It can easily be verified that Eq. (28) is in agreement with Eq. (25) up to the first-order terms in ω_0. Thus, whenever the frequency of band width variation is reasonably small by comparison with the mean band width, Eq. (28) provides an adequate approximation to the impulsive response of the network. A higher order approximation may be obtained, if necessary, by continuing the iteration beyond F_2.

An Extension of Wiener's Theory of Prediction*

Lotfi A. Zadeh and John R. Ragazzini
Department of Electrical Engineering, Columbia University, New York, New York

The theory of prediction described in this paper is essentially an extension of Wiener's theory. It differs from the latter in the following respects.
1. The signal (message) component of the given time series is assumed to consist of two parts, (a) a non-random function of time which is representable as a polynomial of degree not greater than a specified number *n* and about which no information other than *n* is available; and (b) a stationary random function of time which is described statistically by a given correlation function. (In Wiener's theory, the signal may not contain a non-random part except when such a part is a known function of time.)
2. The impulsive response of the predictor or, in other words, the weighting function used in the process of prediction is required to vanish outside of a specified time interval $0 \leq t \leq T$. (In Wiener's theory T is assumed to be infinite.)
The theory developed in this paper is applicable to a broader and more practical class of problems than that covered in Wiener's theory. As in Wiener's theory, the determination of the optimum predictor reduces to the solution of an integral equation which, however, is a modified form of the Wiener-Hopf equation. A simple method of solution of the equation is developed. This method can also be applied with advantage to the solution of the particular case considered by Wiener. The use of the theory is illustrated by several examples of practical interest.

I. INTRODUCTION

PREDICTION—in the broad sense of the term—consists essentially of estimating the values of some function of time on the basis of a time series which may or may not contain random errors. For instance, a typical problem in prediction is as follows. Given a time series $e_1(t)$ which is composed of a signal $s(t)$ and a random disturbance (noise) $N(t)$; provide an estimate of $s(t+\alpha)$, α being a positive constant, as a continuous function of time. More generally, the quantity to be estimated may be a functional of $s(t)$ such as ds/dt, $\int s\,dt$, etc. In forming such estimates the mathematical operations that may be employed are usually limited by practical considerations. Thus, in most cases the operator furnishing the estimate must be linear and fixed in addition to the obvious requirement of being physically realizable. The physical counterpart of such an operator is what is commonly known as a predictor or an estimator.

It is evident that a function of time cannot be predicted intelligently unless sufficient *a priori* information is available about both the function and the errors. The nature of such information, as well as the characteristics of the signal and noise, can assume a variety of forms. Of these the more common ones have been investigated in recent years with the result that for certain classes of time series it is now possible to design predictors which make optimum use of the *a priori* information concerning the signal and the noise. Thus, when the given time series is stationary and the correlation functions of the signal and noise are known, one can use Wiener's theory[1] to arrive at the specifications of the optimum predictor, that is, one minimizing the mean-square value of the prediction error. On the other hand, when, as is often the case in practice, the given time series is non-stationary, the available theories of prediction, notably Phillips and Weiss' theory,[2] do not lead to the best possible predictor except for a narrow class of time series. It is possible, however, to extend Wiener's theory in many different directions thereby making it applicable to a wider class of problems than is covered by either Wiener's or Phillips and Weiss' theories in their present forms. One such extension is discussed in this paper. A feature of the extension is that the signal (message) is assumed to consist of a stationary component superimposed on a non-random function of time which is known to be representable as a polynomial of degree less than or equal to a specified number *n*. It will also be shown that the general method developed for treating this problem can be applied with advantage to the solution of many cases of practical

[1] N. Wiener, "The extrapolation, interpolation, and smoothing of stationary time series," Report of the Services 19, Research Project DIC-6037, M.I.T. (February, 1942). Published in book form by John Wiley and Sons, Inc., New York (1949).
[2] R. S. Phillips and P. R. Weiss, "Theoretical calculation on best smoothing of position data for gunnery prediction," Report 532, Rad. Lab., M.I.T. (February, 1944).

interest as well as the particular case considered by Wiener.

II. FORMULATION

Consider a given time series $e_1(t)$ which is the sum of a function $s(t)$ (signal) and a stationary random disturbance $N(t)$. Let $s^*(t)$ be the quantity to be estimated and let $s^*(t)$ be related to $s(t)$ through a given linear operator $K(p)$, i.e.,

$$s^*(t) = K(p)s(t). \tag{1}$$

$K(p)$ may be thought of as the system function of an *ideal predictor*, i.e., a predictor capable of perfect prediction in the absence of noise. In many cases, particularly those involving actual prediction, the operator $K(p)$ is not physically realizable so that the process of estimation cannot be carried out exactly even in the absence of random disturbances.

Frequently it will be convenient to use a different, though equivalent representation of Eq. (1), i.e.,

$$s^*(t) = \int_{-\infty}^{\infty} k(\tau)s(t-\tau)d\tau, \tag{2}$$

where τ is the variable of integration and $k(t)$ represents the impulsive response of the ideal predictor. $K(p)$ shall be referred to as the ideal prediction operator. As a matter of convenience, the more common of the many possible forms which $K(p)$ and $k(t)$ can assume are given in Table I.

Like all theories of prediction, the theory to be described applies only to a special class of time series. The time series to be considered in the work which follows will be assumed to consist of a signal $s(t)$ and noise $N(t)$, with the signal being composed of a random component $M(t)$ superposed upon a non-random function of time $P(t)$, i.e.,

$$s(t) = M(t) + P(t). \tag{3}$$

The assumptions made concerning the characteristics of $P(t)$, $M(t)$, and $N(t)$, are as follows:

(a) $P(t)$ is assumed to be representable as a polynomial in t of degree not higher than a specified number n.

(b) $M(t)$ and $N(t)$ are stationary functions of time described respectively by their auto-correlation functions $\psi_M(\tau)$ and $\psi_N(\tau)$.
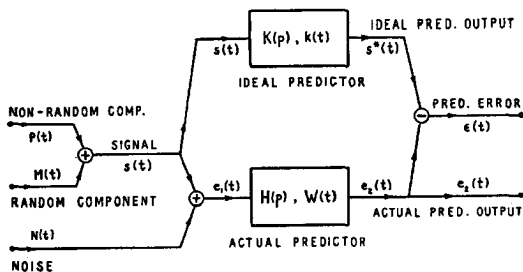

FIG. 1. Flow diagram of prediction process.

TABLE I. Common forms of the prediction operators $K(p)$ and $k(t)$.

Relation between $s^*(t)$ and $s(t)$	Significance of the quantity to be estimated	$K(p)$	$k(t)$
$s^*(t) = s(t)$	Present value of $s(t)$	1	$\delta(t)$
$s^*(t) = \dot{s}(t)$	Present value $\dot{s}(t)$	p	$\delta^{(1)}(t)$
$s^*(t) = \ddot{s}(t)$	Present value of $\ddot{s}(t)$	p^2	$\delta^{(2)}(t)$
$s^*(t) = s(t-\alpha)$	Past or future value of $s(t)$ depending respectively on whether α is a positive or negative constant	$e^{-\alpha p}$	$\delta(t-\alpha)$

Note.—$\delta(t)$ denotes a unit impulse at $t=0$, and $\delta^{(\nu)}(t)$ stands for the νth derivative of $\delta(t)$ with respect to t (time).

(c) $M(t)$ and $N(t)$ have zero mean and are uncorrelated. This assumption is introduced only for the purpose of simplification and is not essential to the analysis. The condition expressed by (c) prevails in most practical cases.

Referring to Fig. 1, these inputs are shown being applied to the actual predictor whose system function is $H(p)$ and whose impulsive response is $W(t)$. The output of the predictor, $e_2(t)$, may be expressed in operational form

$$e_2(t) = H(p)e_1(t) \tag{4}$$

or, alternatively, in the form of a superposition integral

$$e_2(t) = \int_0^{\infty} W(\tau)e_1(t-\tau)d\tau. \tag{5}$$

An important characteristic of the actual predictor is the so-called prediction or estimation error ϵ, which is defined as the difference between the output of the predictor and the quantity to be estimated, $s^*(t)$. Equation-wise this is:

$$\epsilon = e_2(t) - s^*(t). \tag{6}$$

If there were no noise and if $K(p)$ were physically realizable there would be no prediction error and $H(p)$ would be identical with $K(p)$. This, of course, is the trivial case of the prediction problem. In what follows it will be assumed that either because of the presence of noise or physical unrealizability of $K(p)$, or both, $H(p)$ cannot be the same as $K(p)$.

The available *a priori* information about $s(t)$ and $N(t)$ is assumed to consist of n, $\psi_M(\tau)$, and $\psi_N(\tau)$. The problem is to specify the system function or the impulsive response of a predictor that would minimize in a certain sense the prediction error $\epsilon = e_2(t) - s^*(t)$. By analogy with Wiener's theory it will be postulated that the optimum predictor is the one in which: (a) the ensemble mean of ϵ is equal to zero (for all values of t), and (b) the ensemble variance of ϵ is a minimum. Denoting the ensemble average by the symbol $\langle \ \rangle_{\text{Av}}$, these conditions read:

(a) $\langle \epsilon \rangle_{\text{Av}} = 0$ or, equivalently, $\langle e_2(t) \rangle_{\text{Av}} \equiv \langle s^*(t) \rangle_{\text{Av}}$, (7)

(b) $\sigma^2 = \langle \epsilon^2 \rangle_{\text{Av}} = \text{minimum},$ (8)

646

where σ^2, the ensemble variance of ϵ, is equal to the mean-square value of the prediction error. In what follows, conditions (a) and (b) will be used as the basis for the determination of the optimum predictor.

III. DETERMINATION OF THE IMPULSIVE RESPONSE OF THE OPTIMUM PREDICTOR

It will be recalled that the output of a predictor may be expressed in the form of a superposition integral

$$e_2(t) = \int_0^\infty W(\tau)e_1(t-\tau)d\tau \qquad (9)$$

where τ is a dummy variable and $W(t)$ represents the impulse response of the predictor. In practice it is usually found necessary to restrict the duration of sampling of the input time series to a finite constant T, meaning in other words that $W(t)$ must be zero outside the interval $0 \leq t \leq T$. To place this property in evidence Eq. (9) will be written in the following form:

$$e_2(t) = \int_0^T W(\tau)e_1(t-\tau)d\tau. \qquad (10)$$

In the limiting case where the duration of sampling is infinite $(T \to \infty)$ Eq. (10) becomes identical with Eq. (9).

By hypothesis,

$$e_1(t) = P(t) + M(t) + N(t). \qquad (11)$$

Substituting Eq. (11) into Eq. (10) and making use of the identity

$$P(t-\tau) \equiv P(t) - \tau \dot{P}(t) + \frac{\tau^2}{2!}\ddot{P}(t) + \cdots$$
$$+ (-1)^n \frac{\tau^n}{n!}P^{(n)}(t), \qquad (12)$$

it is found that $e_2(t)$ may be expressed as:

$$e_2(t) = \mu_0 P(t) - \mu_1 \dot{P}(t) + \frac{\mu_2}{2!}\ddot{P}(t) + \cdots + (-1)^n \frac{\mu_n}{n!}P^{(n)}(t)$$
$$+ \int_0^T W(\tau)M(t-\tau)d\tau + \int_0^T W(\tau)N(t-\tau)d\tau, \qquad (13)$$

where μ_0, μ_1, μ_2, etc., designate the moments of $W(t)$, i.e.,

$$\mu_\nu = \int_0^T \tau^\nu W(\tau)d\tau, \quad \nu = 0, 1, 2, \cdots n. \qquad (14)$$

Since $M(t)$ and $N(t)$ are stationary (with zero mean), it follows that the ensemble means of $e_2(t)$ and $s^*(t)$ depend only on the non-random component of the signal, i.e.,

$$\langle e_2(t) \rangle_{Av} = \int_0^T W(\tau)P(t-\tau)d\tau \qquad (15)$$

or

$$\langle e_2(t) \rangle_{Av} = \mu_0 P(t) - \mu_1 \dot{P}(t) + \frac{\mu_2}{2!}\ddot{P}(t) + \cdots$$
$$+ (-1)^n \frac{\mu_n}{n!}P^{(n)}(t), \qquad (16)$$

and

$$\langle s^*(t) \rangle_{Av} = \langle K(p)s(t) \rangle_{Av} \qquad (17)$$

or

$$\langle s^*(t) \rangle_{Av} = K(p)P(t). \qquad (18)$$

Comparing Eqs. (16) and (18), condition (a) is reduced to

$$K(p)P(t) \equiv \mu_0 P(t) - \mu_1 \dot{P}(t) + \frac{\mu_2}{2!}\ddot{P}(t) + \cdots$$
$$+ (-1)^n \frac{\mu_n}{n!}P^{(n)}(t). \qquad (19)$$

Equation (19), being an identity, determines the values of μ_0, μ_1, $\cdots$, and μ_n. In other words, *the ideal prediction operator $K(p)$ determines through Eq. (19) the first $n+1$ moments of the impulsive response of the optimum predictor.*

As an illustration of the foregoing statement consider a case where the quantity to be estimated is the derivative of $s(t)$, i.e., $s^*(t) = \dot{s}(t)$. For this case Eq. (19) reduces to

$$\dot{P}(t) \equiv \mu_0 P(t) - \mu_1 \dot{P}(t) + \frac{\mu_2}{2!}\ddot{P}(t) + \cdots$$
$$+ (-1)^n \frac{\mu_n}{n!}P^{(n)}(t), \qquad (20)$$

and a term by term comparison of the left-hand and right-hand sides of Eq. (20) yields:

$$\mu_0 = \int_0^T W(\tau)d\tau = 0$$

$$\mu_1 = \int_0^T \tau W(\tau)d\tau = -1$$

$$\mu_2 = \int_0^T \tau^2 W(\tau)d\tau = 0 \qquad (21)$$

$$\cdots$$

$$\mu_n = \int_0^T \tau^n W(\tau)d\tau = 0.$$

These, therefore, are the $n+1$ constraints which the impulsive response of a derivative estimating network must satisfy.

As the second example consider a case where $K(p)s(t) = s(t-\alpha)$, α being a positive or negative constant. For

this case Eq. (19) reads

$$P(t-\alpha) \equiv \mu_0 P(t) - \mu_1 \dot{P}(t) + \frac{\mu_2}{2!} \ddot{P}(t) + \cdots$$

$$+ (-1)^n \frac{\mu_n}{n!} P^{(n)}(t). \quad (22)$$

Rewriting $P(t-\alpha)$ as

$$P(t-\alpha) \equiv P(t) - \alpha \dot{P}(t) + \frac{\alpha^2}{2!} \ddot{P}(t) + \cdots$$

$$+ (-1)^n \frac{\alpha^n}{n!} P^{(n)}(t), \quad (23)$$

and making in Eq. (22) a term-by-term comparison of the coefficients of $P(t)$, $\dot{P}(t)$, etc., it is easily found that:

$$\mu_0 = \int_0^T W(\tau) d\tau = 1$$

$$\mu_1 = \int_0^T \tau W(\tau) d\tau = \alpha \qquad (24)$$

$$\cdots$$

$$\mu_n = \int_0^T \tau^n W(\tau) d\tau = \alpha^n,$$

which thus represent the constraints imposed upon $W(t)$ in case the quantity to be estimated is $s(t-\alpha)$.

The problem that remains to be solved is that of minimizing σ^2. For this purpose it will be necessary to develop an explicit expression for σ^2 in terms of $W(t)$ and the auto-correlation functions of the signal and noise. Assuming that condition (a) is satisfied, it follows from inspection of Eqs. (6), (13), and (19) that the prediction error is given by the expression

$$\epsilon = \int_0^T W(\tau)[M(t-\tau) + N(t-\tau)] d\tau - K(p) M(t) \quad (25)$$

or equivalently

$$\epsilon = \int_0^T W(\tau)[M(t-\tau) + N(t-\tau)] d\tau$$

$$- \int_{-\infty}^{\infty} k(\tau) M(t-\tau) d\tau \quad (26)$$

where $k(t)$ is the impulsive response of the ideal predictor. The mean-square value of ϵ may be written as

$$\sigma^2 = \langle \epsilon^2 \rangle_{\text{Av}} = \lim_{L \to \infty} \frac{1}{L} \int_0^L \epsilon^2 dt \quad (27)$$

or

$$\sigma^2 = \lim_{L \to \infty} \frac{1}{L} \int_0^L dt \left\{ \int_0^T W(\tau)[M(t-\tau) + N(t-\tau)] d\tau \right.$$

$$\left. - \int_{\infty}^{\infty} k(\tau) M(t-\tau) d\tau \right\}^2. \quad (28)$$

A typical term of Eq. (28) such as

$$\lim_{L \to \infty} \frac{1}{L} \int_0^L dt \left[\int_0^T W(\tau) M(t-\tau) d\tau \right]^2 \quad (29)$$

is expressible in the form of a triple integral

$$\int_0^T \int_0^T d\tau_1 d\tau_2 W(\tau_1) W(\tau_2)$$

$$\times \lim_{L \to \infty} \frac{1}{L} \int_0^L M(t-\tau_1) M(t-\tau_2) dt \quad (30)$$

which in view of the definition of the auto-correlation function of $M(t)$, i.e.,

$$\psi_M(\tau) = \lim_{L \to \infty} \frac{1}{L} \int_0^L M(t) M(t-\tau) dt \quad (31)$$

may be written as

$$\int_0^T \int_0^T W(\tau_1) W(\tau_2) \psi_M(\tau_1 - \tau_2) d\tau_1 d\tau_2. \quad (32)$$

Proceeding similarly in the case of other terms, Eq. (28) reduces finally to the following expression:

$$\sigma^2 = \int_0^T \int_0^T W(\tau_1) W(\tau_2) [\psi_M(\tau_1 - \tau_2)$$

$$+ \psi_N(\tau_1 - \tau_2)] d\tau_1 d\tau_2$$

$$- 2 \int_{\infty}^{\infty} \int_0^T W(\tau_1) k(\tau_2) \psi_M(\tau_1 - \tau_2) d\tau_1 d\tau_2$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(\tau_1) k(\tau_2) \psi_M(\tau_1 - \tau_2) d\tau_1 d\tau_2 \quad (33)$$

where, to recapitulate: τ_1, τ_2 = dummy variables; $W(t)$ = impulsive response of the predictor; $\psi_M(\tau)$ = auto-correlation function of $M(t)$ [$M(t)$ is the stationary part of the input signal]; $\psi_N(\tau)$ = auto-correlation function of $N(t)$ [$N(t)$ is the input noise]; $k(t)$ = impulsive response of the ideal predictor.

Returning to the problem of minimization of σ^2 it will be noted first that the last term in Eq. (33) is independent of $W(t)$ and hence, insofar as minimization of σ^2 is concerned, need not be considered. Second, it will be recalled that $W(t)$ is subject to the $n+1$ constraints

648

expressed by Eq. (14); therefore, the problem of minimizing σ^2 with respect to the class of $W(t)$'s satisfying Eq. (14) reduces essentially to an isoperimetric problem in the calculus of variations. Following the standard approach to such problems, one is led to minimizing the following expression:

$$I = \sigma^2 - 2\lambda_0\mu_0 - 2\lambda_1\mu_1 - \cdots - 2\lambda_n\mu_n \qquad (34)$$

or, more explicitly

$$I = \int_0^T W(\tau_1)d\tau_1 \left\{ \int_0^T W(\tau_2)[\psi_M(\tau_1-\tau_2) \right.$$

$$+ \psi_N(\tau_1-\tau_2)]d\tau_2 - 2\int_{-\infty}^{\infty} k(\tau_2)\psi_M(\tau_1-\tau_2)d\tau_2$$

$$\left. - 2\lambda_0 - 2\lambda_1\tau_1 - \cdots - 2\lambda_n\tau_1^n \right\} \qquad (35)$$

where the constants λ_0, λ_1, $\cdots$, λ_n, are the Lagrangian multipliers. Proceeding in the usual manner, that is, setting the variation of I equal to zero, it is easily found that I and hence σ^2 is a minimum provided $W(t)$ satisfies the following integral equation:

$$\int_0^T W(\tau)[\psi_M(t-\tau)+\psi_N(t-\tau)]d\tau = \lambda_0 + \lambda_1 t + \cdots$$

$$+ \lambda_n t^n + \int_{-\infty}^{\infty} k(\tau)\psi_M(t-\tau)d\tau, \quad 0 \leq t \leq T. \quad (36)$$

This equation together with the $n+1$ constraints expressed by Eq. (14) provides the basis for the determination of the optimum predictor. It will be observed that in the particular case where $n=0$, $T=\infty$, and $k(t)=\delta(t+\alpha)$ [$\delta(t)$ standing, as usual, for a unit impulse at $t=0$], Eq. (36) reduces to

$$\int_0^{\infty} W(\tau)[\psi_M(t-\tau)+\psi_N(t-\tau)]d\tau$$

$$= \psi_M(t+\alpha), \quad t>0 \quad (37)$$

which is essentially the integral equation of Wiener's theory. On the other hand, in the special case where $M(t) \equiv 0$, Eq. (36) reduces to

$$\int_0^T W(\tau)\psi_N(t-\tau)d\tau = \lambda_0 + \lambda_1 t + \cdots + \lambda_n t^n,$$

$$0 \leq t \leq T \quad (38)$$

which is the integral equation of Phillips and Weiss' theory. Thus, the integral equations of Wiener's, and Phillips and Weiss' theories are special cases of Eq. (36).

IV. SOLUTION OF THE INTEGRAL EQUATION

In the general case where $\psi_M(\tau)$ and $\psi_N(\tau)$ are prescribed but otherwise arbitrary auto-correlation functions, the complicated nature of the integral equation makes it appear that the solution of Eq. (36) is a formidable problem. In reality, the problem is not as difficult as it may seem, for by using a procedure to be described, the general case can be reduced to a special case which has a simple solution.

Preliminary to the discussion of this procedure it will be expedient to introduce the spectral densities of $M(t)$, $N(t)$, and $M(t)+N(t)$. Denoting these by $S_M(\omega^2)$, $S_N(\omega^2)$, and $S(\omega^2)$, respectively, and recalling that the spectral density of a function is the inverse Fourier transform of its auto-correlation function,[1] it follows that

$$S_M(\omega^2) = \int_{-\infty}^{\infty} \psi_M(\tau)e^{-i\omega\tau}d\tau \qquad (39)$$

$$S_N(\omega^2) = \int_{-\infty}^{\infty} \psi_N(\tau)e^{-i\omega\tau}d\tau \qquad (40)$$

and

$$S(\omega^2) = S_M(\omega^2) + S_N(\omega^2). \qquad (41)$$

Now the spectral density function $S(\omega^2)$ may be factored[1] into the product of two conjugate factors

$$S(\omega^2) = G(j\omega) \cdot G(-j\omega) \qquad (42)$$

such that both $G(j\omega)$ and $1/G(j\omega)$ are analytic in the right half of the $j\omega$-plane. Usually $S(\omega^2)$ is assumed to be a rational function of ω^2 of the form

$$S(\omega^2) = [A(\omega^2)/B(\omega^2)], \qquad (43)$$

where $A(\omega^2)$ and $B(\omega^2)$ are polynomials in ω^2. For such cases the process of factorization is quite straightforward as can be seen from the following examples:

(a) $\quad S(\omega^2) = \omega^2; \qquad G(j\omega) = j\omega.$

(b) $\quad S(\omega^2) = \dfrac{1}{\omega^2 + \omega_0^2}; \quad G(j\omega) = \dfrac{1}{j\omega + \omega_0}.$

(c) $\quad S(\omega^2) = \dfrac{\omega^2 + a^2}{\omega^4 + b^2\omega^2 + c^4};$

$$G(j\omega) = \frac{j\omega + a}{(j\omega)^2 + j\omega(b^2 + 2c^2)^{\frac{1}{2}} + c^2}.$$

To summarize, a rational spectral density function may be written as

$$S(\omega^2) = |G(p)|^2_{p=j\omega} \qquad (44)$$

where $G(p)$ is of the form:

$$G(p) = \frac{Q(p)}{R(p)} = \frac{a_0 + a_1 p + \cdots + a_m p^m}{b_0 + b_1 p + \cdots + b_l p^l}, \qquad (45)$$

and the polynomials $Q(p)$ and $R(p)$ do not have any zeros in the right half of the p-plane.
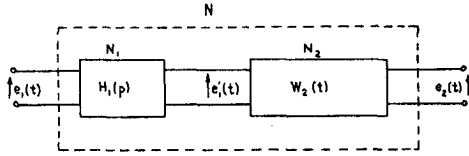
FIG. 2. Division of the predictor into the component networks N_1 and N_2.

The outline of the procedure used[3] for the solution of Eq. (36) can be best explained with reference to Fig. 2. The predictor N is assumed to be composed of two networks N_1 and N_2. The function of N_1 is to suitably modify some of the characteristics of the input time series $e_1(t)$, while that of N_2 is to provide the desired prediction through operating on the time series $e_1'(t)$, which is the output of N_1. It will be seen later that it is possible to choose N_1 in such a manner that the determination of the impulsive response of N_2 becomes an easily solvable problem. Then, once $W_2(t)$ (the impulsive response of N_2) is determined, the impulsive response of N, $W(t)$, can easily be found from the relation

$$W(t) = H_1(p)W_2(t), \quad (46)$$

where $H_1(p)$ is the system function of N_1. The choice of $H_1(p)$ and the problem of determination of $W_2(t)$ are discussed in the sequel.

It is evident that the problem of determination of $W_2(t)$ is similar to that of the determination of $W(t)$, except that the characteristics of the input time series are different for the two problems. An inspection of the integral equation (36) shows that it can be solved rather easily when the input to the predictor consists of a polynomial in t and a stationary component whose spectral density is a polynomial in ω^2. Therefore, in order to make the determination of $W_2(t)$ a simple problem, it is necessary to provide N_2 with an input which has this property. It is not difficult to verify that such a condition will obtain if, and only if, the system function of N_1 is chosen to be

$$H_1(p) = R(p), \quad (47)$$

where $R(p)$ is the denominator of $G(p)$ [cf. Eq. (45)]. With this choice of $H_1(p)$ the input to N_2 will consist of a polynomial in t of the same degree[4] as $P(t)$, and a stationary component $M'(t) + N'(t)$ whose spectral density is

$$S'(\omega^2) = |H_1(j\omega)|^2 S(\omega^2) \quad (48)$$

or, in view of Eqs. (43), (44), and (47),

$$S'(\omega^2) = A(\omega^2), \quad (49)$$

[3] The appendix of a report by Bode, Blackman, and Shannon, "Data smoothing and prediction in fire-control systems," Research and Development Board, Washington, D. C. (August, 1948), contains a brief exposition of a method which is similar in certain respects to the method described here.

[4] It is tacitly assumed that $R(p)$ does not have a zero at the origin or, in other words, that $S(\omega^2)$ does not have a pole at zero frequency.

where $A(\omega^2)$ is the numerator of $S(\omega^2)$. It will be noted that $A(\omega^2)$ is a polynomial of the form

$$A(\omega^2) = \gamma_0 + \gamma_1\omega^2 + \cdots + \gamma_m\omega^{2m}, \quad (50)$$

and correspondingly the auto-correlation function of $M'(t) + N'(t)$ is

$$\psi_M'(\tau) + \psi_N'(\tau) = \gamma_0\delta(\tau) - \gamma_1\delta^{(2)}(\tau) + \cdots + (-1)^m\gamma_m\delta^{(2m)}(\tau), \quad (51)$$

where $\delta^{(\nu)}(\tau)$ represents the impulse function of νth order [i.e., the νth derivative of the unit impulse function $\delta(\tau)$].

In addition to $\psi_M'(\tau) + \psi_N'(\tau)$, a number of other quantities associated with the input to N_2 enter the integral equation satisfied by $W_2(t)$. The significance of each of these quantities, as well as their expressions, are as follows:

(a) $S_M'(\omega^2) =$ spectral density of $M'(t)$
$$= S_M(\omega^2)|R(j\omega)|^2. \quad (52)$$

(b) $\psi_M'(\tau) =$ auto-correlation function of $M'(t)$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_M(\omega^2)|R(j\omega)|^2 e^{j\omega\tau}d\omega. \quad (53)$$

(c) $k'(t) =$ ideal impulsive response for N_2
$$= [1/R(p)]k(t). \quad (54)$$

In terms of these quantities the integral equation satisfied by $W_2(t)$ reads:

$$\int_0^\infty W_2(\tau)[\psi_M'(t-\tau) + \psi_N'(t-\tau)]d\tau$$

$$= \lambda_0' + \lambda_1't + \cdots + \lambda_n't^n$$

$$+ \int_{-\infty}^\infty k'(\tau)\psi_M'(t-\tau)d\tau, \quad t \geq 0. \quad (55)$$

It will be noticed that in the case of $W_2(t)$ the upper limit of the integral is infinity, while in the case of $W(t)$ [cf. Eq. (36)] it is T. The explanation for this difference is that $W_2(t)$ need not vanish for $t > T$, even though $W(t)$ is required to do so. Thus in general, $W_2(t)$ will be piecewise analytic in the interval $0 < t < \infty$ as is illustrated in Fig. 3. Denoting the parts of $W_2(t)$ extending over the intervals $0 \leq t \leq T$ and $T < t < \infty$ by $U(t)$ and $V(t)$, respectively, the relation connecting $W(t)$ and $W_2(t)$ [cf. Eq. (46)] may be rewritten in the following form:

$$W(t) = R(p)U(t) \quad (56a)$$

and

$$0 = R(p)V(t). \quad (56b)$$

These relations show that $W(t)$ is completely determined by the part of $W_2(t)$ which extends over the interval $0 \leq t \leq T$; the form of $W_2(t)$ outside this interval is irrelevant to the determination of $W(t)$.

Returning to the integral equation (55), it will be noted that the range of integration $0 \leq \tau < \infty$ may be divided into two parts, $0 \leq \tau \leq T$ and $T < \tau < \infty$, involving $U(t)$ and $V(t)$, respectively. Since $V(t)$ is determined by Eq. (56b) to within a finite number of constants, the integral equation in question degenerates into an integral equation involving only $U(t)$:

$$\int_0^T U(\tau)[\psi_M'(t-\tau) + \psi_N'(t-\tau)]d\tau = \lambda_0' + \lambda_1' t + \cdots$$

$$+ \lambda_n' t^n + \int_{-\infty}^{\infty} k'(\tau)\psi_M'(t-\tau)d\tau, \quad 0 \leq t \leq T. \quad (57)$$

Upon substitution of Eqs. (52), (53), and (54), and performing minor simplifications, Eq. (57) reads

$$\int_0^T U(\tau)[\gamma_0 \delta(t-\tau) - \gamma_1 \delta^{(2)}(t-\tau) + \cdots$$

$$+ (-1)^m \gamma_m \delta^{(2m)}(t-\tau)]d\tau = \lambda_0' + \lambda_1' t + \cdots + \lambda_n' t^n$$

$$+ \frac{1}{2\pi} \int_{-\infty}^{\infty} S_M(\omega^2) K(j\omega) R(-j\omega) e^{j\omega t} d\omega. \quad (58)$$

Making use of the identity

$$\int_0^T U(\tau) \delta^{(2\nu)}(t-\tau)d\tau \equiv p^{2\nu} U(t), \quad (59)$$

Equation (58) may be rewritten as

$$[\gamma_0 - \gamma_1 p^2 + \cdots + (-1)^m \gamma_m p^{2m}]U(t)$$

$$= \lambda_0' + \lambda_1' t + \cdots + \lambda_n' t^n$$

$$+ \frac{1}{2\pi} \int_{-\infty}^{\infty} S_M(\omega^2) K(j\omega) R(-j\omega) e^{j\omega t} d\omega. \quad (60)$$

Since in this equation the left-hand side operator is simply $A(-p^2)$ [cf. Eq. (50)], the integral equation (57) finally reduces to the following differential equation:

$$A(-p^2)U(t) = \lambda_0' + \lambda_1' t + \cdots + \lambda_n' t^n$$

$$+ \frac{1}{2\pi} \int_{-\infty}^{\infty} S_M(\omega^2) K(j\omega) R(-j\omega) e^{j\omega t} d\omega. \quad (61)$$

The general solution of this equation is of the form:

$$U(t) = A_0' + A_1' t + \cdots + A_n' t^n + B_1' \exp(\alpha_1 t)$$

$$+ B_2' \exp(\alpha_2 t) + \cdots + B_{2m}' \exp(\alpha_{2m} t)$$

$$+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{A(\omega^2)} K(j\omega) R(-j\omega) e^{j\omega t} d\omega,$$

$$0 \leq t \leq T, \quad (62)$$

where A_0', A_1', $\cdots$, A_n' and B_1', B_2', $\cdots$, B_{2m}' are as

yet undetermined constants, and α_1, α_2, $\cdots$, α_{2m} are the roots of the characteristic equation

$$A(-p^2) = 0. \quad (63)$$

In brief, Eq. (62) provides an explicit expression for $U(t)$ involving $2m+n+1$ undetermined constants. Availability of such an expression reduces the problem of determination of $W(t)$ to a relatively routine matter which is discussed in the following section.

V. DERIVATION OF AN EXPLICIT EXPRESSION FOR $W(t)$

Recalling that $W(t)$ is related to $U(t)$ through the operational relation

$$W(t) = R(p)U(t), \quad (56a)$$

and substituting $U(t)$ as given by Eq. (62) into Eq. (56a), it is readily found that in the most general case $W(t)$ is given by the following expression:

$$W(t) = [u(t) - u(t-T)]\Big\{ A_0 + A_1 t + \cdots + A_n t^n$$

$$+ B_1 \exp(\alpha_1 t) + \cdots + B_{2m} \exp(\alpha_{2m} t)$$

$$+ \frac{1}{2\pi} R(p) \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{A(\omega^2)} K(j\omega) R(-j\omega) e^{j\omega t} d\omega \Big\}$$

$$+ C_1 \delta(t) + \cdots + C_{l-m} \delta^{(l-m-1)}(t)$$

$$+ D_1 \delta(t-T) + \cdots + D_{l-m} \delta^{(l-m-1)}(t-T) \quad (64)$$

where the A's, B's, C's, and D's are as yet undetermined constants, and the unit step functions $u(t)$ and $u(t-T)$ are used simply to indicate that $W(t)$ is zero outside the interval $0 \leq t \leq T$. The impulse functions contained in the expression for $W(t)$ arise from operation by $R(p)$ on the discontinuities of $U(t)$ at $t=0$ and $t=T$. It will be observed that the order of these impulse functions does not exceed $l-m-1$, which is one-half the order of the zero of $S(\omega^2)$ at infinity minus one. This is due to the fact that the first $m-1$ derivatives of $U(t)$ vanish at $t=0$ and $t=T$. It is not difficult to verify that if this would not have been the case, the mean-square error at the output of N_2 would be infinite.

Having obtained the general expression for $W(t)$ in the form of Eq. (64), there remains the problem of determination of the $2l+n+1$ unknown constants. These can be found in the following manner:
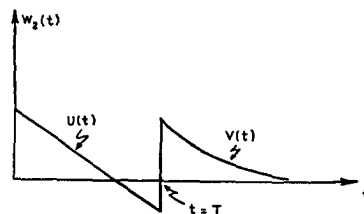

FIG. 3. Form of the impulsive response of N_2.

1. Substituting $W(t)$ as given by Eq. (64) into the integral equation (36) and requiring that the equation be satisfied identically, leads to $2l$ linear homogeneous equations in the A's, B's, C's, and D's.

2. Substituting $W(t)$ as given by Eq. (64) into the $n+1$ moment equations

$$\int_0^T \tau^\nu W(\tau)d\tau = \mu_\nu, \quad \nu = 0, 1, 2, \cdots, n \qquad (65)$$

yields other $n+1$ linear equations. These $n+1$ equations, together with the $2l$ equations obtained in (1), provide a system of $2l+n+1$ linear equations in the unknown constants. Solution of this system gives the values of the A's, B's, C's, and D's and thus completes the process of determination of $W(t)$.

It should be remarked that in some cases it is advantageous to deal with the system function $H(p)$ of the predictor, rather than with its impulsive response $W(t)$. In such cases one can use a transformed form of the integral equation (36) which is as follows:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} H(j\omega)S(\omega^2)e^{j\omega t}d\omega = \lambda_0 + \lambda_1 t + \cdots + \lambda_n t^n$$
$$+ \frac{1}{2\pi}\int_{-\infty}^{\infty} S_M(\omega^2)K(j\omega)R(-j\omega)e^{j\omega t}d\omega. \quad (66)$$

Using Eq. (64), the solution of this equation may be written directly as

$$H(p) = \int_0^T (A_0 + A_1 t + \cdots + A_n t^n)e^{-pt}dt$$

$$+ \frac{B_1}{p+\alpha_1} + \frac{B_2}{p+\alpha_2} + \cdots + \frac{B_{2m}}{p+\alpha_{2m}}$$

$$- \left\{ \frac{B_1 \exp(\alpha_1 T)}{p+\alpha_1} + \frac{B_2 \exp(\alpha_2 T)}{p+\alpha_2} + \cdots \right.$$

$$\left. + \frac{B_{2m} \exp(\alpha_{2m}T)}{p+\alpha_{2m}} \right\} e^{-pT} + \frac{1}{2\pi}R(p)\int_0^T dt e^{-pt}$$

$$\times \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{A(\omega^2)}K(j\omega)R(-j\omega)e^{j\omega t}d\omega$$

$$+ C_1 + C_2 p + \cdots + C_{l-m}p^{l-m-1}$$

$$+ (D_1 + \cdots + D_{l-m}p^{l-m-1})e^{-pT}. \quad (67)$$

The undetermined constants involved in this expression are found in the same manner as in the case of $W(t)$, that is, $H(p)$ as given by Eq. (67) is substituted into the integral equation (66) and the resulting expression is treated as an identity. The $2l$ linear relations between

$A_0, A_1, A_2, \cdots$, etc., which are obtained in this manner are adjoined to the $n+1$ relations resulting from Eq. (65); then the system of linear equations in the unknown constants is solved for $A_0, A_1, A_2, \cdots$, etc.

In order to facilitate application of the techniques described in the preceding sections, a summary of the procedure for the determination of $W(t)$ (or $H(p)$) is given in Section VI. Furthermore, actual use of the procedure is illustrated by a few practical examples at the end of the section.

VI. SUMMARY OF THE PROCEDURE FOR DETERMINATION OF $W(t)$ AND $H(p)$

The complete expressions for $W(t)$ (the impulsive response of the optimum predictor) and $H(p)$ (the system function of the optimum predictor) are given by Eqs. (64) and (67). In order to avoid the necessity for reference to preceding sections, the meaning of all symbols appearing in these equations is given:

$u(t) =$ unit step function.
$T =$ duration of sampling (settling time).
$A_0, A_1, \cdots, A_n, B_1, B_2, \cdots, B_{2m}, C_1, C_2, \cdots, C_{l-m}, D_1, D_2, \cdots,$
$\quad D_{l-m} =$ undetermined constants.
$n =$ degree of the polynomial component of the input signal.
$S_M(\omega^2) =$ spectral density of $M(t)$ [$M(t)$ is the stationary part of the input signal].
$S(\omega^2) =$ spectral density of $M(t)+N(t)$ [$N(t)$ is the input noise].
$A(\omega^2) =$ numerator of $S(\omega^2)$.
$B(\omega^2) =$ denominator of $S(\omega^2)$.
$Q(j\omega) =$ a factor of $A(\omega^2)$ containing all the zeros in right half of the $j\omega$-plane.
$R(j\omega) =$ a factor of $B(\omega^2)$ containing all the zeros in the left half of the $j\omega$-plane.
$2l =$ degree of $B(\omega^2)$.
$2m =$ degree of $A(\omega^2)$.
$\alpha_1, \alpha_2, \cdots, \alpha_{2m} =$ roots of the characteristic equation $A(-p^2)=0$.
$\delta(t) =$ unit impulse function.
$\delta^{(\nu)}(t) = \nu$th derivative of $\delta(t)$.

The undetermined constants occurring in the expression for $W(t)$ [and $H(p)$] can be found in the following manner.

1. $W(t)$ as given by Eq. (64) is substituted into the integral equation (36) and the resulting expression is treated as an identity. This gives $2l$ homogeneous linear equations in the unknown constants. Same equations can be obtained by substituting $H(p)$, as given by Eq. (67), into the integral equation (66).

2. $W(t)$ as given by Eq. (64) is substituted into the $n+1$ constraint equations

$$\int_0^T \tau^\nu W(\tau)d\tau = \mu_\nu, \quad \nu = 0, 1, \cdots, n. \qquad (14)$$

where the μ_ν are determined by the choice of the prediction operator $K(p)$ [cf. Eq. (19)]. The resulting $n+1$ linear equations in the unknown constants are adjoined to the $2l$ equations obtained from (1). The set of $2l+n+1$ linear equations thus obtained is solved for the undetermined constants $A_0, A_1, \cdots, D_{l-m}$. This concludes the process of determining $W(t)$ [or $H(p)$].

652

Example 1. Wiener's Theory

Wiener's theory is, in the main, a study of the particular case in which $P(t) \equiv 0$, $T = \infty$ and $K(p) = e^{-\alpha p}$. For this case Eq. (64) gives

$$W(t) = u(t)\left\{ \frac{1}{2\pi} R(p) \right.$$

$$\times \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{A(\omega^2)} K(j\omega) R(-j\omega) e^{j\omega(t-\alpha)} d\omega$$

$$\left. + B_1 \exp(\alpha_1 t) + \cdots + B_{2m} \exp(\alpha_{2m} t) \right\}. \quad (68)$$

The exponential terms appearing in Eq. (68) may be made to vanish through a slight rearrangement of the factors in the first term of Eq. (68). The resulting expression for $W(t)$ is the same as that obtained by using Wiener's theory, namely,

$$W(t) = u(t) \frac{1}{2\pi} \frac{R(p)}{Q(p)}$$

$$\times \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{Q(-j\omega)} R(-j\omega) e^{j\omega(t-\alpha)} d\omega. \quad (69)$$

The rearrangement amounts, essentially, to choosing a particular solution of Eq. (61) which differs from the one chosen before by the exponential terms of Eq. (68). The same result may be achieved directly by choosing $H_1(p)$ [cf. Eq. (47)] as

$$H_1(p) = R(p)/Q(p).$$

With this choice of $H_1(p)$ [in place of the one expressed by Eq. (47)] the various quantities entering Eq. (55) become:

$$S'(\omega^2) = 1,$$
$$\psi_M'(\tau) + \psi_N'(\tau) = \delta(\tau),$$
$$S_M'(\omega^2) = S_M(\omega^2) |R(j\omega)/Q(j\omega)|^2,$$
$$K'(j\omega) = K(j\omega) Q(j\omega)/R(j\omega),$$

and hence the integral equation (55) reduces to

$$W_2(t) = u(t) \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_M(\omega^2)}{Q(-j\omega)} K(j\omega) R(-j\omega) e^{j\omega t} d\omega; \quad (55a)$$

Eq. (69) then follows immediately from the relation connecting $W(t)$ and $W_2(t)$ [cf. Eq. (46)].

Example 2

The assumptions made here are as follows:

1. $M(t) \equiv 0.$

2. $n = 1.$

3. $\psi_N(\tau) = e^{-a|\tau|}; \quad S_N(\omega^2) = \dfrac{2a}{\omega^2 + a^2}.$

The choice of the prediction operator is left open.

Solution

For this case $A(-p^2) = 2a$, and hence $\alpha_1 = \cdots = \alpha_{2m} = 0$. Also, $l = 1$, $M \equiv 0$; hence Eq. (64) gives

$$W(t) = A_0 + A_1 t + C_1 \delta(t) + D_1 \delta(t - T), \quad (70)$$

and

$$H(p) = \frac{A_0}{p}(1 - e^{-pT})$$

$$+ \frac{A_1}{p}\left(\frac{1}{p} - \frac{e^{-pT}}{p} + T e^{-pT} \right) + C_1 + D_1 e^{-pT}. \quad (71)$$

Substituting $H(p)$ as given by Eq. (71) into the integral equation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2a}{a^2 - p^2} H(p) e^{pt} dp = \lambda_0 + \lambda_1 t, \quad 0 \le t \le T, \quad (72)$$

and requiring that this equation be satisfied identically, leads to the following relations:

$$a A_0 - A_1 - a^2 C_1 = 0, \quad (73)$$

$$a A_0 + (aT + 1) A_1 - a^2 D_1 = 0. \quad (74)$$

Furthermore, substituting $W(t)$ as given by Eq. (70) into the constraint equations

$$\int_0^T W(\tau) d\tau = \mu_0 \quad (75)$$

and

$$\int_0^T \tau W(\tau) d\tau = \mu_1, \quad (76)$$

yields

$$A_0 T + A_1(T^2/2) + C_1 + D_1 = \mu_0, \quad (77)$$

and

$$A_0(T^2/2) + A_1(T^3/3) + D_1 T = \mu_1. \quad (78)$$

The unknown constants A_0, A_1, C_1, and D_1 can be readily found from the solution of Eqs. (73), (74), (77), and (78). Thus,

$$A_0 = \mu_0 \frac{4a(a^2 T^2 + 3aT + 3)}{(a^2 T^2 + 6aT + 12)(aT + 2)}$$

$$- \mu_1 \frac{6a^2}{a^2 T^2 + 6aT + 12}, \quad (79)$$

$$A_1 = -\mu_0 \frac{6a^2}{a^2 T^2 + 6aT + 12} + \mu_1 \frac{12a^2}{T(a^2 T^2 + 6aT + 12)}, \quad (80)$$

$$C_1 = \mu_0 \frac{2(2a^2 T^2 + 9aT + 12)}{(a^2 T^2 + 6aT + 12)(aT + 2)}$$

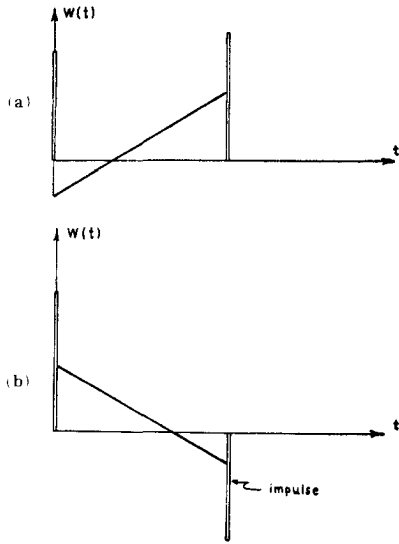$$- \mu_1 \frac{6(aT + 2)}{T(a^2 T^2 + 6aT + 12)}, \quad (81)$$

FIG. 4. Shapes of optimum weighting functions. (a) For the best value of $s(t)$. (b) For the best value of $s(t-\alpha)$ (α-negative).

and

$$D_1 = -\mu_0 \frac{2aT(aT+3)}{(a^2T^2+6aT+12)(aT+2)}$$

$$+ \mu_1 \frac{6(aT+2)}{T(a^2T^2+6aT+12)}. \quad (82)$$

Whenever the stationary part of the input signal is zero [i.e., $M(t) \equiv 0$], the mean-square value of the prediction error assumes the simple form

$$\sigma^2 = \mu_0\lambda_0 + \mu_1\lambda_1 + \cdots + \mu_n\lambda_n, \quad (83)$$

which may be readily established by substituting Eq. (36) into the general expression for σ^2 [cf. Eq. (33)]. For the particular case under consideration Eq. (83) gives

$$\sigma^2 = \mu_0{}^2 \frac{8(a^2T^2+3aT+3)}{(a^2T^2+6aT+12)(aT+2)}$$

$$+ \mu_1{}^2 \frac{24a}{T(a^2T^2+6aT+12)}$$

$$- \mu_0\mu_1 \frac{24a}{(a^2T^2+6aT+12)}. \quad (84)$$

The expressions given above are valid for any choice of the prediction operator. For the particular case in which the predictor is called upon to furnish the best possible estimate of the present value of $\dot{s}(t)$, the values of μ_0 and μ_1 are, respectively,

$$\mu_0 = 0$$
$$\text{[cf. Eq. (21)]} \quad (85)$$
$$\mu_1 = -1.$$

On the other hand, in the case of the estimation of $s(t+\alpha)$ (i.e., the value of $s(t)$ α-seconds in the future)

$$\mu_0 = 1$$
$$\text{[cf. Eq. (24)]} \quad (86)$$
$$\mu_1 = -\alpha.$$

The shapes of $W(t)$ for these two particular cases are illustrated in Fig. 4.

Example 3

The case to be considered here is the same as that treated in Example 2, except that the auto-correlation function of $N(t)$ is assumed to be of the form

$$\psi_N(\tau) = e^{-a|\tau|} \cos\omega_0\tau, \quad (87)$$

with the associated spectral density function being

$$S_N(\omega^2) = \frac{2a(a^2+\omega_0{}^2+\omega^2)}{\omega^4 + 2(a^2-\omega_0{}^2)\omega^2 + (a^2+\omega_0{}^2)^2}. \quad (88)$$

This form of spectral density function is of considerable practical importance since it provides a reasonably good approximation to many of the actual spectra encountered in practice.

Solution

By Eq. (64), the weighting function for this case is of the form

$$W(t) = A_0 + A_1 t + B_1 e^{bt} + B_2 e^{-bt}$$
$$+ C_1\delta(t) + D_1\delta(t-T), \quad 0 \leq t \leq T, \quad (89)$$

where $b = (a^2+\omega_0{}^2)^{\frac{1}{2}}$. Substituting $W(t)$ as given by Eq. (89) into Eq. (36), and requiring that Eq. (36) be satisfied by $W(t)$ establishes four linear algebraic equations between the six constants $A_0, A_1, B_1, B_2, C_1, D_1$. These are:

$$-2ab^2A_0 + 2(a^2-\omega_0{}^2)A_1 - b^3B_1 + b^3B_2 + C_1 = 0,$$

$$2\omega_0{}^2b^2A_0 - 4a\omega_0{}^2A_1 + b^3(b-a)B_1 + b^3(b+a)B_2 = 0,$$

$$-2ab^2A_0 + 2(\omega_0{}^2-a^2-aTb^2)A_1$$
$$+ \frac{e^{bT}}{2b}B_1 - \frac{e^{-bT}}{2b}B_2 + D_1 = 0, \quad (90)$$

$$2\omega_0{}^2b^2A_0 + 2\omega_0{}^2(2a+b^2T)A_1 + b^3(b+a)e^{bT}B_1$$
$$+ b^3(b-a)e^{-bT}B_2 = 0.$$

The constraints imposed by the prediction operator $K(p)$ are given by Eq. (14); they are:

$$\int_0^T W(\tau)d\tau = \mu_0, \quad (75)$$

and

$$\int_0^T \tau W(\tau)d\tau = \mu_1. \quad (76)$$

654

The requirement that $W(t)$ must satisfy Eqs. (75) and (76) leads to two additional linear equations. These are:

$$2bTA_0+bT^2A_1+2(e^{bT}-1)B_1-2(e^{-bT}-1)B_2 \\ +2bC_1+2bD_1=\mu_0,$$

and

$$3b^2T^2A_0+2b^2T^3A_1+b[e^{bT}(bT-1)+1]B_1 \\ -6[e^{-bT}(bT+1)-1]B_2+6b^2D_1=\mu_1. \quad (91)$$

The coefficients A_0, A_1, B_1, B_2, C_1, D_1 of Eq. (89) are the solutions of the six linear equations (90) and (91).

The expression for the mean-square value of the prediction error can be conveniently expressed in terms of A_0 and A_1. Thus, making use of Eq. (83) it is readily found that

$$\sigma^2=(2a/b^2)(\mu_0A_0+\mu_1A_1). \quad (92)$$

This completes essentially the solution of the problem.

Potential Flow into Circumferential Openings in Drain Tubes*

Don Kirkham

Iowa State College Agricultural Experiment Station, Department of Agronomy and Department of Physics, Ames, Iowa

(Received December 22, 1949)

A theoretical analysis of the effect of the spaces between drain tube units as used in the artificial drainage of soil is given. The problem is one of potential flow; therefore, the results are applicable to heat flow, etc. The basic problem solved is that for axially symmetric flow from an external cylindrical boundary at constant potential to a series of equal, equally spaced openings at a lower potential, all located axially on, and comprising a part of, the otherwise impervious drain tube. The radii of the open sections and impermeable sections of the drain tube are equal. The basic problem is extended to obtain the solution to the practical problem—the seepage of ground water into drain tubes beneath a horizontal water table. The exact solution of the basic problem is not suitable for numerical work. Accordingly, approximate solutions of specified uncertainty are derived and are utilized for tabulation of numerical results. As an example, the analysis shows, in the case of 6 in. diameter drain tubes having 1 ft. long impermeable sections and buried 4 ft. deep in uniformly permeable soil, that increasing the openings from $\frac{1}{32}$ in. width to $\frac{1}{4}$ in. width will increase the flow 36 percent; while embedding the tubes in gravel, to make the $\frac{1}{32}$ in. openings of effectively infinite width, will increase the flow 180 percent.

IN the drainage of soil, for agricultural use, for foundations, roads, or for dams, excess ground water is generally removed by drain tubes installed two or more feet below the soil surface. These tubes consist, in most cases, of 1 ft. long sections of impervious pipe, either tile or cement, fitted together axially, except for a space left between individual units, to permit the water to enter. In practice the opening between the units is usually small. It may be less than 1/64 in. or, as much as $\frac{1}{4}$ in. But if the tubes are embedded in gravel, as is sometimes the case, the open space between the pipe units becomes effectively infinite. This is a consequence of the negligible loss of head which results when water seeps through gravel as compared with water seeping through soil. In this paper the effect of width of opening between the pipe units on drainage rate will be analyzed. As the problem is one of potential theory, the results will be applicable to other physical problems as flow of heat, electricity, etc.[1]

In the actual drainage problem the drain tubes are installed in a nearly horizontal position in the soil.

Therefore if the soil is water-saturated to the surface of the ground, the soil surface is an equipotential plane, and the flow problem may be solved by the method of images.

AXIAL FLOW

Before applying the method of images it is convenient to solve a simpler problem: that of axially symmetrical flow for the system indicated in Fig. 1. Here, because of symmetry, only a section $ABCD$ need be considered. This section is bounded by inner and outer cylindrical surfaces at $r=a$ and $r=b$; by a plane $z=0$, midway between an open space of width $2c$; and a plane $z=s$, $2s$ being the distance between centers of openings.

Taking ϕ as the potential function, V_a as the constant potential over the openings of the drain tube, and V_b the constant potential over the outer cylinder, the boundary conditions are—I: $\phi=V_a$, $r=a$, $0\leq z\leq c$; II: $\partial\phi/\partial r=0$, $r=a$, $c<z\leq s$; III: $\partial\phi/\partial z=0$, $z=s$; IV: $\phi=V_b$, $r=b$; V: $\partial\phi/\partial z=0$, $z=0$.

The exact solution of $\nabla^2\phi$ which satisfies the above conditions can be formulated. This exact solution, is not, however, tractable in calculations. Therefore, a simpler, approximate solution is sought. Simplification is looked for through substitution of approximately correct,

---

[1] M. Muskat, *The Flow of Homogeneous Fluids through Porous Media* (McGraw-Hill Book Company, Inc., New York, 1937, or J. W. Edwards, Inc., Ann Arbor, Michigan, 1946), p. 140.