throughout the room. This assumption is reasonably valid if the room is reverberant enough to be a good listening room, if the sound source is not highly directional, and if the room dimensions are many wave lengths long. Or mathematically

$$W_a = Ia \tag{13}$$

where I = sound intensity in watts per square centimeter and W_a = acoustic power radiated by the source. The acoustic power equals the product of the speaker efficiency and electrical signal input power. If the room absorption is known, the speaker efficiency in this room may be determined by measuring the average sound intensity in the room.

By combining Eqs. (12) and (13) we get

$$I = 620\frac{W_a T}{V} \text{ watt per sq cm} \tag{14}$$

28b. Room Power Requirements. If we know the desired sound intensity, the acoustic input power W_a required to produce it may be obtained either from Eq. (13) by knowing the total room absorption or

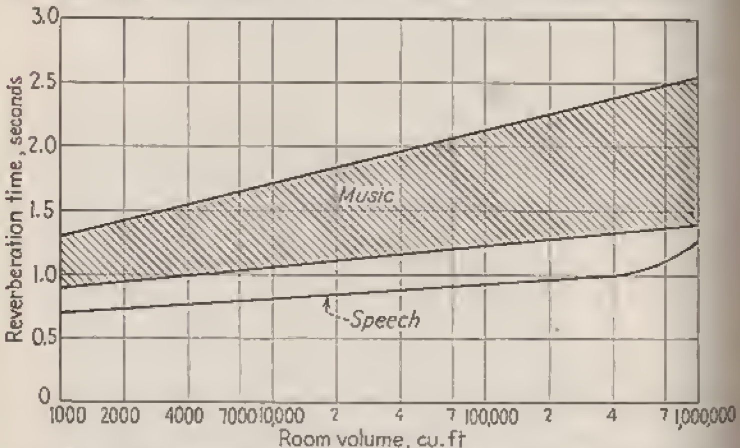


FIG. 36.—Dependence of optimum reverberation time for speech and range of reverberation time for music on room volume.

from Eq. (14) by knowing the room volume and reverberation time. Desirable reverberation times in terms of room volume for speech and music are shown in Fig. 36. The values for speech are seldom realized except in acoustically treated rooms. Typical schoolrooms with average attendance, for example, usually have reverberation times well up toward the upper music range.

Speech articulation increases rapidly with intensity up to an intensity level of 40 db or 10^{-12} watt per sq cm and more slowly to 50 db or 10^{-11} watt per sq cm. If room noises are present, the speech intensity should exceed these by at least 10 db. In conversational speech the person speaking radiates about 10^{-5} watt. Loud speaking requires 10^{-3} watt. If a loud-speaker is to simulate a person speaking loudly, its acoustic output should be at least 10^{-3} watt which, for a 1 per cent efficient loud-speaker radiating all its output into the room, means an electrical input of one-tenth watt.

There is considerable difference of opinion on what constitute acceptable levels of reproduced sound. Values of electrical power input which

have been suggested for theater use are shown in Fig. 37. These are based on the use of speaker systems which have average system efficiencies of 25 per cent. The trend is toward larger inputs to get enhanced dramatic value in the reproduction.

29. Acoustic System Characteristics. When a sound receiver is included in the room with a source, then we must consider the reaction of the room on it. The most common receiver is a listener. Because of the difficulty, however, of making objective measurements of what is going on in the listener's central nervous system, it is more convenient,

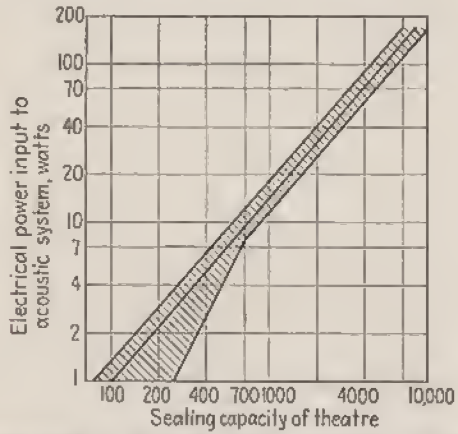


FIG. 37.—Recommended amplifier output for motion-picture reproduction. Speaker system efficiency assumed to be 25 per cent so acoustic input to theater is one quarter indicated electrical input. Trend is toward higher values.

although only approximately correct, to substitute one or more microphones for the listener.

For sake of simplicity, assume we have a loud-speaker as a source and a microphone as a receiver or sink, each with two accessible terminals. Since no source of energy is assumed in the room, these four leads may be considered the terminals of a passive quadripole or four-terminal network. From circuit theory we know that the measurement of three independent quantities will completely specify the performance of this quadripole at any one frequency. By analogy with the purely electrical case we may define the ratio (complex) of the current through a specified microphone load to the input voltage of the speaker as a transfer admittance.

A curve giving the magnitude of this quantity for a speaker and a microphone mounted in diagonally opposite corners of a rectangular 18- by 20- by 11-ft. room is shown in Fig. 35. This transfer admittance is what might be called the response curve of the loud-speaker measured in this room with designated locations for the source and microphone and with the particular microphone employed. As would be expected from the complicated equivalent circuit of the room, this transfer admittance varies by a large factor with frequency and exhibits a large number of maxima and minima at high frequencies. In passing it may be noted that the impedance of the boundary of this room varied appreciably with frequency and therefore the location of the resonant and antiresonant frequencies of the "electro-acousto-electrical network" do not occur at the frequencies predicted by constant boundary impedances.

Considering the matter in this light, we see that this response curve depends on the type of speaker, microphone, their location in the room, the geometry of the room, and the impedance of the room boundaries and therefore the impedance of the entire dynamical system seen from the acoustic terminals of the speaker and microphone.

The loud-speaker supplies maximum energy to the room when the impedance seen by the diaphragm looking into the room is the conjugate of the impedance looking back into the diaphragm with the speaker connected to its generator or amplifier. Since the latter impedance is usually high, the speaker supplies maximum power when the room impedance is high, which occurs (by definition) when the ratio of the pressure to particle velocity is high, *i.e.*, when the speaker is near a pressure maximum. A pressure-actuated microphone gives maximum response at a pressure maximum. The maximum peaks in the transfer admittance of Fig. 35 therefore occur when both the microphone and speaker are near pressure maxima. No resonance pressure maximum occurs at the speaker below the lowest resonant frequency of the room, and good l-f response is therefore hard to obtain in small rooms.

The apparent damping or Q of a mode of the system may be obtained by steady-state sharpness of resonance methods corresponding to those employed in circuit investigations. With negligible dissipation due to losses in the source, air, and sink, the damping of a particular mode is an indication of the absorption of energy at the boundary. From this damping coefficient the effective absorption coefficient of the boundary under the conditions determined by the mode may also be obtained. The simplest result is obtained when the wave front is parallel to the walls on which it is incident. The same apparent absorption is obtained from the transient decay of the same mode.

Likewise, by analogy with the electrical case, we may think of the transient current which flows through the microphone load when a unit d-c potential is suddenly supplied to the speaker terminals as the transfer indicial admittance of this electroacoustoelectrical network. Viewed in this light the transient response of the speaker itself (a small part of the dynamical system) or the transfer indicial admittance of the loud-speaker and microphone mounted in a free field where no reflections are present takes on much less significance.

We know experimentally that any room which is considered a good acoustical listening environment has appreciable reverberation and therefore that the rate of decay of the energy in the resonators of the room is only moderate. Experimentally it has been found that the rate of decay of the modes of the speaker itself measured in a free field is of the same order. If the loud-speaker is loosely coupled to the room, *i.e.*, if the room impedance seen by its diaphragm is small as compared with the impedance seen looking into the diaphragm, then we may loosely think of the loud-speaker as converting the unit d-c e.m.f. into a number of damped sinusoidal terms (one corresponding to each mode of the speaker), which in turn excite the room. The spectral composition of a single damped sinusoid for two rates of decay is shown in Fig. 7. From these we see that, if the rate of decay is large, the number of room modes excited may be large, because of the broad frequency spread of the energy exciting the room. Actually, of course, we should only think of the dynamical system as a whole and the above explanation as a simplification of the problem.

Experimental curves of this transfer indicial admittance are what one would predict from the theory. If the speaker is loosely coupled to the room, if its damping when it looks into a small acoustical impedance is low, and if an undamped resonant frequency of the speaker lies near one of the resonant modes of the room, the transient term looks like a typical one for two coupled circuits. That is, there are two prominent rates of decay containing the combined effect of the two important resonant frequencies (the speaker and room). On the other hand, if the speaker is highly damped when looking into a high acoustical impedance, if the driving point impedance of the room at an undamped natural frequency of the diaphragm is high, and if the room is large or its dimensions are so chosen that there are a number of resonant frequencies near an undamped resonant frequency of the speaker diaphragm, then the transient term consists of the superposition of a large number of damped sinusoids. In this case it may be seen that elimination of the term due to the loud-speaker would cause a negligibly small change in the apparent transient response of the system. This was verified in an unreported investigation conducted 8 years ago, in which it was shown experimentally that, if the fundamental speaker mode was eliminated by the use of a properly chosen electrical network, the aural result in reproduced speech and music was small unless the damping of the speaker radiating into a free field was unusually small.

The more important practical implications of the above (see also Art. 7, Mutual Radiation Impedance) are the following: (*a*) The loud-speaker should preferably be mounted in the corner of the room. In this position the greatest number of room resonances are "excited" and the most energy is supplied to the room. (*b*) The *average* l-f radiation is a maximum when the speaker is as near the floor (or ceiling) as possible and in the room corner. Next most desirable location is near floor (or ceiling) and side wall. (*c*) At any one frequency, maximum radiation is obtained when the room impedance seen by the diaphragm is the conjugate of the impedance seen looking back into the diaphragm. That is for the location which makes the *combined* speaker and room resonate. Such a maximum may not be obtained if the longest room dimension is less than roughly a half wave length long. (*d*) Because of this and the small number of resonant frequencies which occur in small rooms at low frequencies, small rooms do not normally permit the best l-f reproduction. (*e*) Corner positions also permit improved h-f response because of the smaller solid angle the radiation has to cover. (*f*) The l-f transient response of the speaker itself is not so important as is generally supposed because the transient response of the room helps obscure this distortion.

OBJECTIVE LOUD-SPEAKER TESTS

The following more important characteristics of a loud-speaker must be determined in any complete test: response-frequency, efficiency-frequency, directional, impedance, and distortion.

30. Response-frequency Characteristic (Steady State). A response-frequency curve of a speaker is a curve graphically depicting the sound produced at a designated position in the medium, the electrical input and acoustic environment being specified. Frequency discrimination is the most important form of distortion in many loud-speakers, and the response curve attempts to indicate quantitatively the amount present. Since the ear is primarily responsive to the sound pressure, the ordinate of the curve is made proportional to it or to its average value in a specified region.

The response curve is obtained by connecting the loud-speaker to a variable frequency source of specified internal impedance and constant specified internal voltage. The pressure at one or more points in the medium is measured as the frequency is varied slowly enough so the resulting measurement does not differ appreciably from the steady-state value.

A "free-field" response curve is made outdoors in the absence of unintended reflecting surfaces and is probably the most useful single curve showing the loud-speaker performance for outdoor applications. Curves of this type are valuable because (1) the direct incident sound from the source in various directions may be accurately determined and a close estimate made of the direct sound indoors; (2) the acoustic environment is relatively simple since only intended reflecting surfaces are included. The efficiency of the unit may then be accurately obtained for this environment at some frequencies and estimated at others, since the impedance seen by the diaphragm will change slowly with frequency (*i.e.*, the transfer admittance of the speaker, air, and microphone is a smooth curve). (3) The specified test conditions may be duplicated relatively easily at various laboratories permitting significant comparison of test results. The construction of identical test rooms, however desirable, would be difficult, partly because of differences of opinion on an "average" room and partly for economic reasons.

By "intended" reflecting surfaces is meant those that are an intended part of the radiating system. Frequently a cabinet or enclosure is measured outdoors in the absence of all reflecting surfaces, *i.e.*, radiating into a solid angle of 4π steradians or a complete sphere. This is usually undesirable since most enclosures are intended to operate on a floor and against a wall which adds two intended reflecting surfaces and primary images (see Radiation Impedance). If the enclosure is intended to operate in a room corner, there are three important primary images. The impedance seen looking into these reduced solid angles of π and $\pi/2$ steradians, respectively, is very different from the 4π case and usually results in an error of the order of 6 to 8 db at low frequencies. The error is especially large in the case of compound sources such as vented enclosures. The measured non-linear distortion usually differs by a much larger factor.

Outdoor measurements into solid angles of π and $\pi/2$ steradians are made by constructing large rigid non-absorbing surfaces.

Other intended parts of the speaker such as the baffle, horn, enclosure, etc., should, of course, be specified. The normal impedance or the impedance looking into the signal terminals of the speaker with the acoustical load (acoustical environment), used when the response curve was obtained, should be plotted. Both the angle and modulus of this impedance are required if the response of the speaker with any source impedance other than that employed in the test is to be calculated.

If a space average of the pressure is obtained by moving the microphone or by using multiple microphones, details of the method should be given. If a warble tone or noise generator is used to get a "moving frequency average" of the transfer admittance, the spectral composition of the source should be specified. These expedients and the one involving motion of the loud-speaker, which is usually unsatisfactory, are recommended only for indoor measurements when the room does not provide approximately free-field conditions.

Normal listening-room measurements are made with the loud-speaker mounted in its intended position in a typical listening room. As noted under Room Acoustics, the room impedance seen by the loud-speaker depends on the characteristics and location of the source itself (diaphragm sizes, locations, and modes of vibration), the geometry of the room, and the impedance of its boundaries. This means that the energy supplied the room depends on the particular room and speaker location chosen. This is frequently used as an argument against this type of test. Since the results obtained in reasonably similar rooms, with similar speaker locations in each, differ by only a moderate amount, this disadvantage does not outweigh the many important advantages of this type of test, some of which are (1) the impedance seen by the loud-speaker (including cabinet or enclosure), averaged over a small frequency interval, is closer to the average impedance seen under operating conditions than the impedance seen under the usual (4π steradians) outdoor test conditions; (2) calculation of the indoor from the outdoor performance is only of academic interest when hundreds of response curves are to be obtained, because of the labor involved; (3) ready comparison of the results of objective and subjective or listening tests in the same room is possible if the room is a good listening room; (4) one is not at the mercy of the weather; and most important (5) test facilities are readily provided in almost any organization.

Three large laboratories measured one speaker and plotted what they would publish as the response-frequency graph of the speaker. The results are shown in Fig. 38. This does not indicate any error in measurement. Actually different things were measured in each case. The curves indicate that response curves must be interpreted with great care and then only by a person familiar with the many factors involved. No speaker expert thinks of choosing a speaker solely or even largely on the basis of a response curve.

31. Efficiency-frequency Characteristic. If the free-field-pressure response at a sufficient number of points on a spherical surface centered on the diaphragm is obtained, the total acoustical output may be calculated.

The efficiency-frequency curve of a speaker corresponds to the response-frequency curve except that the ordinate indicates the efficiency (usually "absolute" or system). In a typical listening environment and listener location the direct incident sound energy, which would be approximately indicated by the free-field response-frequency curve at the listener's location with respect to the speaker, is only a small fraction of the reflected sound energy. A curve which gives the pressure, averaged over the useful listening region, then indicates the probable pressure the listener will experience. If absorption at the room boundary is independent of frequency this will be proportional to the total energy emitted by the speaker. Efficiency-frequency or space-averaged response-frequency curves are therefore the most useful in interpreting indoor operation. Outdoor response-frequency curves at various angles off the speaker axis, with the speaker radiating into approximately the solid angle it will see indoors, are also desirable since the listener, owing to his ability to localize sounds, weights the direct incident sound energy particularly at high frequencies more heavily than the same energy if in a reflected wave.

SUBJECTIVE LOUD-SPEAKER TESTS

32. Listening or Subjective Tests. Listening tests are a necessary part of the complete test of a loud-speaker. While objective measurements

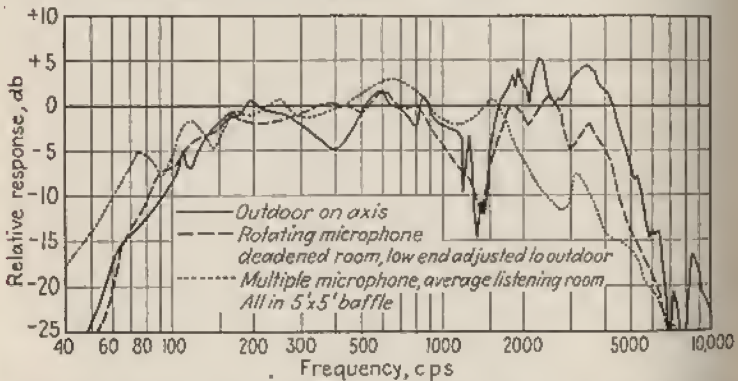


Fig. 38.—Response of one speaker as measured by three different companies, showing the futility of relying on response curves without a complete knowledge of the specific test, acoustic environment, and method.

are valuable in design work and in quantitatively determining some performance criteria, they cannot at present completely specify the subjective performance. Because of the apparent simplicity of listening tests many important factors are frequently neglected, with the result that many tests are meaningless and others actually misleading.

In both indoor and outdoor tests all precautions should be taken that are used in objective tests. The only essential difference is that the listener is substituted for the microphone. The properties of the ear and listener must therefore be considered in interpreting the results.

33. Relative-loudness Efficiency. The most common test is one to determine the relative-loudness efficiency of two speakers. An attenuator in the amplifier which does not alter its response is adjusted (usually with a relay which also switches the speakers) to attenuate the input to the louder speaker by the amount required to make the speakers equally loud. The required attenuation of the louder in decibels is their relative loudness efficiency in decibels. The relative loudness will depend primarily on the spectral composition of the test signal, the response-frequency characteristic of the speakers, and on the sound

intensity. Tests on the speaking and singing voice and various types of music are usually averaged. A valuable signal source for this and response-frequency tests is a "flat" noise source, or one in which the energy is uniformly distributed with frequency. This particular spectral composition ensures energy at frequencies at which significant differences in the speaker response may occur.

34. Response-frequency Characteristic. Apparent subjective response-frequency tests may be made with the same signal sources used in the loudness tests. Since the listener is not mobile, "space-averaging" methods employed with microphones cannot be used and "frequency-averaging" methods are employed. While noise sources are occasionally used in objective tests, they have unfortunately been neglected in subjec-

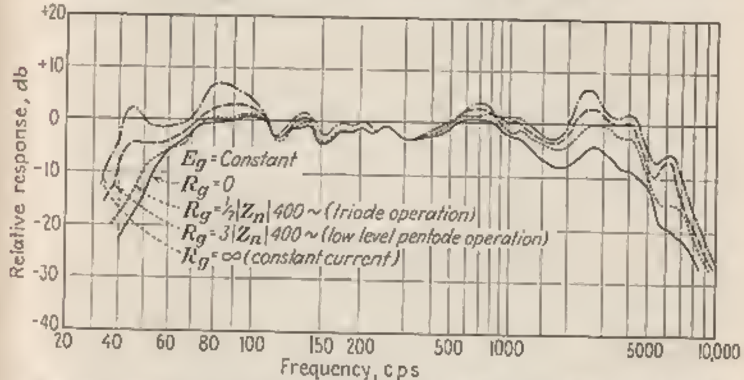


Fig. 39.—Variation in total sound power radiated by a bass reflex enclosure into a room as source impedance (impedance match) is varied. Generator or source voltage constant for each curve but arbitrarily raised as the source resistance was raised to maintain constant acoustic output at 400 cycles.

tive tests, where they are of special value because the trained ear can quickly appraise response differences which are missed if the signal source contains no energy at the frequencies at which differences occur.

35. Distortion Characteristic. Except with a single- or double-frequency input (the latter to determine intermodulation) it is difficult to determine the distortion characteristic of the speaker itself. With one or two simultaneously applied frequencies the input to the speaker is readily determined when the normal impedance of the speaker is known. This is not true of a signal of random energy distribution, and therefore with such a signal the apparent input to the speaker is not readily determined unless the normal impedance is relatively independent of frequency. When the *speaker* distortion characteristic is desired, the amplifier should be capable of supplying many times the rated input power to the speaker without distortion because of the high ratio of peak to average energy in speech and music (see these sections). Much overload charged to speakers is amplifier overload.

What is usually measured is the *system* distortion characteristic. Since amplifier overload almost invariably occurs at about the level at which speaker overload occurs in an economically planned system, what is measured is the combined system overload. In this case the speaker with the most restricted h-f response (other factors being equal) will have the best *system* overload rating since h-f distortion products are annoying.

There is no standard for speaker input power rating, but in practice a speaker rated at X watts will "handle" the output of an X-watt amplifier.

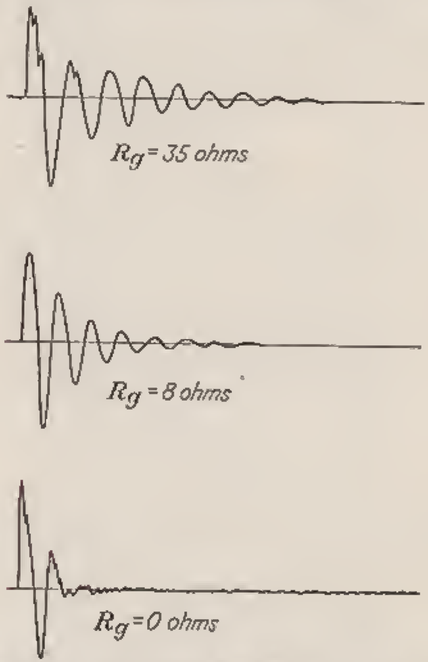
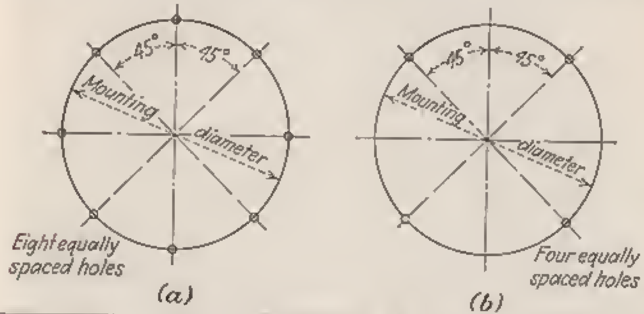


Fig. 40.—Response of moving-coil speaker with 8-ohm voice coil to an electrical pulse as source resistance is varied. Minor irregularities in curves are due to h-f modes of vibration of diaphragm.

which is not overloaded, with speech or music (complex wave) input (see Motors, p. 903; and Art. 18, Permanent and Electromagnets). The rating does not indicate the power the speaker will handle at a single frequency but takes advantage of the *normal* (no heavy bass or treble accentuation) spectral composition of speech and music. To avoid the trouble in determining the signal level across the speaker, with its variation in normal impedance, the grid voltage on the output stage may be measured. The signal input is raised until perceptible distortion results. The peak grid voltage is then measured with an indicator having a negligible time constant, such as a cathode-ray tube. A resistance

equal in value to the magnitude of the nominal loud-speaker impedance is substituted for the speaker. The power dissipated in the resistance load with a 400-cycle signal having the same maximum value as the maximum signal is the system (since it includes the output stage) speaker input power rating.

36. Outdoor Tests. These should be conducted so the listener subtends the same or a known angle with each source. The sources should be mounted so their mutual-radiation impedance does not influence the



Nominal Speaker Size, In.	Hole Arrangement	Mounting Diameter, In.	Minimum Hole Diameter, In.	Baffle* Hole Diameter, In.
3½	B	3 15/16	9/64	3⅛
4	B	4 11/16	1/8	3¾
5	A	4 15/16	1/8	4⅛
5¾	A	5⅜	1/8	4¾
6½	A	6⅛	1/8	5⅝
8	A	7⅝	7/32	6¾
10	B	9⅝	7/32	8¾
12	B	11 9/16	1/4	10¾
15	A	14 9/16	17/64	13¼

* Not an adopted standard.

Fig. 41.—RMA standard mounting dimensions for electrodynamic speakers.

result. Separating the sources by several times the diameter of the cone or horn mouth usually suffices. The energy absorbed by the unused speaker will be more nearly independent of frequency if the voice coil of the unused speaker is short-circuited. Unwanted reflecting surfaces should be avoided.

37. Indoor Tests. The speakers should be separated by several times the diameter of the cone or horn mouth to minimize mutual-radiation impedance. This is particularly true if the speakers are mounted on a common open baffle. Some coupling between the sources will always exist because of the transfer admittance between the two in the room (see Room Acoustics). It is important to mount the speakers symmetrically with respect to the room and listener in order to provide similar coupling between each source, the room, and the listener.

References

General:

American Standards Association, Tentative Standards, Z24.1, Z24.2, Z24.3 and Z24.4.
Bell System Tech. Jour. (see Cum. Index, vols. **1–10**).
CRANDALL, I. B.: "Theory of Vibrating Systems and Sound."
Electronics (McGraw-Hill Publishing Company, Inc.).
GEIGER, H., and KARL SCHEEL: "Handbuch der Physik" (vol. 8, Akustik), 1927.
I.R.E. Electroacoustic Standards, 1938 (also *Proc. I.R.E.*).
Jour. Acoustical Soc. Amer. (see Cum. Index, vols. 1–10).
MCLACHLAN, N. W.: "Loudspeakers."
MORSE, P. M.: "Vibration and Sound."
OLSON, H. F.: "Elements of Acoustical Engineering."
———, and F. MASSA: "Applied Acoustics." 1939 ed.
RAYLEIGH, LORD J. W. S.: "Theory of Sound."
Soc. Mot. Picture Engrs., Jour.
STEWART, G. W., and R. B. LINDSAY: "Acoustics."
WEIN-HARMS: "Handbuch der Experiment Physik" (vol. 17, parts 1–3).

Speech, Music, and Hearing:

BEASLEY, W. C.: Characteristics and Distribution of Impaired Hearing in the Population of the United States, *Jour. Acoustical Soc. Amer.*, **12**, 114–121, 1940; also National Health Survey, *Hearing Study Series Bulls.* 1–7.

FLETCHER, H.: "Speech and Hearing"; also Physical Characteristics of Speech and Music, *Bell System Tech. Jour.*, **10**, 349, 1931; STEINBERG, J. C.: Articulation Testing Methods, *Bell System Tech. Jour.*, **10**, 806, 1929; and MUNSEN, W. A.: *Jour. Acoustical Soc. Amer.*, **5**, 83, 1933.

SIVIAN, L. J.: Speech Power and Its Measurement, *Bell System Tech. Jour.*, **10**, 646, 1939; and DUNN, H. K., and S. D. WHITE: "Absolute Amplitudes and Spectra of Musical Instruments and Orchestras."

SNOW, W. B.: Audible Frequency Ranges of Music, Speech and Noise, *Jour. Acoustical Soc. Amer.*, **3**, 155–166, 1931.

STEVENS, S. S., and H. DAVIS: "Hearing, Its Psychology and Physiology."

Loud-speakers:

BOSTWICK, L. G.: *Bell System Tech. Jour.*, **8**, 135, 1929. (Tests.)
COOK, E. D.: *Gen. Elec. Rev.*, **33**, 505, 1930. (Tests.)
GREAVES, V. F., F. W. KRANZ, and W. D. KROZIER: The Kyle Condenser Loudspeaker, *Proc. I.R.E.*, **18**, 1142, 1929.
HALL, W. M.: *Jour. Acoustical Soc. Amer.*, **3**, 552, 1932. (Horns.)
KLAPMAN, S. J.: Interaction Impedance of a System of Circular Pistons, *Jour. Acoustical Soc. Amer.*, **11**, 289, 1940.
KNOWLES, H. S.: *Electronics*, **4**, 154, 1932; also **6**, 240, 1933. (Tests.)
OLNEY, B.: *Proc. I.R.E.*, **10**, 1113, 1931. (Tests.)
PEDERSEN, P. O.: *Jour. Acoustical Soc. Amer.*, **6**, 227, 1935; and **7**, 64, 1935. (Subharmonic theory.)
PHELPS, W. D.: *Jour. Acoustical Soc. Amer.*, **12**, 68, 1940. (Horn losses.)
RICE, C. W., and E. W. KELLOG: *Jour. A. I. E. E.*, **44**, 982; also 1015, 1925. (Moving-coil speaker.)
STENZEL, H.: *Elektrische Nach. Tech.*, **4**, 239, 1927; **6**, 165, 1929; and **7**, 90, 1930; also *Zeit. Tech. Physik*, **10**, 569, 1929; and *Ann. Physik* **11**, 947, 1930. (Directivity of radiators.)
WEGEL, R. L.: *Jour. A. I. E. E.*, **49**, 791, 1921. (Theory.)
WOLFF, I., and L. MALTER: *Phys. Rev.*, **33**, 1061, 1929. (Mutual radiation impedance.)
Also *Jour. Acoustical Soc. Amer.*, **2**, 201, 1930. (Directivity.)

Room Acoustics:

Jour. Acoustical Soc. Amer., Symposium, July, 1940, Bibliography (see also Cum. index, vols. 1–10).
KNUDSEN, V. O.: "Architectural Acoustics," 1932.
SABINE, P. E.: "Acoustics and Architecture," 1932.
WATSON, F. R.: "Acoustics of Buildings," 1930; and *Jour. Acoustical Soc. Amer.*, **3**, 14–43, 1931.

INDEX

A

D

S

T

RADIO
ENGINEERING
HANDBOOK

HENNEY

THIRD
EDITION

McGRAW-HILL
BOOK COMPANY