Gen Relativ Gravit (2009) 41:581–660
DOI 10.1007/s10714-009-0760-7

GOLDEN OLDIE

Republication of: Relativistic cosmology

George F. R. Ellis

Published online: 30 January 2009

An editorial note to this paper and a biography can be found via doi:10.1007/s10714-009-0759-0.

Original paper: G. F. R. Ellis: Relativistic cosmology, In: Sachs, R.K. (ed.) Proceedings of the International School of Physics "Enrico Fermi", Course 47: General relativity and cosmology, pp. 104–182. Academic Press, New York and London (1971). Reprinted with the kind permission of Elsevier Ltd, the current owner of the Academic Press copyright, and of the author.

Editorial responsibility: A. Krasiński, e-mail: akr@camk.edu.pl.

G. F. R. Ellis (✉)
Mathematics Department, University of Cape Town,
Rondebosch 7701, Cape Town, South Africa
e-mail: George.Ellis@uct.ac.za

Relativistic Cosmology.

G. F. R. Ellis

Department of Applied Mathematics and Theoretical Physics
University of Cambridge - Cambridge

1. – Introduction.

The aim of cosmology is to determine the large-scale structure of the physical universe.

One rather vague assumption is at the basis of virtually all present-day work on the subject. We shall call this (following BONDI [1]) the *Copernican Principle*. It is the assumption:-

« We do not occupy a privileged position in space-time ». We may apply this to two different kinds of experiment. First, considering laboratory experiments, we may interpret the principle as implying that local physical laws are the same everywhere in the universe. Secondly, considering astronomical observations, we may assume that our view of the universe is not a preferred picture. (We can, if we wish, proceed without this second assumption; the result is to obtain, beside the picture of the universe discussed in Sect. 7, more general situations which can be regarded as rather unlikely models of the universe.)

The over-all picture we observe is well known (see BONDI [1], MCVITTIE [2]): on a moderately large scale ($\gtrsim 3\cdot 10^8$ light years) the distribution of clusters of galaxies is approximately isotropic about us. The general motion is an over-all expansion, random velocities relative to this general motion being rather small. Thus we are able to determine an average velocity vector which represents to a good approximation the over-all motion of matter in our « local » vicinity; the Copernican principle then suggests that we can assume the existence of such a vector at each point of space-time. The principle will also be taken to imply that corresponding to the very nearly isotropic distribution of matter and background radiation about us, any observer moving with the average velocity vector will see a very nearly isotropic distribution of matter and radiation on his celestial sphere.

Our aim will be to determine a cosmological model which will correctly predict the results of astronomical observations, and whose behaviour is determined by those physical laws which describe the behaviour of matter on scales up to those of clusters of galaxies. To find such laws, we remember that within the solar system the dominant long-range force is gravity. (We shall assume over-all electrical neutrality of heavenly bodies.) Accepting (with some reservations until the consequences of Dicke's solar oblateness measurements are completely understood) that general relativity is the best theory of gravitation available for phenomena up to the scale of the solar system, we shall extrapolate by assuming it is also, to a good approximation, the law valid for gravity acting on scales about that of clusters of galaxies. If this eventually proves to lead to predictions inconsistent with observations, we may have to modify or abandon general relativity; however we would like to avoid *ad hoc* modifications introduced solely for cosmological convenience! In these lectures we shall therefore assume that general relativity is valid. (The alternative theories so far proposed may be studied by much the same methods as will be used here to study the conventional theory.)

On this view, the geometry of space-time will be determined by its energy content. We may describe the matter and radiation content of the universe in two convenient ways: by using a particle distribution function (as discussed in Ehlers' lectures) or by using a fluid approximation. It is the latter course (first introduced by EINSTEIN in 1917 [3], and followed in nearly all the classical cosmological discussions) we follow in these lectures: a continuum approximation is used, and the average velocity vector may be thought of as representing the velocity of fluid « particles », which are in this case clusters of galaxies.

In Sect. 2-5, we shall obtain dynamical relations valid in any cosmological model, and illustrate these relations by applying them to some particular cosmological models. There is a very close parallel between fluid dynamics in Newtonian theory and in general relativity, and we shall emphasize this by giving corresponding relations in parallel. The basis of this close correspondence is that the fluid equations we shall use are based on *relative* motion. Corresponding to the applications we have in mind, the Newtonian fluid will always be a self-gravitating fluid.

In Sect. 6, we shall derive observational relations valid for any cosmological model and give as an illustration of the methods used, their application to the standard isotropic cosmological models. (The optical relations valid in special relativity are a special case of these observational relations.)

In Sect. 7, we shall present a broad over-all picture of the observable universe, on the basis of recent observations. This over-all picture serves as a background for more detailed observational discussions.

Conventions used are as follows:

in *general relativity*,

arbitrary coordinates x^a (Latin letters run from 1 to 4) are used in the 4-dimensional space-time with metric tensor g_{ab} (signature $(- + + +)$). A semi-colon (;) denotes the covariant derivative with respect to the metric g_{ab}; so $g_{ab;c} = 0$.

in *Newtonian theory*,

unless otherwise stated, fixed curvilinear co-ordinates x^ν (Greek letters run from 1 to 3) and a natural time co-ordinate t is used in the (3+1)-dimensional space-time. A comma (,) denotes the covariant derivative with respect to the flat 3-space metric $h_{\mu\nu}$ (signature $(+ + +)$); so $h_{\mu\nu,\sigma} = 0$.

Square brackets denote skew-symmetrization and round brackets denote symmetrization (SCHOUTEN [4]), so $T_{(ab)} = \frac{1}{2}(T_{ab} + T_{ba})$,

$$S_{[abc]} = \tfrac{1}{6}(S_{abc} + S_{cab} + S_{bca} - S_{acb} - S_{bac} - S_{cba}) .$$

The general-relativity notation and units we use are the same as those used by EHLERS in this volume; in particular, note that we use units in which the speed of light is unity. As a notational convenience, we shall often use the general-relativity name of a quantity to refer to its Newtonian analogue as well.

This article is a revised and corrected version of lectures given at Varenna. It presupposes a knowledge of the basic ideas and observations of cosmology (see for example BONDI [1], RINDLER [5]). It should be supplemented by review articles written from other view-points, such as those by HECKMANN and SCHÜCKING [6, 7], ZEL'DOVICH [8], DAVIDSON and NARLIKAR [9], RINDLER [10], SCHÜCKING [11] and SANDAGE [12]; it is intended to be complementary to the articles by EHLERS and SCIAMA in this volume. I should like to thank S. W. HAWKING, J. EHLERS and M. A. H. MACCALLUM for many useful discussions, M. A. H. MACCALLUM for assistance in preparing these notes, and M. STRANGLEMAN for patiently and carefully typing them.

2. – Kinematical quantities.

2·1. *The velocity vector field.* – We assume the existence of a unique vector field representing the average velocity of matter (or of some particular component of matter) at each point of space-time. Thus there is defined,

in *general relativity*:

the normalized 4-velocity field u^a; so

$$u_a u^a = -1 . \tag{2.1}$$

in *Newtonian theory*:

the 3-velocity

$$v^\nu .$$

To express these vectors in simple form we use

normalized comoving co-ordinates (y^ν, s),

Lagrangian co-ordinates (y^ν, t),

defined locally as follows. In some arbitrarily chosen space-section (in Newtonian theory, one of the natural space sections) of space-time, label the fluid particles by co-ordinates y^ν. At all later times, label the same particles by the same co-ordinate values, so that the fluid flow lines in space-time are the curves $\{y^\nu = \text{const}\}$. Determine the time co-ordinate by measuring proper time, from the initial space section, along the flow lines. In terms of these co-ordinates, the velocity vector takes the form

$$u^a \stackrel{*}{=} \delta^a{}_0 .$$

$$v^\nu \stackrel{*}{=} 0 .$$

In terms of

general co-ordinates x^a,

co-ordinates (x^ν, t), where x^ν are fixed curvilinear co-ordinates,

the velocity vectors are given by

$$u^a = \left.\frac{\mathrm{d}x^a}{\mathrm{d}s}\right|_{y^\nu = \text{const}}$$

$$v^\nu = \left.\frac{\mathrm{d}x^\nu}{\mathrm{d}t}\right|_{y^\sigma = \text{const}}$$

2'1.1. The projection tensor.

A $\{3+1\}$ splitting of space-time into $\{\text{space} + \text{time}\}$ is

determined at each point by u^a. The tensor

$$h_{ab} := g_{ab} + u_a u_b \tag{2.2}$$

is, at each point, a projection tensor into the rest space of an observer moving with 4-velocity u^a.

given and absolute. The Euclidean metric tensor

$$h_{\mu\nu}$$

is given in local orthonormal axes by

$$\delta_{\mu\nu} .$$

These tensors are 3-dimensional projection tensors:

$$h_{ab} u^b = 0, \quad h_a{}^c h_c{}^b = h_a{}^b, \quad h_a{}^a = 3 .$$

$$h_\mu{}^\nu h_\nu{}^\sigma = h_\mu{}^\sigma, \quad h_\mu{}^\mu = 3 .$$

By (2.2), the expression for ds^2 in general relativity is:

$$ds^2 := g_{ab}\,dx^a\,dx^b = h_{ab}\,dx^a\,dx^b - (u_a\,dx^a)^2;$$

an observer at the space-time point x^a moving with 4-velocity u^a assigns to the event $x^a + dx^a$ a spatial separation $(h_{ab}\,dx^a\,dx^b)^{\frac{1}{2}}$, and a time separation $|u_a\,dx^a|$ from himself. Note that the 3-planes defined in each tangent space by h_{ab} do *not* in general mesh together to form 3-surfaces in space-time; the condition that they do so is given in Subsect. 4.4.1.

2.1.2. Volume elements. Effective volume elements in the instantaneous rest spaces of the co-moving observers are given by

$\eta_{abcd}\,u^d$	$\eta_{\mu\nu\sigma}$

where the totally-skew pseudotensors η are defined by

$\eta^{abcd} = \eta^{[abcd]}, \quad \eta^{1234} = (-g)^{-\frac{1}{2}},$ $g := \det(g_{ab})\,.$	$\eta^{\mu\nu\sigma} = \eta^{[\mu\nu\sigma]}, \quad \eta^{123} = (h)^{-\frac{1}{2}},$ $h := \det(h_{\mu\nu})\,.$

(In terms of an orthonormal frame, these are the alternating quantities ε^{abcd}, $\varepsilon^{\mu\nu\sigma}$.)

The identities

$\eta^{abcd}\eta_{efgh} = -4!\,\delta_e{}^{[a}\delta_f{}^{b}\delta_g{}^{c}\delta_h{}^{d]},$	$\eta^{\alpha\beta\gamma}\eta_{\mu\nu\sigma} = 3!\,\delta^{[\alpha}{}_{\mu}\delta^{\beta}{}_{\nu}\delta^{\gamma]}{}_{\sigma},$

and the further identities obtained from these by contracting pairs of indices, are frequently useful in proving formulae given in the sequel. (The — sign in the general-relativity case appears because the determinant g is negative.)

2.2. *Time derivatives.* – We write the effective time derivative of a tensor T, measured by an observer moving with the standard velocity, as $\dot{T}$. Thus

(2.3) $\quad \dot{T}^{a\ldots b}{}_{c\ldots d} := T^{a\ldots b}{}_{c\ldots d;e}\,u^e$ is the covariant derivative along the particle world lines.	(2.3′) $\quad \dot{T}^{\alpha\ldots\beta}{}_{\gamma\ldots\delta} := \partial T^{\alpha\ldots\beta}{}_{\gamma\ldots\delta}/\partial t + T^{\alpha\ldots\beta}{}_{\gamma\ldots\delta,\varepsilon}\,v^{\varepsilon},$ is the convective derivative for motion with the fluid.

2·3. *The acceleration vector.* – We can represent the combined effects of gravitational and inertial forces on the fluid by the vectors

(2.4) $\dot{u}^a := u^a{}_{;b} u^b$

the « acceleration vector », which is spacelike as (2.1) implies $\dot{u}^a u_a = 0$.

(2.4′) $a_\nu := (v_\nu)^{\cdot} + \Phi_{,\nu}$

where Φ is the gravitational potential (force $= -\operatorname{grad}\Phi$).

In co-moving co-ordinates, we find

$$\dot{u}^a \stackrel{*}{=} \Gamma^a{}_{00}$$

$$a_\nu \stackrel{*}{=} \frac{\partial \Phi}{\partial x^\nu}\,.$$

Note that even in Newtonian theory, we are unable to separate the gravitational and inertial parts of a_ν invariantly if the density of matter does not go to zero at infinity; see the discussions of HECKMANN and SCHÜCKING [13], BONDI [1], TRAUTMANN [14], and references cited there.

2·4. *Relative motion of neighbouring particles.*

2·4.1. The relative position vector. Consider the world-lines of neighbouring particles 0 and G, labelled by co-moving co-ordinates y^ν and $y^\nu + \delta y^\nu$ respectively. The vector X^a with components given in co-moving co-ordinates by $X^\nu \stackrel{*}{=} \delta y^\nu$ (and in general relativity, $X^0 \stackrel{*}{=} 0$) joins these same two world-lines at all times; we therefore call it a *connecting vector.*

Using general co-ordinates in general relativity, this vector has components $X^a = (\partial x^a/\partial y^\nu)\,\delta y^\nu$. If we calculate $\dot{X}^a$, we find

$$\dot{X}^a = X^a{}_{;b} u^b = \{\partial(\partial x^a/\partial y^\nu\, \delta y^\nu)/\partial x^b + \Gamma^a{}_{bc}(\partial x^c/\partial y^\nu\, \delta y^\nu)\}\,\frac{\partial x^b}{\partial s} =$$

$$= \{\partial(\partial x^a/\partial s)/\partial x^c + \Gamma^a{}_{cb}\,\partial x^b/\partial s\}\,\frac{\partial x^c}{\partial y^\nu}\,\delta y^\nu = u^a{}_{;b} X^b\,,$$

since $\partial^2 x^a/\partial y^\nu \partial s = \partial^2 x^a/\partial s\, \partial y^\nu$ and $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$.

(Geometrically, this is the statement and that the Lie derivative of $\boldsymbol{X}$ with respect to $\boldsymbol{u}$ vanishes.) A similar calculation shows that in Newtonian theory, $\dot{X}_\nu = v_{\nu,\mu} X^\mu$. Thus a connecting vector obeys the differential equation

(2.5) $\dot{X}_a = u_{a;b} X^b\,.$

(2.5′) $\dot{X}_\nu = v_{\nu,\mu} X^\mu\,.$

In general, a connecting vector will not (in general relativity) lie in the rest space of an observer moving with 4-velocity u^a. We obtain the *relative position vector* $X^a_\perp$ of G with respect to O, by projecting the connecting vector X_a into the rest-frame of O:

$$X_{\perp a} := h_a{}^b X_b . \qquad\qquad X_{\perp\nu} := h_\nu{}^\mu X_\mu = X_\nu .$$

2·4.2. Relative velocity. The *relative velocity vector* V^a of G with respect to 0 is defined to be the spatial part of the observed rate-of-change of the relative position vector:

(2.6) $$V^a := h^a{}_b (X_\perp{}^b)^{\boldsymbol{\cdot}} ,$$

so

$$V^a = h^a{}_b (X^c h_c{}^b)_{;d} u^d .$$

(2.6′) $$V^\mu := h^\mu{}_\nu (X_\perp{}^\nu)^{\boldsymbol{\cdot}} ,$$

so

$$V^\mu = (X^\mu)^{\boldsymbol{\cdot}} .$$

Since X^a obeys the eq. (2.5), it follows (*) that

(2.7a) $$V^a = v^a{}_b X_\perp{}^b ,$$

where

(2.7b) $$v_{ab} := h_a{}^c h_b{}^d u_{c;d} ;$$

(2.7a′) $$V^\mu = v^\mu{}_\nu X_\perp{}^\nu ,$$

where

(2.7b′) $$v_{\mu\nu} := v_{\mu,\nu} ;$$

in each case, the relative velocity vector of a neighbouring particle is related to the relative position vector of that particle by a linear transformation, and the tensor determining that transformation is the spatial gradient of the velocity vector.

To examine this relation further, we split v_{ab} into its symmetric and skew-symmetric parts:

(2.8) $$v_{ab} = \theta_{ab} + \omega_{ab} ,$$

where

$$\theta_{ab} = \theta_{(ab)} , \qquad \omega_{ab} = \omega_{[ab]} ,$$

so

$$\theta_{ab} u^b = 0 = \omega_{ab} u^b ,$$

(2.8′) $$v_{\mu\nu} = \theta_{\mu\nu} + \omega_{\mu\nu} ,$$

where

$$\theta_{\mu\nu} = \theta_{(\mu\nu)} , \qquad \omega_{\mu\nu} = \omega_{[\mu\nu]} ,$$

(*) The essential properties of the covariant derivative are that it i) is linear, ii) obeys Leibniz' rule, and iii) preserves the metric, irrespective of whether indices are up or down. For example,

$$(\tfrac{1}{2} T_{ab} + f g_{ab})_{;c} = \tfrac{1}{2} T_{ab;c} + f_{;c} g_{ab} ;$$

$(S^{\nu\mu} X_\mu)_{,\sigma} = S^{\nu\mu}{}_{,\sigma} X_\mu + S^{\nu\mu} X_{\mu,\sigma}$. We occasionally need the expressions involving the Christoffel symbols $\Gamma^a{}_{bc}$, for example in deriving (2.5), (2.5′) and in proving the relations

$$\eta^{abcd}{}_{;e} = 0 = \eta^{\mu\nu\sigma}{}_{,\varkappa} , \qquad k^a{}_{;a} = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_a} (\sqrt{|g|}\, k^a) .$$

and the symmetric part into its trace and its trace-free part:

(2.9) $\theta_{ab} = \sigma_{ab} + \frac{1}{3}\theta h_{ab}$,

where

$\sigma^a{}_a = 0$, so $\sigma_{ab} = \sigma_{(ab)}$,

$\sigma_{ab} u^b = 0$ and $\theta = u^a{}_{;a}$.

(2.9′) $\theta_{\mu\nu} = \sigma_{\mu\nu} + \frac{1}{3}\theta h_{\mu\nu}$,

where

$\sigma^\mu{}_\mu = 0$, so $\sigma_{\mu\nu} = \sigma_{(\mu\nu)}$,

and $\theta = v^\nu{}_{,\nu}$.

The quantities we have defined determine the first derivatives of the velocity vector completely:

(2.10) $u_{a;b} = \omega_{ab} + \sigma_{ab} + \frac{1}{3}\theta h_{ab} - \dot{u}_a u_b$.

(2.10a) $v_{\mu,\nu} = \omega_{\mu\nu} + \sigma_{\mu\nu} + \frac{1}{3}\theta h_{\mu\nu}$,

(2.10b) $\partial v_\nu / \partial t = a_\nu - v_{\nu,\mu} v^\mu - \Phi_{,\nu}$.

We now split the relative position vector $X_{\perp_a}$ of G relative to 0 into a *relative distance* δl and a *direction* n^a; then $X_{\perp_a} = n^a \delta l$ where $n_a n^a = 1$, so $(\delta l)^2 = h^{ab} X_a X_b$ (and $n^a u_a = 0$ in general relativity).

Equating the right-hand sides of (2.6) and (2.7), we now find that *the rate-of-change of relative distance* is

(2.11) $$\frac{(\delta l)^{\cdot}}{(\delta l)} = \theta_{ab} n^a n^b = \sigma_{ab} n^a n^b + \frac{1}{3}\theta ,$$

in both cases, and that the *rate-of-change of direction* is

(2.12) $$h_a{}^b (n_b)^{\cdot} = \left(\omega_a{}^b + \sigma_a{}^b - (\sigma_{cd} n^c n^d) h_a{}^b\right) n_b .$$

These two equations are equivalent to (2.7).

Applying these equations to a cosmological model (2.11) is a generalized Hubble law allowing for possible anisotropic expansion; it is valid for distances large enough to ensure random velocities are small compared with velocities associated with the average motion of matter, but small enough for the Hubble relation to be linear and also for the change in distance of the galaxies to be relatively small during the time of light travel between the galaxies and the observer. Thus we might expect its range of validity to be roughly from 50 to 500 Mpc. Relation (2.12), which we might expect to be valid for roughly the same length scales, gives the rate-of-change of direction of neighbouring fluid elements with respect to

a Fermi-propagated basis of orthonormal vectors:

(2.13) $$h_a{}^b(e_b)^\bullet = 0$$

for each basis vector e_b,

$$e^a e_a = 1\,, \qquad e^a u_a = 0\,.$$

a nonrotating basis of orthonormal vectors:

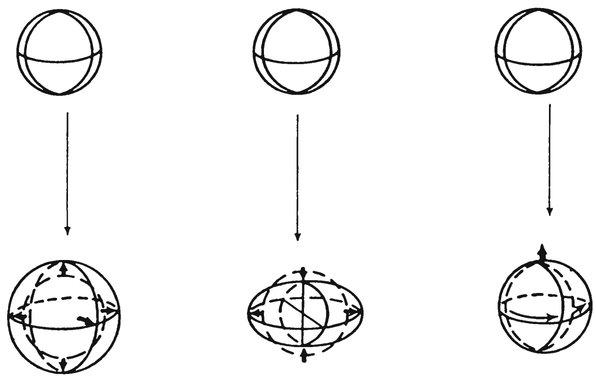
(2.13′) $$(e_\nu)^\bullet = 0$$

for each basis vector e_ν,

$$e_\nu e^\nu = 1\,.$$

Thus this relation gives the rate-of-change of position in the sky of neighbouring clusters of galaxies with respect to axes at rest in a local inertial rest frame (a rest frame determined by gyroscopes or other modern refinements of Newton's bucket experiment). [Detailed discussions showing that (2.13) corresponds to a dynamically nonrotating set of axes may be found in Synge [15] and Trautmann [16].]

2·5. *The kinematic quantities.* – It is probably easiest to understand the implications of (2.11), (2.12) by considering how a sphere of fluid particles changes during the elapse of a small increment in proper time, choosing 0 at the centre of the sphere. The results are indicated in Fig. 1.



a) b) c)

Fig. 1. – During a small time interval, *a*) the action of θ alone transforms a fluid sphere to a similar sphere of different volume but with the same orientation. *b*) The action of σ_{ab} alone distorts the sphere, leaving its volume constant and the directions of the principal axes of shear unchanged. *c*) The action of ω_a alone is to give a rigid rotation leaving one direction (the axis of rotation) fixed. As time progresses the directions of the principal axes of shear and of the axis of rotation will, in general, change.

2.5.1. Volume expansion. From (2.11) we see that θ_{ab}, the *expansion tensor*, determines the rate-of-change of distance of neighbouring particles (in our application, clusters of galaxies) in the fluid. The isotropic part of the expansion is determined by θ, the *volume expansion*. We may define a representative length l by the equation

$$\frac{l^{\cdot}}{l} = \frac{1}{3}\theta; \tag{2.14}$$

l represents completely the volume behaviour of the fluid, and corresponds to the radius function $R(t)$ in the Robertson-Walker (homogeneous and isotropic) world models. From it we can at any time t define the « Hubble constant » $H := l^{\cdot}/l = \frac{1}{3}\theta$, and the « Deceleration parameter » $q := -(l^{\cdot\cdot}/l)(1/H^2)$. If we plot the curve $l(t)$ which represents the direction-averaged rate-of-change of distance of clusters of galaxies as a function of proper time, H is the slope of this curve (units: $[\text{time}]^{-1}$) and q represents the curvature of this curve (in dimensionless units; q is positive if the expansion is slowing down).

2.5.2. Shear. The *shear tensor* σ_{ab} determines the distortion arising in the fluid flow, leaving the volume invariant. The directions of the principal axes of shear (*i.e.* of any eigenvector of the tensor σ_{ab}) are unchanged by the distortion, but all other directions are changed. The magnitude of σ_{ab} is the *shear* σ, defined by $\sigma^2 = \frac{1}{2}\sigma^{ab}\sigma_{ab}$; $\sigma^2 \geqslant 0$, and $\sigma = 0 \Leftrightarrow \sigma_{ab} = 0$.

2.5.3. Vorticity. The *vorticity tensor* ω_{ab} determines a rigid rotation of clusters of galaxies with respect to a local inertial rest frame. We may also represent vorticity by the vorticity vector ω_a, where

$$\omega^a := \frac{1}{2}\eta^{abcd}u_b\omega_{cd} \Leftrightarrow \omega_{ab} = \eta_{abcd}\omega^c u^d\,, \tag{2.15}$$

so $\omega^a u_a = 0 = \omega^{ab}u_b$, $\omega^a\omega_{ab} = 0$.

$$\omega^{\nu} := \frac{1}{2}\eta^{\nu\mu\sigma}\omega_{\mu\sigma} \Leftrightarrow \omega_{\mu\nu} = \eta_{\mu\nu\sigma}\omega^{\sigma}\,, \tag{2.15'}$$

so $\omega_{\mu\nu}\omega^{\nu} = 0$.

The direction of this vector is the axis of the rotation of the matter, since if we choose n_a in the direction of ω_a when vorticity alone is nonzero, we find $h_a{}^b(n_b)^{\cdot} = 0$; this direction is left invariant by the rotation. Its magnitude is the *vorticity* ω, where $\omega := (\omega^a\omega_a)^{\frac{1}{2}} = (\frac{1}{2}\omega^{ab}\omega_{ab})^{\frac{1}{2}}$; as ω_{ab} is spacelike, $\omega \geqslant 0$ and $\omega = 0 \Leftrightarrow \omega_a = 0 \Leftrightarrow \omega_{ab} = 0$.

We may note that general-relativistic cosmological models are no more and no less « anti-Machian » than Newtonian cosmological models, in the sense that in both cases a local dynamical rest-frame is nonrotating with respect to a direction average over fairly distant galaxies, if and only if $\omega = 0$.

2˙5.4. General motion. In a general fluid flow, both ω and σ will be nonzero. It is then not immediately obvious from (2.12), (except in the special case in which the vorticity vector is a shear eigenvector), if there exists any direction left invariant by the rotation. However we can show that such a direction exists by the following general argument.

On a unit 2-sphere which represents the sky, mark the positions of neighbouring fluid particles at some instant. Then mark on the sphere at the position of each particle, an arrow representing the observed rate-of-change of position of that particle. The set of all such arrows forms a vector field on the 2-sphere, and any vector field on a 2-sphere must have a fixed point; in this case, the fixed point corresponding to a direction e_a left invariant by the fluid motion. Now (2.12) shows that $-e_a$ is also left invariant by the motion. Hence if we can use a fluid approximation to represent the motion of clusters of galaxies, there exists at least one direction in the sky such that the positions in the sky of nearby clusters in that direction and in the opposite direction are instantaneously fixed in a local inertial rest frame. (As the universe evolves this direction itself will, in general, change.)

2˙5.5. Kinematic quantities in the universe. Since ω_a, σ_{ab} and θ determine the relative motion of galaxies in a cosmological model, we should like to determine the values of these quantities in the observed universe. In principle we can compare observations with the theoretical expressions (2.11) and (2.12), (the left-hand side of (2.11) is observable in terms of red-shift and brightness measurements, cf. Subsect. 6˙2.1), to evaluate θ, ω_a and σ_{ab} at the present time t_0. The value $\theta_0 = 3H_0 = 3 \times (1.3 \cdot 10^{10}\,\mathrm{y})^{-1}$ is probably correct to within a factor 2 (cf. the lectures in this volume by Sciama and Burbidge), but we can only obtain rather poor limits on ω_0 and σ_0 from direct observations. The condition that the systematic motion of galaxies is *away* from us in all directions (there are no directions in which we see a systematic blue-shift effect in galactic spectra) imposes the restriction $\sigma_0 < \frac{1}{3}\theta_0$. More detailed examination of the direct evidence gives us the limits

$$\omega_0 \lesssim \tfrac{1}{3}\theta_0\,, \qquad \sigma_0 \lesssim \tfrac{1}{4}\theta_0 \tag{2.16}$$

(Kristian and Sachs [17]). We can obtain much better limits on ω_0 and σ_0 by *indirect* arguments (see Subsect. 7˙2).

2˙5.6. Other applications. While we have been considering these quantities primarily in a cosmological context, they can of course be defined in other situations where the fluid approximation is valid in astrophysics. Thus in a static (nonrotating) star model, we would have $\theta = \sigma = \omega = 0$. In the study of our own galaxy, where stars may be taken to constitute the fluid particles, observations indicate that $\theta \simeq 0$ near us, while Oort's constants A and B suffice (as the flow is approximately two-dimensional) to determine the shear and vorticity respectively (FREEMAN [18]; VOLTJER [19]).

Further discussion of these concepts in Newtonian theory may be found, for example, in SERRIN [20] and BATCHELOR [21]. A systematic development in the general-relativistic case is given in Ehlers' review article [22].

3. – Conservation of energy and momentum.

3·1. *The average velocity.* – We will take the average velocity vector to be determined by the condition that it is the « baricentric » velocity. An observer moving with an arbitrary 4-velocity can define from the rest-masses m_* and 3-velocities V_* of particles he observes in a (3-) volume element $\mathrm{d}V$, the mass-current density vectors

(3.1) $$J^a = \left(\frac{1}{\mathrm{d}V}\sum_{\substack{\text{particles}\\ \text{in d}V}} m_* V_*, \frac{1}{\mathrm{d}V}\sum_{\substack{\text{particles}\\ \text{in d}V}} m_*\right).$$

(3.1′) $$J = \frac{1}{\mathrm{d}V}\sum_{\substack{\text{particles}\\ \text{in d}V}} m_* V_* .$$

Then the average velocity vector is given by

defining it as the (space-time) direction u^a of J^a:

$$J^a = \varrho u^a , \qquad u^a u_a = -1 .$$

Then ϱ is the particle rest-mass density measured by an observer travelling with 4-velocity u^a.

first defining the density ϱ of the matter:

$$\varrho = \frac{1}{\mathrm{d}V}\sum_{\substack{\text{particles}\\ \text{in d}V}} m_* .$$

Then the velocity v^ν is defined by

$$J^\nu = \varrho v^\nu .$$

Conservation of particle rest-mass now implies the conservation equations

(3.2) $$(\varrho u^a)_{;a} = 0 \Leftrightarrow \dot{\varrho} + \varrho\theta = 0 ,$$

(3.2′) $$\frac{\partial \varrho}{\partial t} + (\varrho v^\nu)_{,\nu} = 0 \Leftrightarrow \dot{\varrho} + \varrho\theta = 0 ,$$

which can be integrated, using (2.14), to give

(3.3) $$\varrho l^3 =: M , \qquad \dot{M} = 0 .$$

This is the equation of conservation of matter along the world-lines, implying that the total rest mass $\int_V \varrho\, \mathrm{d}V$ in any small comoving volume element V is constant.

There are two situations in which this discussion has to be modified in the relativistic case. The first is the case of particles of zero rest mass (photons, neutrinos, etc.) when the definition of u^a proceeds as before (*) if we define J^a as the number-current density instead of the mass-current density (3.1); then ϱ is the particle number density, which is again conserved in general. The second situation is that in which the fluid is far from thermal equilibrium (**), in particular when nuclear reactions or pair production are taking place. The ϱ will not be conserved, so it will be appropriate to define u^a (and a new corresponding conserved quantity ϱ) from a baryon or charge current density, instead of the mass-current density used here.

A more thorough discussion of these alternatives, and of equations (3.1)-(3.3), may be found in Ehlers' lectures.

3.2. *Conservation of energy and momentum.* – In Newtonian theory, the momentum conservation equation is the Navier-Stokes equation

$$(V_\nu)^\cdot = -\Phi_{.\nu} - \frac{1}{\varrho}(p_{.\nu} + \pi_\nu{}^\mu{}_{.\mu}) ,$$

where p is the isotropic pressure of matter and $\pi_{\mu\nu}$ is the (trace-free) anisotropic pressure; we can rewrite this as

$$\varrho a_\nu + (p_{.\nu} + \pi_\nu{}^\mu{}_{.\mu}) = 0 . \tag{3.4}$$

The thermal balance eq. ((3.13′) below) is deduced separately.

In general relativity, the energy-momentum tensor of matter can be split up uniquely with respect to any timelike vector u^a (and so in particular, with respect to the average velocity vector u^a) to give

$$T_{ab} = \mu u_a u_b + (q_a u_b + u_a q_b) + p h_{ab} + \pi_{ab} , \tag{3.5}$$

where $q_a u^a = 0$, $\pi^a{}_a = 0$, $\pi_{ab} u^b = 0$. Then μ is the total (relativistic) energy density of matter measured by u^a; the specific internal energy density ε of the fluid is defined by $\mu = \varrho(1 + \varepsilon)$. q_a is the energy flux relative to u^a (which will represent processes such as diffusion and heat conduction), p is the isotropic pressure, and π_{ab} is the anisotropic matter pressure (due to processes such as viscosity).

(*) Providing the particles do not all move in precisely the same direction.

(**) If the material is too far from thermal equilibrium a fluid description may be inappropriate (cf. Stewart [23]).

The relativistic equations of conservation of energy and momentum are

$$T^{ab}{}_{;b} = 0 . \tag{3.6}$$

Using the decomposition (3.5) of T_{ab} and the decomposition of $u_{a;b}$ given in Sect. 2, we can rewrite eqs. (3.6) as the equations

$$\dot{\mu} + (\mu + p)\theta + \pi^{ab}\sigma_{ab} + q^a{}_{;a} + q^a \dot{u}_a = 0 , \tag{3.7a}$$

$$(\mu + p)\dot{u}_a + h_a{}^c(p_{;c} + \pi_c{}^b{}_{;b} + \dot{q}_c) + (\omega_a{}^b + \sigma_a{}^b + \tfrac{4}{3}\theta h_a{}^b)q_b = 0 . \tag{3.7b}$$

3˙3. *Equations of state.* – We only get physics into the picture when we specify further the properties of T_{ab}. We may do this by giving a prescription for defining ϱ and T_{ab} (in Newtonian theory, ϱ, p, $\pi_{\mu\nu}$ and the heat-flow vector q_ν) from a particle distribution function which obeys suitable equations (c.f. Ehlers' lectures); in the fluid approximation however, we do so by giving equations restricting ϱ, p, μ, q_a and π_{ab}.

3˙3.1. General restrictions. A general restriction one would normally put on the matter is that its energy density be positive. The restrictions of this kind we shall require the fluid to obey are

$$\mu + p > 0 , \tag{3.8a}$$

$$\mu + 3p > 0 . \tag{3.8b}$$

$$\varrho > 0 . \tag{3.8'}$$

(These restrictions will be used in Sect. 3˙4 and 5˙1.1). One would not expect (3.8′) to be violated under any circumstances; (3.8*a*) and (3.8*b*) can only be violated, assuming μ is positive, if the pressure takes extremely large negative values. Thus if μ is 1 g/cm^3, (3.8*a*), (3.8*b*) can only be violated if $p \leqslant -10^{15}$ atm (GEROCH [24]).

Further qualitative restrictions one would usually impose are:

(3.9*a*): that the fluid should be stable against local mechanical instability; and, in the relativistic case,

(3.9*b*): that the speed of sound should be less than the speed of light.

3˙3.2. Phenomenological equations. To obtain detailed equations of state, we compare one-component fluids in General relativity and in Newtonian theory, assuming the bulk viscosity is negligible. Defining

$$v = \frac{1}{\varrho}, \tag{3.10}$$

so v is the specific volume of the fluid (*), we assume there is an equation of state of the form

$$\varepsilon = \varepsilon(p, v) \tag{3.11}$$

where ε is the specific internal energy density of the fluid. Then we can define the temperature $T(p, v)$ and the specific entropy $S(p, v)$ by

$$\mathrm{d}\varepsilon + p\,\mathrm{d}v = T\,\mathrm{d}S\,. \tag{3.12}$$

The equation of conservation of thermal energy is then

(3.13) $\varrho T\dot{S} = -(\pi_{ab}\sigma^{ab} + q^a{}_{;a} + \dot{u}_a q^a)\,,$ | (3.13′) $\varrho T\dot{S} = -(\pi_{\mu\nu}\sigma^{\mu\nu} + q^\nu{}_{,\nu})\,,$

[(3.13) is just (3.7a) rewritten in terms of the new variables], and we find that the phenomenological equations

(3.14a) $\pi_{ab} = -\lambda\sigma_{ab}\,,$ | (3.14a') $\pi_{\mu\nu} = -\lambda\sigma_{\mu\nu}\,,$

(3.14b) $q_a = -\varkappa h_a{}^b(T_{;b} + T\dot{u}_b)$ | (3.14b') $q_\nu = -\varkappa T_{,\nu}\,,$

are necessary if the rate of generation of entropy is never negative; further the heat conduction coefficient $\varkappa(p, v)$ and the viscosity coefficient $\lambda(p, v)$ must obey the restrictions

$$\varkappa \geqslant 0\,, \qquad \lambda \geqslant 0\,.$$

[A bulk viscosity coefficient would give a contribution $-\eta\theta$ to the pressure, where $\eta\ (\geqslant 0)$ is the coefficient of bulk viscosity.]

A more detailed discussion of the relativistic thermodynamics leading to the identifications (3.14) may be found in Ehlers' review paper [22] and references cited there. We may note that the difference between the energy conservation eqs. (3.13), (3.13′) and the momentum conservation eqs. (3.4), (3.7b) are *special*-relativistic in origin, arising partly from the inertia assigned by special relativity to all forms of energy, and partly from the $3 + 1$ splitting of space-time given by h_{ab}.

(*) Volume per unit mass, per particle, or per baryon if J_a is defined as the mass, number or baryon-current density vector respectively.

3·4. *Perfect fluids.* – A *perfect fluid* is characterized by negligible heat conduction and viscosity:

$$\pi_{ab} = q_a = 0 \Leftrightarrow \varkappa = \lambda = 0 \; .$$

In this case

$$T_{ab} = \mu u_a u_b + p h_{ab} \, ,$$

and u_a is uniquely defined as the timelike eigenvector of the Ricci tensor.

The momentum-conservation eqs. (3.4), (3.7*b*) now show that the acceleration of the fluid is determined by the spatial pressure gradient:

(3.15) $$\dot{u}^a = -\frac{h_a{}^b p_{;b}}{\mu + p} \, .$$

(3.15′) $$a_\nu = -\frac{p_{;\nu}}{\varrho} \, .$$

Conditions (3.8*a*), (3.8′) ensure that these equations are determinate, and that the acceleration is always away from a high-pressure region towards a neighbouring low-pressure region. In the relativistic theory, the inertial-mass density of the fluid is $\mu + p$; a given spatial pressure gradient is therefore relatively less efficient in producing an acceleration in a general-relativistic fluid than in a corresponding Newtonian fluid, the inertial-mass density being enhanced by the contribution $p + \varepsilon\varrho$.

We may estimate that

(3.16*a*) $$\pi_{ab}|_0 \simeq 0 \simeq q_a|_0 \, , \qquad p|_0 \simeq 0$$

in any reasonable cosmological model, since the random velocities of galaxies are small at the present time (*). Then it is a plausible inference from the Copernican principle, but does *not* necessarily follow, that

(3.16*b*) $$\dot{u}_a|_0 \simeq 0 ;$$

(3.16*b*′) $$a_\nu|_0 \simeq 0 ;$$

i.e. that the acceleration of the velocity field representing the average motion of galaxies is negligible at the present time. We shall assume this is true.

For a perfect fluid, eq. (3.13) is:

$$\dot{S} = 0 \Leftrightarrow \text{entropy is constant along the flow lines of the fluid.}$$

Then (3.11), (3.12) show that there is only one independent thermodynamic

(*) Random galactic velocities of order 1000 km/s imply $p/\varrho c^2 \simeq 10^{-5}$. If there is a large flux of neutrinos or gravitational waves, these estimates could be seriously incorrect.

variable along each world line; by (3.10), this can be chosen to be ϱ. As the change of ϱ along a world-line is determined by (3.2), or equivalently by (3.3) we see that

(3.17) the change of thermodynamic state along the flow lines of a perfect fluid is determined (for a given equation of state) by the average length l alone.

In the general-relativity case, the original form (3.7*a*) of (3.13) is

$$\dot{\mu} + (\mu + p)\theta = 0 \,. \tag{3.18}$$

Restriction (3.8*a*) shows that compressing the fluid increases its energy density.

We may often assume an equation of state of the form $p = p(\mu)$. Given any such equation of state, restrictions (3.9*a*), (3.9*b*) are satisfied if $1 > \partial p/\partial\mu > 0$ (Curtis [25]; see also Harrison, Thorne, Wakano and Wheeler [26]), and (3.18) gives μ as a function of l along the world lines of the fluid.

3·5. *The matter and radiation content of the universe.* – To obtain an idea of the possible thermal histories of the universe, we represent the matter content of the universe by two main components.

3·5.1. Matter content. The *matter content* of the universe (galaxies and a possible intergalactic gas) can plausibly be represented by a perfect fluid (or mixture of perfect fluids with the same 4-velocity) with equations of state

$$p = \alpha\varrho^{\gamma} \,, \qquad T = \beta p/\varrho \,, \tag{3.19}$$

determining the pressure p and temperature T, where α, β, γ are constants. Then (3.2) shows that

$$p = \frac{\text{const}}{l^{3\gamma}} \,, \qquad T = \frac{\text{const}}{l^{3(\gamma-1)}} \,,$$

govern the change of temperature and pressure along the world lines of the fluid. For a monatomic gas, which is probably a good approximation at least at the present time, $\gamma = \frac{5}{3}$ and so

$$T_{\text{matter}} \propto \frac{1}{l^2} \tag{3.20}$$

gives the change of temperature as the volume changes.

For the relativistic case, the total energy density is the sum of the rest-mass energy and the internal energy:

$$\mu = \varrho + \frac{p}{\gamma - 1}$$

(TOOPER [27]), eqs. (3.2) and (3.18) being automatically consistent for this form of μ. The restrictions (3.9) are satisfied if $2 \geqslant \gamma > 1$. The change of μ along the world lines is clearly given by

$$\mu = \frac{(\varrho_0 l_0^3)}{l^3} + \left(\frac{P_0}{\gamma - 1}\right)\frac{l_0^{3\gamma}}{l^{3\gamma}}.$$

For recent times, we find that, to a good approximation (see (3.16) above) $p \simeq 0$ so that the energy-momentum tensor of matter in recent times is represented fairly well (for dynamical purposes) by regarding the matter as a *pressure-free fluid* (« *dust* »), with

$$p = 0 , \qquad \mu = \varrho , \qquad \dot{u}_a = 0$$

and

(3.21) $$\mu_{\text{matter}} \propto \frac{1}{l^3}.$$

(This is the approximation first introduced by EINSTEIN [3].)

3.5.2. Radiation content. The *radiation content* of the universe consists of many components, but the thermodynamically and dynamically dominant component appears to be the microwave radiation. We will accept this radiation as being black-body radiation at a temperature of about 3° K (this is its most probable interpretation; see Sciama's lectures for a discussion). We may represent this radiation as a perfect fluid with pressure $p = \frac{1}{3}\mu$ (*) and with the same 4-velocity as the matter. This is so at early stages of the universe's history because we would then expect (see next Section) the radiation to be in collisional equilibrium with the matter in the universe (the matter and radiation together can then be regarded as a 1-component fluid with $p \simeq \frac{1}{3}\mu$). At later stages the radiation propagates freely but may still be regarded as a perfect fluid moving with the matter 4-velocity because it is very nearly an isotropic radiation field with respect to that 4-velocity (cf. discussion in Sect. 7).

(*) The stress-tensor of any fluid consisting of particles with zero rest mass is $T_{ab} = \sum_{\text{particles}} p_* k_{*a} k_{*b}$ where k_{*a} is null for each particle, so $T^a{}_a = 0$; but from (3.5), this is the condition $p = \frac{1}{3}\mu$.

With the condition $p = \frac{1}{3}\mu$, (3.18) shows that

(3.22) $$\mu_{\text{radiation}} \propto \frac{1}{l^4}.$$

Defining the radiation temperature as usual by the condition $\mu = aT^4$ (a a constant), we see that the radiation temperature obeys the law

(3.23) $$T_{\text{radiation}} \propto \frac{1}{l}.$$

We note that the radiation obeys the perfect gas equations of state (3.19) with $\gamma = \frac{4}{3}$.

We have no information as to the isotropy of neutrinos of cosmological origin. These would be freely propagating at the present time. If they are now moving isotropically, they also obey (3.23). If they move anisotropically, then a fluid approximation is no longer adequate for the discussion (cf. MISNER [28]).

3.6. *The thermal history of the universe.* – So far, we have implicitly assumed there is no effective interaction between the two fluids (the matter and the radiation) in our cosmological model (or between different matter components). This is a good approximation at the present time, but will be untrue at certain earlier times. However we can obtain a good idea of the possible thermal histories of the universe from the information already at our disposal; despite the interactions, (3.17) is still approximately true (*).

Suppose that l decreases, as we go backwards in time, continuously to zero. Then the temperature of matter and radiation both rise indefinitely, according to (3.20) and (3.23). However when the temperature is above about 3000 °K ($l/l_0 \leqslant 1/1000$) the matter is ionized; Thomson scattering of the radiation and free electrons puts the electrons and radiation into close thermal contact, so the matter becomes an opaque plasma. The matter (the ions and electrons are in close contact because of Coulomb forces) and radiation must both have the same temperature for smaller values of l. Since the radiation (present density about 10^{-33} g/cm^3) has a much greater thermal capacity than the matter (present density between 10^{-29} and 10^{-31} g/cm^3), they both then obey the radiation temperature law (3.23).

As the temperature increases, successive interaction processes become important: nuclear reactions are important from 10^8 to 10^9 °K; electron-positron pair production and annihilation are important at $\sim 6 \cdot 10^9$ °K, with cor-

(*) Essentially because we still have an equation of state approximately of the form $p = p(\mu)$ for the mixture.

responding cooling and heating of the radiation gas; neutrino reactions become so rapid when $T \gtrsim 2 \cdot 10^{10}$ °K, that the plasma is then opaque to neutrinos. At these temperatures, the baryons, electron-positron pairs, radiation and neutrinos are in thermal equilibrium and form an ultra-relativistic perfect gas obeying (3.23). At about 10^{12} °K, large numbers of π- and μ-meson pairs are created and annihilated; known physics is a rather poor guide as to what happens at these and higher temperatures.

With this information, we can draw a diagram (Fig. 2) showing the possible thermal histories of the universe, by giving T as a function of the average radius l. Corresponding to each function $l(t)$, we obtain a thermal history from Fig. 2; for example, if (as we go back in time) l goes to a minimum and then increases again, T goes to a maximum and decreases. We can only find the

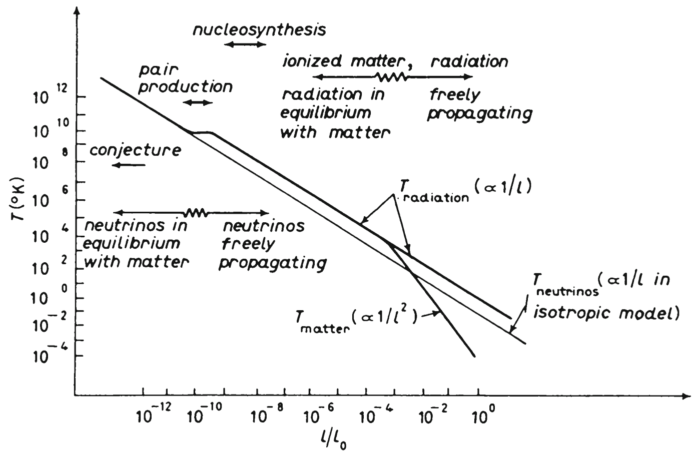


Fig. 2. – Diagram of temperature T as a function of average length l in a cosmological model. The curve for neutrinos is only valid for isotropic world models; the rest of the Figure will differ somewhat in anisotropic models, but remains essentially the same.

actual function $l(t)$ in any particular cosmological model, and so the time scales available for the various interactions and the actual thermal history in that model, from the field equations. The details of Fig. 2 depend on these time scales; if they were radically different from the time scales in the spherically symmetric models, the times available for equilibrium to be attained at each stage of expansion would be different, and various aspects of the diagram (such as the temperatures at which photons and neutrinos decouple from the matter) might change significantly.

The assumptions we have made in obtaining this diagram will be valid at nearly all times, and for any reasonable degree of inhomogeneity and anisotropy.

However (as has been pointed out by MISNER [28, 29]) there are at least two stages when irreversible processes certainly cannot be ignored: namely, when the neutrinos decouple and when the photons decouple. At these stages the matter behaves as a viscous fluid, and this might play an important role in the process of galaxy formation. The accompanying viscous heating arising (c.f. eq. (3.13)) from anisotropic fluid motion, would affect Fig. 2 at these times. (We note that the role of bulk viscosity has not yet been clarified. It may be important at some stages of the expansion of the universe, cf. a forthcoming paper by ANDERSON and STEWART.)

The detailed form of this diagram depends critically on astrophysical processes in the universe: in particular some arguments (see Sciama's lectures) suggest that an intergalactic gas might be reheated after its temperature had cooled well below the radiation temperature, at about the time that galaxy formation occurred in the universe. In this case the universe would again be filled by a ionised plasma (at a temperature of about 10^5 °K) at recent times.

The thermal history of the universe (derived for the isotropic case) is discussed in many articles, see for example DICKE, PEEBLES, ROLL and WILKINSON [30], PEEBLES [31], ALPHER, GAMOW and HERMANN [32], WAGONER, FOWLER and HOYLE [33] and HARRISON [34]. The role of irreversible processes in an anisotropic universe is discussed by MISNER [28, 29].

4. – The field equations.

4·1. *The curvature tensor.* – In general relativity, the curvature of space-time is represented by the Riemann curvature tensor R_{abcd}, which has the symmetry properties

$$R_{[ab][cd]} = R_{abcd} = R_{cdab}\,, \qquad R_{a[bcd]} = 0\,. \tag{4.1}$$

This tensor (which has 20 independent components) can be algebraically separated into the Ricci tensor R_{ab}, defined by

$$R_{ab} := R^c{}_{acb}\,, \tag{4.2}$$

and the Weyl tensor C_{abcd} (the « conformal curvature tensor ») defined by

$$C^{ab}{}_{cd} := R^{ab}{}_{cd} - 2g^{[a}{}_{[c}\,R^{b]}{}_{d]} + \frac{R}{3}\,g^{[a}{}_{[c}g^{b]}{}_{d]}\,, \tag{4.3}$$

where the Ricci scalar R is $R := R^a{}_a = R^{ab}{}_{ab}$. It follows from (4.3) that the Weyl tensor has all the Riemann tensor symmetries (4.1) and the additional property

$$C^{ab}{}_{ad} = 0\ .$$

Thus we may think of R_{ab} (with 10 independent components) as the trace of R_{abcd}, and of C_{abcd} (also with 10 independent components) as its trace-free part (*).

The quantity we may regard as corresponding to R_{abcd} in Newtonian theory is the second derivative $\Phi_{,\mu,\nu}$ of the gravitational potential Φ. This can be separated algebraically into its trace $\Phi^{,\nu}{}_{,\nu}$ and its trace-free part $E_{\mu\nu}$:

$$E_{\mu\nu} := \Phi_{,\mu,\nu} - \tfrac{1}{3} h_{\mu\nu} \Phi^{,\sigma}{}_{,\sigma}\ . \tag{4.4}$$

4.1.1. The role of the field equations. The field equations, including a cosmological constant Λ, are (on choosing units suitably),

the Einstein equations:	the Poisson equation:
(4.5) $(R_{ab} - \frac{1}{2} R g_{ab}) + \Lambda g_{ab} = T_{ab} \Leftrightarrow$ $\Leftrightarrow R_{ab} = (T_{ab} - \frac{1}{2} T g_{ab}) + \Lambda g_{ab}\ .$	(4.5′) $\Phi^{,\nu}{}_{,\nu} + \Lambda = \varrho/2 \Leftrightarrow$ $\Leftrightarrow \Phi^{,\nu}{}_{,\nu} = \varrho/2 - \Lambda\ .$

In each case, the field equations determine the « trace » part of the gravitational field algebraically at each point of space-time from the matter content at that point. The « trace-free » part is related to the matter content by differential equations (see (4.21), (4.22)) and is determined by these equations in conjunction with suitable boundary conditions, initial conditions, kinematic conditions or symmetry requirements; we shall generically call such restrictions « boundary conditions ».

In Newtonian theory, we have to drop the boundary condition

$$\Phi \to 0 \text{ at } \infty \tag{4.6}$$

if we are to get any reasonable cosmological models at all (see BONDI [1] and references given there; this follows from (4.5′) and is closely related to the difficulty (Subsect. 2.3) in separating a_ν invariantly into gravitational and inertial parts). In fact to get Newtonian analogues *a*) to many general-relativistic cosmological models, and *b*) to the general-relativistic expression for gravitational radiation reaction, we must also drop the condition

$$E_{\mu\nu} \to 0 \text{ at } \infty\ , \tag{4.7}$$

(*) For further details of this decomposition, see JORDAN, EHLERS and KUNDT [35].

which might seem to be a suitable generalization of (4.6). (For *a*) see HECKMANN and SCHÜCKING [36] and Sect. **5** below; for *b*) see the lectures by THORNE in this volume.)

4·2. *The Ricci identities for the velocity vector.* – In general relativity, arbitrary vector fields obey Ricci's identity (which we may regard as defining the curvature tensor). Applying this identity to the vector field u^a,

$$u_{a;d;c} - u_{a;c;d} = R_{abcd}\, u^b \,. \tag{4.8}$$

The corresponding Newtonian identities are

$$\partial(v_{\nu,\mu})/\partial t = (\partial v_\nu/\partial t)_{.\mu}\,, \tag{4.9a}$$

$$v_{\mu.\nu.\sigma} = v_{\mu.\sigma.\nu}\,. \tag{4.9b}$$

(In orthogonal curvilinear co-ordinates these are $\partial^2 v_\nu/\partial t\,\partial x^\mu = \partial^2 v_\nu/\partial x^\mu\,\partial t$, $\partial^2 v_\mu/\partial x^\nu\,\partial x^\sigma = \partial^2 v_\mu/\partial x^\sigma\,\partial x^\nu$).

Multiplying (4.8) by u^d in the general-relativity case, we get

$$(u_{a;c})^{\cdot} - \dot{u}_{a;c} + u_{a;d}\,u^d{}_{;c} + R_{abcd}\,u^b u^d = 0\,.$$

Projecting and remembering (2.7*b*), this equation is

$$h_a{}^c h_b{}^d (v_{cd})^{\cdot} - \dot{u}_a\,\dot{u}_b - h_a{}^c h_b{}^d\,\dot{u}_{c;d} + v_{ad}\,v^d{}_b + R_{acbd}\,u^c u^d = 0\,. \tag{4.10}$$

In Newtonian theory operating on (2.10*b*) by $_{.\nu}$ and using (4.9*a*), we find

$$\partial(v_{\mu.\nu})/\partial t + (v_{\mu.\nu})_{.\sigma}\,v^\sigma + v_{\mu.\sigma}\,v^\sigma{}_{.\nu} + \Phi_{.\mu.\nu} = a_{\mu.\nu};$$

this can be rewritten

$$(v_{\mu\nu})^{\cdot} - a_{\mu.\nu} + v_{\mu\sigma}\,v^\sigma{}_\nu + \Phi_{.\mu.\nu} = 0\,. \tag{4.11}$$

Equations (4.10), (4.11) are propagation equations for $v_{\mu\nu}$ along the flow lines of the fluid; the similarity of these equations is essentially the similarity of the « geodesic deviation equations » in Newtonian theory and in general relativity (see PIRANI [37], pp. 260-269, for a clear account of this correspondence).

4.2.1. **Raychaudhuri's equation.** Contracting eqs. (4.10), (4.11) we obtain propagation equations for θ.

$g^{ab}\times$(4.10), (3.5) and (4.5) implies	$h^{\mu\nu}\times$(4.11) and (4.5′) implies
(4.12) $\theta^{\bullet}+\frac{1}{3}\theta^2-\dot{u}^a{}_{;a}+2(\sigma^2-\omega^2)+ \frac{1}{2}(\mu+3p)-\Lambda=0$,	(4.12′) $\theta^{\bullet}+\frac{1}{3}\theta^2-a^{\nu}{}_{.\nu}+2(\sigma^2-\omega^2)+ \frac{1}{2}\varrho-\Lambda=0$,

which is Raychaudhuri's equation (Raychaudhuri [38, 39]). By (2.14), $\theta^{\bullet}+\frac{1}{3}\theta^2=3l^{\bullet\bullet}/l$, so we can rewrite this equation in the form

(4.13) $3l^{\bullet\bullet}/l=2(\omega^2-\sigma^2)+\dot{u}^a{}_{;a}- \frac{1}{2}(\mu+3p)+\Lambda$.	(4.13′) $3l^{\bullet\bullet}/l=2(\omega^2-\sigma^2)+a^{\nu}{}_{.\nu}- \frac{1}{2}\varrho+\Lambda$.

This shows how the second derivative of the curve $l(t)$ is determined directly at each space-time point by the matter density at that point, with the Λ-term acting as a constant repulsive force; rotation tends to hold the matter apart (as we might expect, representing a «centrifugal» effect); a pure distortion tends to pull the world-lines together; and acceleration affects the average distance of the world lines through its divergence.

The main difference between the Newtonian and general-relativistic cases lies in the fact that while the active gravitational mass density is ϱ in Newtonian mechanics, it is $\mu+3p=\varrho+\varepsilon\varrho+3p$ in general relativity. It is this additional pressure and internal energy contribution to the gravitational force which is the major cause of the problem of gravitational collapse in general relativity. Thus if we consider a static star model filled with a perfect fluid and take $\Lambda=0$, (4.12) becomes

$\dot{u}^a{}_{;a}=\frac{1}{2}(\mu+3p)$,	$a^{\nu}{}_{.\nu}=\frac{1}{2}\varrho$,

where the acceleration is determined from the pressure gradient by (3.15). The extra terms in the relativistic case show that the pressure which tries to balance the star through acceleration tends to defeat itself, since it contributes directly to the gravitational field which tends to cause the star to collapse. A contributing factor is the relative inefficiency of a given spatial pressure gradient in causing acceleration, in the relativistic case (see Subsect. 3.4).

We can evaluate Raychaudhuri's equation in a cosmological model at the present time t_0. Remembering the definitions of q and H (Subsect. 2.5.1) we find from (2.16) and (3.16) (we make the plausible inference $\dot{u}^a{}_{;a}|_0\simeq 0$; this is *not* a direct consequence of (3.16)) that (*)

$$q_0 \simeq \frac{\mu_0}{6H_0^2} - \frac{\Lambda}{3H_0^2}, \tag{4.13a}$$

with possible maximum errors of $-\frac{2}{3}$, $+\frac{2}{3}$ arising from the limits (2.16) on shear and vorticity (but the actual errors due to these terms are probably negligible, see Sect. 7).

If we believe that $\Lambda = 0$, the (4.13*a*) becomes

$$q_0 \simeq \frac{\mu_0}{6H_0^2}, \tag{4.13b}$$

showing directly the deceleration caused by the matter content of the universe. It is rather difficult to estimate q_0 observationally (cf. Burbidge's lecture). We have indications that $q_0 > 0$; more specifically, that $q_0 = 1 \pm \frac{1}{2}$. However we can only be reasonably certain (**) of fairly broad observational limits, say

$$|q_0| < 5 . \tag{4.14a}$$

On the other hand, we can find a lower limit for μ_0 on the basis of the observed matter in the universe; this lowest estimate is $\mu_0/6H_0^2 \simeq 10^{-2}$. There could exist intergalactic gas which is so far unobserved, giving values up to $\mu_0/6H_0^2 \simeq 1$ (cf. lectures by Sciama in this volume). Thus we can obtain plausible limits

$$10^{-2} \lesssim \frac{\mu_0}{6H_0^2} \lesssim 1 , \tag{4.14b}$$

on the density μ_0. However it is possible that the universe is filled with a large density of matter or energy, (such as neutrinos, (Ruderman [40]), gravitational waves, (Field, Rees and Sciama [41]), condensed stars (Zel'dovich [8]), young galaxies (Partridge and Peebles [42]), or rocks), which is very difficult to detect. The best definite limit we can place on the effective density of such forms of energy comes directly from (4.14*a*) and (4.13*b*), which place a limit $\mu_0 \lesssim 10^{-28}$ g/cm³ on the smoothed-out energy density of such « unobservable » forms of matter.

An alternative viewpoint is to use (4.13*a*) to give limits on Λ in terms of

---

(*) The density μ is often represented by the parameter $\sigma := \mu/6H^2$ (do not confuse with the shear!).

(**) q_0 as defined is obtained by taking a suitable *directional average*, as it is defined from *l*.

an equivalent mass density; we find

$$-250\mu_0 \leqslant \Lambda \leqslant 300\mu_0 \tag{4.14c}$$

if we believe (4.14*a*), (4.14*b*).

4.2.2. Further propagation equations. So far, we have extracted propagation equations for θ from (4.10), (4.11). We can also obtain propagation equations for the shear and vorticity from these relations.

The skew parts of (4.10), (4.11) are equivalent to

$$(4.15) \quad h^a{}_b(l^2\omega^b)^\cdot = \sigma^a{}_b(l^2\omega^b) + \frac{l^2}{2}\eta^{abcd}u_b\,\dot{u}_{c;d}\,.$$

$$(4.15') \quad (l^2\omega^\mu)^\cdot = \sigma^\mu{}_\nu(l^2\omega^\nu) + \frac{l^2}{2}\eta^{\mu\nu\sigma}a_{\nu,\sigma}\,.$$

Substituting into these equations for $\dot{u}_a$ from (3.15), we obtain the usual vorticity conservation laws when we have a perfect fluid with equation of state $p = p(\mu)$ (see Ehlers [22], Godel [43], Synge [44]).

The symmetric, trace-free parts of (4.10), (4.11) are (using (4.12))

$$(4.16) \quad h_a{}^f h_b{}^g(\sigma_{fg})^\cdot - h_a{}^f h_b{}^g\,\dot{u}_{(f;g)} - \dot{u}_a\dot{u}_b + \omega_a\omega_b + \sigma_{af}\sigma^f{}_b + \tfrac{2}{3}\theta\sigma_{ab} + h_{ab}(-\tfrac{1}{3}\omega^2 - \tfrac{2}{3}\sigma^2 + \tfrac{1}{3}\dot{u}^c{}_{;c}) - \tfrac{1}{2}\pi_{ab} + E_{ab} = 0\,,$$

where

$$E_{ac} := C_{abcd}u^b u^d;$$

the Weyl tensor symmetries imply

$$E_{ab} = E_{(ab)}, \quad E^a{}_a = 0\,, \quad E_{ab}u^b = 0\,.$$

$$(4.16') \quad \dot{\sigma}_{\mu\nu} - a_{(\mu,\nu)} + \omega_\mu\omega_\nu + \sigma_{\mu\varkappa}\sigma^\varkappa{}_\nu + \tfrac{2}{3}\theta\sigma_{\mu\nu} + h_{\mu\nu}(-\tfrac{1}{3}\omega^2 - \tfrac{2}{3}\sigma^2 + \tfrac{1}{3}a^\nu{}_{,\nu}) + E_{\mu\nu} = 0\,,$$

where

$$E_{\mu\nu} := \Phi_{,\mu,\nu} - \tfrac{1}{3}h_{\mu\nu}\Phi^{,\sigma}{}_{,\sigma}\,,$$

so

$$E_{\mu\nu} = E_{(\mu\nu)}\,, \quad E^\mu{}_\mu = 0\,.$$

Thus we see that while the expansion is affected directly by the matter at each point, the gravitational field E_{ab} (the « tidal force ») affects the fluid flow by inducing shear in the flow lines; this shear then determines the vorticity propagation and also enters into the expansion equation (tending to cause the world lines to converge). Note the terms in ω_a in (4.16), representing the distorting effect of « centrifugal forces ».

4.2.3. Constraint equations. Three further sets of equations can be obtained from (4.8) and (4.9*b*).

Multiply (4.8) by $g^{ac}h^{ed} \Rightarrow$

(4.17) $h^e{}_b(\omega^{bc}{}_{;c} - \sigma^{bc}{}_{;c} + \frac{2}{3}\theta^{;b}) + (\omega^e{}_b + \sigma^e{}_b)\,\dot{u}^b = q^e$.

Multiply (4.9b) by $h^{\nu\sigma} \Rightarrow$

(4.17′) $\omega^{\nu\mu}{}_{.\mu} - \sigma^{\nu\mu}{}_{.\mu} + \frac{2}{3}\theta^{.\nu} = 0$.

The vorticity vector satisfies a constraint equation:

by (4.8),

$$R^a{}_{[bcd]} = 0 \Rightarrow u_{[a;b;c]} = 0;$$

$\times\eta^{abcd}u_d$ shows

(4.18) $\omega^a{}_{;a} = 2\omega^b\dot{u}_b$.

(4.9b) implies

$$v_{[\mu.\nu.\sigma]} = 0;$$

$\times\eta^{\mu\nu\sigma}$ shows

(4.18′) $\omega^\nu{}_{.\nu} = 0$.

Finally,

(4.8) $\times\eta_{sfdc}u^f$ and symmetrization shows

(4.19) $H_{ad} = 2\dot{u}_{(a}\omega_{d)} - h_a{}^t h_d{}^s(\omega_{(t}{}^{b;c} + \sigma_{(t}{}^{b;c})\,\eta_{s)fbc}u^f$,

where

$$H_{ac} := \frac{1}{2}\eta_{ab}{}^{gh}C_{ghcd}u^b u^d .$$

The Weyl tensor symmetries show

$$H_{ab} = H_{ab},\quad H^a{}_a = 0\ ,\quad H_{ab}u^b = 0\ .$$

(4.9b) $\times\eta^{\varkappa\mu\sigma}$ and symmetrization shows

(4.19′) $(\omega_{(\varkappa}{}^{\nu.\tau} + \sigma_{(\varkappa}{}^{\nu.\tau})\eta_{\mu)\nu\tau} = 0$.

(H_{ab} has no analogue in Newtonian theory).

In the general-relativity case, eqs. (4.12) and (4.15)-(4.19) are equivalent to the Ricci identities (4.8). Through these equations we can completely determine the curvature tensor of space-time, since the Ricci tensor is determined directly by the matter content. and the Weyl tensor is determined by the symmetric, trace-free 3-tensors E_{ab}, H_{ab} (it is given by

$$C_{abcd} = (\eta_{abpq}\eta_{cdrs} + g_{abpq}g_{cdrs})u^p u^r E^{qs} - (\eta_{abpq}g_{cdrs} + g_{abpq}\eta_{cdrs})u^p u^r H^{qs},$$

where

$$g_{abcd} := g_{ac}g_{bd} - g_{ad}g_{bc}) .$$

In the Newtonian case, eqs. (4.12′) and (4.15′), (4.16′) are equivalent to (4.9*a*) and eqs. (4.17′)-(4.19′) are equivalent to (4.9*b*).

4·3. *The Bianchi identities.* – In general-relativity, the curvature tensor has to obey the *Bianchi identities*

$$R_{ab[cd;e]} = 0 \Leftrightarrow C^{abcd}{}_{;d} = R^{c[a;b]} - \tfrac{1}{6} g^{c[a} R^{;b]}; \tag{4.20}$$

these two forms are equivalent only because space-time is 4-dimensional (KUNDT and TRÜMPER [45]). These identities imply the contracted identities

$$R^{ab}{}_{;b} = \tfrac{1}{2} R^{;a} \Leftrightarrow G^{ab}{}_{;b} = 0$$

(which are the eq. (3.7)) and 16 further identities. Substituting into (4.20) for C_{abcd} in terms of E_{ab}, H_{ab} these equations take a form rather similar to Maxwell's equations. If the space-time is filled with a perfect fluid, they are (*)

$$:[\text{« div}\, E \text{ »}]: h^t{}_a E^{as}{}_{;d} h_s{}^d - \eta^{tbpq} u_b \sigma_p{}^d H_{qd} + 3H^t{}_s \omega^s = \tfrac{1}{3} h^t{}_b \mu^{;b}; \tag{4.21a}$$

$$\begin{aligned} :[\text{« } \dot H \text{ »}]: h^{ma} h^{tc} \dot H_{ac} - h_a{}^{(m}\eta^{t)rsd} u_r E^a{}_{s;d} + 2E_q{}^{(t}\eta^{m)bpq} u_b \dot u_p + h^{mt}(\sigma^{ab} H_{ab}) + \\ + \theta H^{mt} - 3H_s{}^{(m}\sigma^{t)s} - H_s{}^{(m}\omega^{t)s} = 0; \end{aligned} \tag{4.21b}$$

$$:[\text{« div}\, H \text{ »}]: h^t{}_a H^{as}{}_{;d} h_s{}^d + \eta^{tbpq} u_b \sigma_p{}^d E_{qd} - 3E^t{}_s \omega^s = (\mu + p)\omega^t; \tag{4.21c}$$

$$\begin{aligned} :[\text{« } \dot E \text{ »}]: h_a{}^m h_c{}^t \dot E^{ac} + h_a{}^{(m}\eta^{t)rsd} u_r H^a{}_{s;d} - 2H_q{}^{(t}\eta^{m)bpq} u_b \dot u_p + h^{mt}\sigma^{ab} E_{ab} + \\ + \theta E^{mt} - 3E_s{}^{(m}\sigma^{t)s} - E_s{}^{(m}\omega^{t)s} = -\tfrac{1}{2}(\mu + p)\sigma^{tm}. \end{aligned} \tag{4.21d}$$

Corresponding to the identities (4.21*a*), (4.21*b*), in Newtonian theory the tensor $E^{\mu\nu}$ satisfies the identities

$$E^{\mu\nu}{}_{,\nu} = \tfrac{1}{3} \varrho^{,\mu}, \tag{4.22a}$$

$$E^{(\mu}{}_{\sigma,\tau} \eta^{\nu)\sigma\tau} = 0, \tag{4.22b}$$

which follow from (4.9*a*). Further the shear tensor and expansion θ satisfy the 8 equations

$$\sigma_{[\nu}{}^{[\mu}{}_{,\tau]}{}^{,\kappa]} + \tfrac{2}{3} h^{[\mu}{}_{[\nu}\, \theta^{,\kappa]}{}_{,\tau]} = 0, \tag{4.22c}$$

(*) This form of the Bianchi identities is due to M. TRÜMPER. No difficulty arises in obtaining these equations for a general fluid; they are rather complex, and we give them in an Appendix. They hold for any energy-momentum tensor; the fluid approximation limits then only through prescribing the equations of state.

which follow from (4.9b) and are in fact completely analogous to (4.21c) and (4.21d). One could see this by substituting into (4.21c) and (4.21d) from (4.19), obtaining second-order identities satisfied by the kinematic quantities; however it is easier to proceed by noting that (4.22c) is the condition $\omega^{\mu}{}_{,\nu\sigma}\eta^{\nu\sigma\tau}=0$ in Newtonian theory (TRÜMPER [46]), while eqs. (4.21c) and (4.21d) are together equivalent to the corresponding general-relativity condition

$$h^a{}_s\,\omega^s{}_{;cd}\,\eta^{cdef}\,u_e = h^a{}_s\,R^s{}_{bcd}\,\omega^b\,\eta^{cdef}\,u_e\,.$$

4.4. *The field equations.*

4.4.1. The case of zero rotation. In this special case, we can obtain a useful alternative form of some of the general-relativity equations. Standard theorems show that

$\omega=0 \Leftrightarrow u_{[a}u_{b;c]}=0$	$\omega=0\Leftrightarrow v_{[\mu,\nu]}=0$
$\Leftrightarrow \exists$ locally functions	$\Leftrightarrow\exists$ locally a function
$f, g: u_a = f g_{;a}\,.$	$g: v_\mu = g_{,\mu}\,.$

Thus in each case, $\omega=0$ is the condition that there locally exist 3-surfaces in space-time (the surfaces $\{g=\text{const}\}$) orthogonal to the velocity vector field. However in general-relativity, u_a is the 4-velocity vector field; so the condition $\omega=0$ is precisely the condition that the instantaneous rest spaces defined at each point by h_{ab} should mesh together to form a set of 3-surfaces in space-time. These surfaces, which are surfaces of simultaneity for all the fluid observers, define a cosmic-time co-ordinate (the function g) determined by the fluid flow. (This time co-ordinate can be locally normalized to measure *proper* time along each world line only if $\dot u=0$.)

We can now use the Gauss-Codacci formulae (SCHOUTEN [4], JORDAN, EHLERS and KUNDT [35]) to relate the curvature tensor of space-time to the curvature of the 3-spaces orthogonal to u^a. If the Ricci tensor of these 3-spaces is $R^*{}_{ab}$, we find

$$R^*{}_{ab} = h_a{}^f h_b{}^g\left(-l^{-3}(l^3\sigma_{fg})^{\boldsymbol{\cdot}} + \dot u_{(f;g)}\right) + \dot u_a\dot u_b + \pi_{ab} + \frac{1}{3}h_{ab}\left(-\frac{2}{3}\theta^2 + 2\sigma^2 + 2\mu + 2\Lambda - \dot u^c{}_{;c}\right), \tag{4.23}$$

which implies that the Ricci scalar R^* of the 3-spaces is

$$R^* = -\frac{2}{3}\theta^2 + 2\sigma^2 + 2\mu + 2\Lambda\,; \tag{4.24}$$

these equations show how the curvature of the 3-spaces $\{g = \text{const}\}$ is determined by the matter content of the space-time through the field equations. They are essentially general-relativistic in character; R^*_{ab} has no Newtonian analogue. We can re-express (4.23) in the language of classical differential geometry (cf. EHLERS [22]): if, in a 3-space $\{g = \text{const}\}$, the Gaussian curvature of the 2-surface formed by all geodesics through the point x orthogonal to the direction e_a ($e^a e_a = 1$, $e^a u_a = 0$) at that point is $K(x, e^a)$, then

$$(4.25)\qquad K(x, e^a) = \{l^{-3}(l^3 \sigma_{fg})^{\cdot} - \dot{u}_{f;g} - \dot{u}_f \dot{u}_g - \pi_{fg}\} e^f e^g + \\ + \tfrac{1}{3}(\sigma^2 - \tfrac{1}{3}\theta^2 + \dot{u}^c{}_{;c} + \Lambda + \mu)\,.$$

If we separate R^*_{ab} into a trace-free part (which is essentially equivalent to E_{ab}) and its trace, then these quantities are related to each other by the Bianchi identities for the 3-spaces,

$$(4.26)\qquad h^d{}_a R^{*ab}{}_c h_b{}^c = \tfrac{1}{2} R^{*;c} h_c{}^d\,.$$

These identities are equivalent to the three 4-space Bianchi identities (4.21*a*), because of (4.23).

4.4.2. The field equations in general. We may regard eqs. (4.12), (4.16) and (4.17) as 9 of the 10 general-relativity field equations. The remaining field equation is one of the four first integrals which exist, when all the other conditions are satisfied, as a consequence of the contracted Bianchi identities (3.7) (*). This equation may be understood as giving a geometrical constraint, but no extra restrictions on the kinematic quantities, if all the other relations are satisfied. We may see this most easily in the case $\omega = 0$. Then the trace-free part of (4.23) is equivalent to (4.16), and the trace (eq. (4.24)) is the tenth field equation, relating R^* to the kinematic quantities and the density of matter. We only have to fulfil this equation at one point of space-time; then it is fulfilled in an open neighbourhood of that point (it is fulfilled in a 3-surface $\{g = \text{const}\}$ because of eqs. (4.21*a*) (which are equivalent to (4.26) if (4.23) is fulfilled); and it is fulfilled on timelike lines through this 3-surface because (3.7*a*) ensures that it is a first integral of the field equation). In effect, the other equations determine a solution of the field equations up to a

(*) Let $Z_{ab} := (R_{ab} - \tfrac{1}{2} R g_{ab}) + \Lambda g_{ab} - T_{ab}$; then the field equations are satisfied when $Z_{ab} = 0$. Now $Z^{ab}{}_{;b} = 0 \Leftrightarrow \partial Z_a{}^4/\partial x^4 = -Z_a{}^\nu{}_{;\nu} - \Gamma^4{}_{4d} Z_a{}^d + \Gamma_{a4}{}^d Z_d{}^4$ shows that if $Z_\mu{}^\nu = 0$ at all times and $Z_a{}^4 = 0$ on an initial surface $\{x^4 = \text{const}\}$, then $Z_a{}^4 = 0$ on an open neighbourhood of the surface; so the equations $Z_a{}^4 = 0$ are four first integrals of the other field equations. For further discussion see, for example, ANDERSON [47], CHOQUET-BRUHAT [48].

constant; eq. (4.24) determines this constant. The situation will be similar in the case $\omega \neq 0$.

In Newtonian theory, the only field equation is (4.12′), the remaining 8 equations which correspond to general-relativity field equations being simply kinematic identities.

We will not give a systematic discussion of the full set of equations here: rather, after briefly discussing the differences between the general-relativity and Newtonian equations, we shall illustrate their use by applying them to simple cosmological models.

4.4.3. Comparison of Newtonian and general-relativity equations. A comparison shows the first-order eqs. (4.12)-(4.19) satisfied by the kinematic quantities in general-relativity and in Newtonian theory are extremely similar, despite there being only one field equation in Newtonian theory and ten in general-relativity. Apart from the extra terms in the momentum equation mentioned in Sect. 3, there are some extra terms in the general-relativity equations which are essentially special-relativistic in origin; these are the term $\dot{u}_a \dot{u}_b$ in (4.16) and the terms in $\dot{u}_a$ in (4.17)-(4.19), resulting from the $3+1$ splitting of space-time given by h_{ab}. There are also extra terms which arise for essentially general-relativistic reasons. The general-relativity equivalents of some of the Newtonian equations contain terms representing the greater geometric freedom resulting from the possibility that space-time can be curved; these are the terms p in (4.12), π_{ab} in (4.16), q_a in (4.17) and H_{ab} in (4.19). However we have substituted for many of these terms (p, π_{ab} and q_a) through the field equations, and these are determined by the other kinematic quantities through equations of state (cf. Sect. 3). The Newtonian and general-relativity equations are most similar in the case of dust when these terms and $\dot{u}_a$ vanish.

The second-order identities (4.21), (4.22), although in 1-1 correspondence, look rather different; the extra terms in the general-relativity equations are so numerous that they dominate the equations. A further important difference arises from the different meanings E_{ab} has in the two theories: $\dot{E}_{ab}$ is locally prescribed by the general-relativity equations, but is left unprescribed in Newtonian theory. Thus, given suitable equations of state, the time development of the system off an initial surface is completely determined in the general-relativity case. For example if we have a perfect fluid with equation of state $p = p(\mu)$, $\dot{\mu}$ is given by (3.18), $p^{\bullet}$ follows from the equation of state, and the rate of change of $\dot{u}_a$ along the world-lines follows from that of p. The rates of change of $\theta, \sigma_{ab}, \omega_a$ are determined by (4.12), (4.16), (4.15); and the time derivatives of E_{ab}, H_{ab} are determined by (4.21b), (4.21d). However in the Newtonian case, the time development of the system is not determined until some suitable restriction has been put on $E_{\mu\nu}$; for example we can choose some particular world line and then prescribe $E_{\mu\nu}$ as an arbitrary function along that world line. Thus having dropped the boundary condition (4.7) we can use the resulting freedom of choice of $E_{\mu\nu}$ (this freedom is essentially a consequence of the infinite speed of propagation of gravitational effects) to influence the fluid flow in a rather arbitrary manner (cf. HECKMANN and SHÜCKING [36]). In particular, we can use this freedom in such a way as to produce Newtonian analogues of relativistic cosmological models. When we do so, we find that the general-relativity integrability conditions are more restrictive than the Newtonian conditions; as we shall see in the next Section, we seem to be able to find Newtonian analogues for each general-relativity solution, but the converse is not true.

5. – Applications of the field equations.

We shall assume in this Section, unless otherwise stated, that the matter and radiation content of the universe can be represented as a perfect fluid with equation of state $p = p(\mu)$. This will be a good approximation at most times (see Sect. 3).

5'1. *The Friedmann (or Robertson-Walker) models.* – These are those general-relativistic models which are locally istropic about every point of space-time. For example, exact isotropy of the number counts and radiation distribution would show $h_a{}^b\mu_{;b} = 0 \Leftrightarrow \dot{\mu}u_a = -\mu_{,a}$; then the equation of state implies $h_a{}^b p_{;b} = 0$. The surfaces $\{\mu = \text{const}\}$ (*) are therefore orthogonal to the fluid vector u_a in this case, which implies $\omega_a = 0$; and the momentum conservation equation (3.15) imply $\dot{u}_a = 0$. An isotropic Hubble law implies $\sigma_{ab} = 0$ (and in fact also implies $\dot{u}_a = 0$; see Sect. 6). With these conditions, (4.17) shows $h_a{}^b\theta_{;b} = 0$. Thus these models can be characterized by the conditions

$$\omega_a = \sigma_{ab} = 0 = \dot{u}_a \,, \tag{5.1a}$$

which imply the further conditions

$$\mu = \mu(t) \,, \qquad p = p(t) \,, \qquad \theta = \theta(t) \,, \tag{5.1b}$$

where the surfaces $\{t = \text{const}\}$ are the surfaces orthogonal to the fluid flow vector. We may choose the co-ordinate t to measure proper time along each world line, and the scale factor l to be a function of t alone; then $l(t)$ is pre-

(*) $\theta \neq 0$ shows $\dot{\mu} \neq 0$ by (3.8*a*), (3.18), so these *are* well-defined surfaces when this condition holds.

cisely the « radius function » $R(t)$ commonly used in describing the Robertson-Walker space-times.

The Newtonian analogues of these models may be defined by (5.1a) which again implies (5.1b). The remaining nontrivial equations are

(5.2a)	$\dot{\mu} + 3(\mu + p)l^{\bullet}/l = 0$,	(5.2a')	$\dot{\varrho} + 3\varrho l^{\bullet}/l = 0$,
(5.2b)	$3l^{\bullet\bullet}/l + \frac{1}{2}(\mu + 3p) - \Lambda = 0$,	(5.2b')	$3l^{\bullet\bullet}/l + \frac{1}{2}\varrho - \Lambda = 0$,
(5.2c)	$E_{ab} = H_{ab} = 0$.	(5.2c')	$E_{\mu\nu} = 0$.

In the general-relativity case, eq. (4.25) reduces to

$$K(x, e^a) = \tfrac{1}{3}(\mu + \Lambda - \tfrac{1}{3}\theta^2) = K(t)\,, \tag{5.2d}$$

which shows that the 3-surfaces orthogonal to u^a are isotropic at each point, and so are 3-spaces of constant curvature $K(t)$. (We may note that (5.2c) is the statement that space-time is conformally flat (the Weyl tensor vanishes). There is in fact a converse to this (due to TRÜMPER):

$$\{E_{ab} = H_{ab} = 0\} \Rightarrow (5.1) \text{ holds}\,,$$

as the Bianchi identities (4.21) show that when $E_{ab} = H_{ab} = 0$, then $h_a{}^b\mu_{,b} = \omega_a = \sigma_{ab} = 0$. Since $p = p(\mu)$ (*),

$$\{h_a{}^b\mu_{,b} = 0\} \Rightarrow \{h_a{}^b p_{,b} = 0\} \Rightarrow \dot{u}_a = 0\,.)$$

Equation (5.2a) shows that

(5.3)	$ll^{\bullet}(\mu + 3p) = -(\mu l^2)^{\bullet}$.	(5.3$'$)	$\varrho ll^{\bullet} = -(\varrho l^2)^{\bullet}$.

Thus when $l^{\bullet} \neq 0$, we can integrate Raychaudhuri's equation (5.2b) to obtain

(5.4a)	$3l^{\bullet 2} - (\mu l^3)/l - \Lambda l^2 = 10E$, $E = \text{const}$	(5.4a')	$3l^{\bullet 2} - (\varrho l^3)/l - \Lambda l^2 = 10E$, $E = \text{const}$

which is the Friedmann equation; it has the form of an energy equation (the term μl^3 is constant if we have a dust-filled universe, and ϱl^3 is constant in the

(*) If we drop this restriction, we also obtain a family of inhomogeneous spaces (SHEPLEY and TAUB [49]).

Newtonian case). We may note that if there are several noninteracting components of matter, (5.2) implies (5.3) independently for each such component; so we may add together the densities μ_i of such components in (5.4), or use these equations with μ representing the total energy density.

In the general-relativity case, (5.2*d*) is equivalent to (5.4*a*) if $K(t)$ is related to E by

(5.4*b*) $$K(t) = -\frac{10E}{3l^2} = \frac{k}{l^2},$$

where for convenience we have introduced the constant $k = -10E/3$; on multiplying $l(t)$ by a suitable number, we can normalize k to one of the values $+1, 0$ or -1.

Evaluating (5.2*d*) at the present time shows that

(5.5*a*) $$K(t_0) = \frac{k}{l_0^2} = H_0^2\left(\frac{\mu_0}{3H_0^2} + \frac{\Lambda}{3H_0^2} - 1\right);$$

this equation can be combined with (4.13*a*) (which is now an exact relation if $p_0 = 0$) to show

(5.5*b*) $$K(t_0) = \frac{k}{l_0^2} = H_0^2(2q_0 - 1) + \Lambda .$$

If the pressure vanishes at all times, we can rewrite the Friedmann equation in the form

(5.4*c*) $$H^2 := \left(\frac{l^\cdot}{l}\right)^2 = \left(\frac{l_0}{l}\right)^3 (2q_0 H_0^2 + \tfrac{2}{3}\Lambda) + \frac{\Lambda}{3} - \left(\frac{l_0}{l}\right)^2 K(t_0) .$$

Given suitable equations of state, we can find $\mu(l)$ (see Sect. 3) and then Friedmann's eq. (5.4) gives the rate of expansion as a function of average length l. We can further integrate the equation $\theta(l) + 3l^\cdot/l$ analytically or numerically to obtain $l(t)$. These equations have been studied extensively in the literature (see, for example, Robertson [50, 51], Bondi [1], Stabell and Refsdal [52], Refsdal, Stabell and de Lange [53], Tauber [54], Rindler [5]) so we shall simply give two comments on their solutions.

5·1.1. Expansion from a singularity. Raychaudhuri's equation (5.2*b*) shows that if $\Lambda \leq 0$ and the energy condition (3.8) holds then, irrespective of the equation of state of matter, $l^{\cdot\cdot} < 0$. This implies $q_0 > 0$. In fact, even if $\Lambda > 0$, the condition $q_0 > 0$ is sufficient to imply $l^{\cdot\cdot} < 0$ at all earlier times (Λ is constant so (3.8) and (3.18) guarantee that $\mu/2$ remains larger than Λ). Therefore either $\Lambda \leq 0$ or $q_0 > 0$ imply there was a singularity in the universe when $l \to 0$ a finite time t_0 ago, where (*) (cf. Fig. 3)

$$t_0 < \frac{1}{H_0} \simeq 1.3 \cdot 10^{10}\ \text{y}\,. \tag{5.6}$$

[In the general-relativity case, the larger the pressure exerted by matter and radiation, the smaller t_0 is.] The conservation equations show that the temperature and density of matter become infinitely large near the singularity.

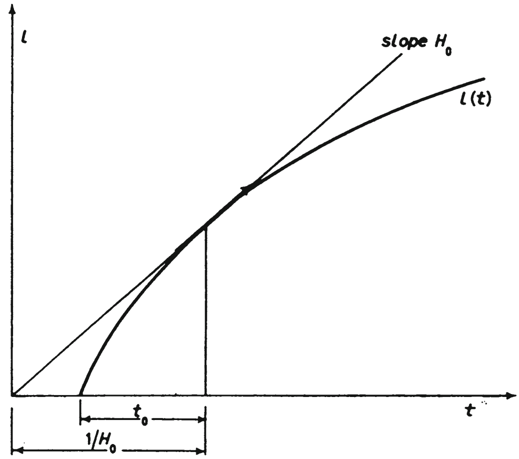


Fig. 3. – If $q_0 > 0$ and the energy conditions $\mu + p > 0$, $\mu + 3p > 0$ are always fulfilled, then the age t_0 of an isotropic universe is less than $1/H_0$.

Raychaudhuri's equation shows similarly that if $\Lambda \leqslant 0$ or $q_0 > 0$, a Robertson-Walker universe evolves very rapidly through its high-temperature phases, most of its lifetime up to the present occurring when $T_{\text{rad}} < 300$ °K. For example, if decoupling takes place a time t_d after the singularity, then $l_d/l_0 \simeq 1/1000$ so $t_d < 10^{-3} H_0^{-1} \simeq 10^7$ y.

5·1.2. Qualitative properties of the solutions. If $\Lambda = 0$, there are (again irrespective of the equation of state as long as $\mu > 0$, $p \geqslant 0$) three possible kinds of solution. These are illustrated in Fig. 4 *a*). The solution either (if $k > 0$) collapses back to a second singularity, or (if $k = 0$) has just sufficient energy to escape such a collapse (so $l \to \infty$ as $t \to \infty$), or (if $k < 0$)

(*) This inequality can be sharpened and used to show, on comparing t_0 with the age of the galaxy, that $q_0 < 5.0$; see Rindler [55].

easily escapes. The second case is the generalized Einstein-de Sitter case, for which $\mu = \frac{1}{3}\theta^2$ at all times. We may use (5.5*a*) to see that the values $\mu_0 > 3H_0^2$, $\mu_0 = 3H_0^2$ and $\mu_0 < 3H_0^2$ correspond to these three cases respectively; and (5.5*b*) shows that, up to a correction term involving p_0, they correspond to $q_0 > \frac{1}{2}$,

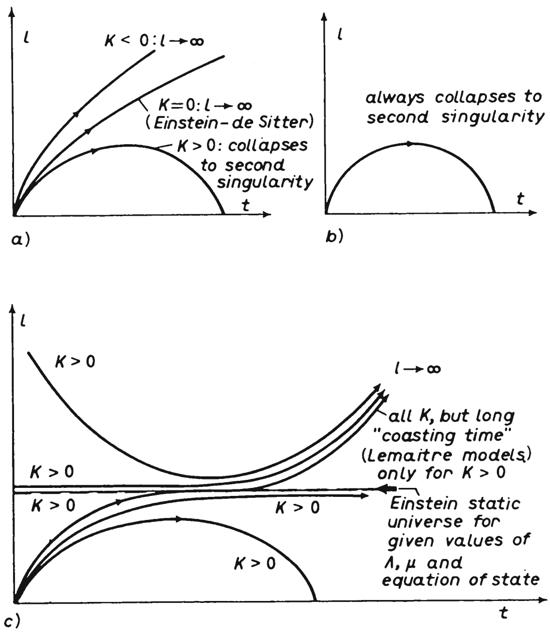


Fig. 4. – The possible characteristics of the function $l(t)$ in Robertson-Walker universes in which $\mu + p > 0$, $\mu + 3p > 0$ at all times. *a*) when $\Lambda = 0$; *b*) when $\Lambda < 0$; *c*) when $\Lambda > 0$. The time reverses of these solution are also solutions.

$q_0 = \frac{1}{2}$ and $q_0 < \frac{1}{2}$. The solutions for which $l \to \infty$ as $t \to \infty$ are asymptotic for large t to solutions with $p = 0$, $\mu = 0$ and $l = H_0 t$ (the *Milne* universe) if $k = -1$ and to the *Einstein-de Sitter* universe ($p = 0, l \propto t^{2/3}$) if $k = 0$.

If $\Lambda < 0$ the solution must collapse back to a second singularity.

If $\Lambda > 0$ there exists an unstable solution with $\theta = 0$, *i.e.* $l = \text{const}$; this in the *Einstein static solution* (Einstein [3]). In this case,

$$(5.7) \quad \begin{cases} \Lambda = \frac{1}{2}(\mu + 3p) > 0\,, \\ K = \frac{1}{2}\mu + p = \text{const} > 0\,. \end{cases} \qquad (5.7') \quad \begin{cases} \varrho = \text{const}\,, \\ \Lambda = \frac{1}{2}\varrho > 0\,. \end{cases}$$

Thus as well as the possibilities arising when $\Lambda = 0$, we may further construct solutions asymptotic, or nearly asymptotic, to the Einstein static universe, obtaining the possibilities shown in Fig. 4 c). Those solutions which expand forever are asymptotic for large l to solutions with $p = 0$, $\mu = 0$ and $l = \exp[H_0 t]$ where $H_0 = \frac{1}{3}\theta = \sqrt{\Lambda}/3$, a constant. The exact general-relativity solution of this kind has $k = 0$, and is the *de Sitter* universe; it is the same space-time which, with different field equations, constitutes the steady-state universe.

We see that solutions in which there is no singularity, and so in which there is a maximum temperature during the evolution of the universe, can only occur for large positive Λ.

5˙1.3. Co-ordinates. It can be shown from (5.1), (5.2*d*) that there exist co-ordinates such that

the metric takes the form

$$ds^2 = -dt^2 + R^2(t)\, d\sigma^2$$

where $d\sigma^2$ is the metric of a 3-space of constant curvature $k = +1$, 0 or -1 (see, for example, HECKMANN and SCHÜCKING [6], ANDERSON [47]),

the co-ordinates of any given fluid particle are

$$x^\nu = R(t)\, c^\nu\,, \qquad c^\nu = \text{const}\,,$$

and the gravitational potential is

$$\Phi = \frac{1}{6}\left(\frac{\varrho(t)}{2} - \Lambda\right) h_{\mu\nu} x^\mu x^\nu$$

(see, for example, BONDI [1]),

where we have written $R(t) \equiv l(t)$ to conform with the notation commonly used in these models. In the general-relativity case, the co-ordinates are co-moving co-ordinates; they can be chosen so that

$$d\sigma^2 = dr^2 + f^2(r)(d\theta^2 + \sin^2\theta\, d\varphi^2)\,,$$

where

$$f(r) = \begin{cases} \sin\ r \\ r \\ \sinh r \end{cases} \qquad \text{if } k = \begin{cases} +1 \\ 0 \\ -1 \end{cases};$$

(these are the co-ordinates we shall use in Sect. 6).

5˙2. *Gödel's solution.* – This solution is characterised by

$$\theta = \sigma_{ab} = \dot{u}_a = 0\,, \quad \omega \neq 0\,, \quad \omega_{a;b} = 0 \quad (\Rightarrow \omega = \text{const})\,.$$

The field equations and 1st-order identities take the form

General relativity	Newtonian theory
(5.8*a*) $2\omega^2 + \Lambda = \frac{1}{2}(\mu + 3p)$,	(5.8*a*′) $2\omega^2 + \Lambda = \frac{1}{2}\varrho$,
(5.8*b*) $E_{ab} = \frac{1}{3}h_{ab}\omega^2 - \omega_a\omega_b$,	(5.8*b*′) $E_{\mu\nu} = \frac{1}{3}h_{\mu\nu}\omega^2 - \omega_\mu\omega_\nu$,
(5.8*c*) $H_{ab} = 0$,	

which imply

$h_d{}^a h_e{}^b E_{ab;c} = 0$, $\mu = \text{const}$, $p = \text{const}$.	$E_{\mu\nu,\sigma} = 0$, $\varrho = \text{const}$.

The second-order identities are

(5.9*a*) $-3E^t{}_s\omega^s = (\mu + p)\omega^t$, (5.9*b*) $-E_s{}^{(m}\omega^{t)s} = 0$.	trivially satisfied.

In the general-relativity case, substituting from (5.8) we find that (5.9*b*) is identically satisfied, but (5.9*a*) gives a new condition:

$$\mu + p = 2\omega^2 . \tag{5.10a}$$

Combining this with (5.8*a*) shows that

$$\tfrac{1}{2}(\mu - p) = -\Lambda . \tag{5.10b}$$

We note that $\Lambda < 0$ for most reasonable equations of state; this is the opposite sign to that in the Einstein static universe.

Given a suitable equation of state $p = p(\mu)$, in general-relativity only one of μ, ω, Λ will determine the other two through eqs. (5.10). (In particular, if $p = 0$ we find $\mu = 2\omega^2 = -2\Lambda$.) In Newtonian theory, we have only to satisfy the restriction (5.8*a*′); two of ϱ, ω, Λ are arbitrary, and Λ can be positive, zero or negative. Thus the family of general-relativity solutions is more restricted than the family of Newtonian solutions.

Further details of these solutions may be found in papers by

GÖDEL [56].	HECKMANN and SCHÜCKING [13].

5.3. *Further solutions.* – Gödel's solution is not a realistic model universe, as the matter in this solution does not expand. We wish to find solutions with $\theta > 0$ which are more complex than the Robertson-Walker solutions. The problem in integrating Raychandhuri's equation to obtain equations like the Friedmann equation is that we require some restriction on the Weyl tensor enabling us to find σ^2 as a function of l. We might try

5.3.1. Solutions with $\sigma = \dot{u} = 0$, $\omega\theta \neq 0$. In this case the vorticity eq. (4.15) shows $(l^2\omega^b)^{\bullet} = 0$ which implies $\omega^2 = \Omega^2/l^4$, $\dot{\Omega} = 0$. Then we can integrate Raychaudhuri's equation to obtain the generalized Friedmann equation

$$3(l^{\bullet})^2 + \frac{2\Omega^2}{l^2} - \frac{(\mu l^3)}{l} - \Lambda l^2 = 10E\,, \qquad E^{\bullet} = 0\,.$$

However it can be shown, using the integrability condition, that

there exist no such general-relativity solutions if $\partial p/\partial\mu \neq 0$ (*) or if $p = 0$ (Ellis [57]).

there exist many such Newtonian solutions (Heckmann and Schücking [13]) which are necessarily homogeneous (Trümper [46]).

Another simple case is.

5.3.2. Solutions with $\omega = \dot{u} = 0$, $\sigma\theta \neq 0$, $K(x, e^a)$ isotropic. In this case the trace-free part of (4.23) (or (4.25)) shows $(l^3\sigma_{ab})^{\bullet} = 0$ which implies $\sigma^2 = \Sigma^2/l^6$, $\dot{\Sigma} = 0$. We can integrate Raychaudhuri's equation to obtain the generalized Friedmann equation.

$$3(l^{\bullet})^2 - \frac{\Sigma^2}{l^4} - \frac{(\mu l^3)}{l} - \Lambda\, l^2 = 10E, \qquad \dot{E} = 0\,.$$

The qualitative behaviour of $l(t)$ is the same as in the Robertson-Walker case, but the expansion time scales associated with these anisotropic solutions are less than those in isotropic solutions.

There are many such solutions in general relativity (see for example Ellis and McCallum [58], Ellis [57]) and in Newtonian theory. For example, the popular Bianchi I space-times with a metric of the form

$$ds^2 = -\,dt^2 + X^2(t)\,dx^2 + Y^2(t)\,dy^2 + Z^2(t)\,dz^2$$

in co-moving co-ordinates (see, for example, Heckmann and Schücking [7], Thorne [59], Jacobs [60, 61]), belong to this family. These spatially homogeneous models have space sections of constant curvature $K = 0$; they are

(*) Then the surfaces $p = \text{const}$ are hypersurfaces orthogonal to the fluid flow vector.

among the simplest homogeneous, anisotropic general-relativity cosmological models. However $E_{ab} \neq 0$ in them so their (spatially homogeneous) Newtonian analogues also have $E_{\mu\nu} \neq 0$.

5·3.3. Perturbations. An important application of the set of equations we have obtained is in the study in a very clear manner of *perturbations of isotropic world models* (HAWKING [62]; but note that there are some misprints in this paper, and that the expressions given for density perturbations with $p = \frac{1}{3}\mu$ are only valid for large wavelengths).

All perturbation studies agree that unless there is a very large positive Λ-term (as in extreme Lemaitre models), there has not been time for galaxies to form from thermal fluctuations in a Robertson-Walker universe (see the seminars by SALPETER and REES).

5·3.4. Viscous effects. Finally, following MISNER [28], we may consider the effect of *viscosity* on the evolution of cosmological models. Consider the spaces of Sect. 5·3.2, but now with a viscosity term present. The trace-free part of (4.23) is (because of (3.14))

$$(l^3 \sigma_{fg})^{\cdot} = l^3 \pi_{fg} = - l^3 \lambda \sigma_{fg} \,.$$

If we assume λ is approximately constant during the time of interest, we can integrate to obtain

$$\sigma^2 = \frac{\Sigma^2 \exp[-2\lambda t]}{l^6}\,,$$

showing that viscosity causes an exponential decay in the shear. This is a strictly general-relativistic effect (the term π_{ab} occurs in (4.16) but not in (4.16′)). Further, terms $-\frac{1}{2}\lambda E_{ab}$ and $-\frac{1}{2}\lambda H_{ab}$ now appear on the right-hand sides of (4.21*d*), (4.21*b*) respectively, showing that viscosity will tend to make the free gravitational field (represented by the Weyl tensor) die away (HAWKING [62]).

This and similar calculations lead us to hope (MISNER [28, 29]; cf. STEWART [23]) that almost any universe model will turn out, after sufficiently realistic physical processes have been considered, to evolve into a state very like a Robertson-Walker universe.

5·3.5. Further solutions. Much of the work in this and the last chapter is based on Ehler's very clear review article [22]. Further applications of the Bianchi identities to study the dynamics of fluids in general relativity may be found in KUNDT and TRÜMPER [45], SZEKERES [63], SHEPLEY and TAUB [49]. Solutions with $H_{ab} = 0$ are discussed by TRÜMPER [64]. Discussion and applications of the Newtonian equations in the cosmological context may be found in HECKMANN and SCHÜCKING [13, 36], HECKMANN [65], RAYCHAUDHURI [39], TRAUTMANN [14], TRÜMPER [46]. Solutions with $E_{\mu\nu} = 0$ are discussed by NARLIKAR [66].

Besides those mentioned in this chapter, many other exact general-relativity solutions have been found; a classification of these space-times according to their symmetries, and further references, may be found in ELLIS [57], STEWART and ELLIS [67], ELLIS and MACCALLUM [58].

6. – Observations in cosmological models.

In this Section we discuss observations in a general curved space-time. Although the sources and observers move with a unique velocity at each space-time point in the applications to cosmological models we have in mind, many of the relations in 6·1-6·4 and 6·6 are valid for sources and observers moving with arbitrary 4-velocities at arbitrary points.

6·1. *The geometric optics approximation.* – The radiation which conveys information in a cosmological model may be represented by a geometric optics solution (*) of Maxwell's equations. The electromagnetic field F_{ab} is regarded as a test field (*i.e.* we can neglect its effect on the curvature of space-time) in a charge- and current-free space-time, and so obeys Maxwell's source-free equations

$$F_{[ab;c]} = 0 \Leftrightarrow \exists \Phi_a : F_{ab} = \Phi_{b;a} - \Phi_{a;b} \,, \tag{6.1a}$$

$$F^{ab}{}_{;b} = 0 \,. \tag{6.1b}$$

The potential Φ_a will be chosen to obey the *gauge* condition

$$\Phi^a{}_{;a} = 0 \,. \tag{6.1c}$$

We assume there exist solutions of these equations of the form

$$\Phi_a = g(\varphi) A_a + \text{small tail terms} \,, \tag{6.2}$$

where *a*) $g(\varphi)$ is an arbitrary function of the phase φ, and *b*) g varies rapidly compared with the amplitude A_a in the sense that

$$g' \, k_{[a} A_{b]} \gg g A_{[a;b]} \,, \tag{6.3}$$

(*) For further discussion of this approximation, see EHLERS [68].

where $g' := \partial g/\partial\varphi$ and we have defined the propagation vector k_a by

$$k_a := \varphi_{;a} \,. \tag{6.4}$$

(*a*) is the condition that arbitrary information can be propagated by the signal (cf. TRAUTMANN [69]), and *b*) is the condition that the signal represents a high-frequency wave with a relatively slowly varying amplitude.) Substituting (6.2) into (6.1*b*), ignoring the tail terms, and equating to zero separately the coefficients of g, $g' = \partial g/\partial\varphi$ and $g'' = \partial^2 g/\partial\varphi^2$ (which we may do as g is arbitrary), we find

$$k^a k_a = 0 \,, \tag{6.5a}$$

$$A_{a;b} k^b = -\tfrac{1}{2} A_a k^b{}_{;b} \,, \tag{6.5b}$$

$$(A^a)^{;b}{}_{;b} + R^a{}_b A^b = 0 \,. \tag{6.5c}$$

The third equation will play no further part in the present discussion; its essential effect is the show that we cannot in general omit the tail terms if (6.2) is to be an exact solution of (6.1).

From (6.1*a*) and (6.3) we find that the electromagnetic field has the approximate form

$$F_{ab} \simeq g'(k_a A_b - A_a k_b) \tag{6.6}$$

and the electromagnetic stress tensor S_{ab}, defined by

$$S_{ab} := F_{ac} F_b{}^c - \tfrac{1}{4} g_{ab} F^{cd} F_{cd} \,,$$

has the form (*)

$$S_{ab} \simeq A^2 (g')^2 k_a k_b \,, \tag{6.7}$$

where we have defined $A^2 := A^a A_a$.

An observer with 4-velocity u^a finds the radiation flux across a surface perpendicular to k^a to be the same as the instantaneous energy density of the radiation, both being equal to

$$\mathfrak{f} = S_{ab} u^a u^b = A^2 (g')^2 (k_a u^a)^2 \,. \tag{6.8}$$

Equation (6.5*a*) implies $k^a k_{a;b} = 0$. However (6.4) shows $k_{a;b} = k_{b;a}$, so

(*) This is the stress-tensor of a particle (the photon) moving with 4-velocity k^a. The gauge condition (6.1*c*) implies $k^a A_a = 0$ and so ensures A_a is spacelike.

we find

$$k_{a;b}k^b = 0 \,. \tag{6.9}$$

Thus the light rays (the curves whose tangent vector field is k^a) are ***null geodesics.*** It follows that light rays are bent by an anisotropic gravitational field. Since this deflection need not be the same for every ray in a small bundle of light rays, such a tube may be differentially bent. Thus a curved space-time will in general distort optical images (SACHS [70]).

6·2. *Red-shifts.* – The rate-of-change of $g(\varphi)$ measured by an observer moving with 4-velocity u^a is $g_{;a}u^a = g'(k_a u^a)$. If observers with 4-velocities u_1^a, u_2^a measure the rate of change of the same signal $g(\varphi)$, these rates of change are in the ratio $(k_a u^a)_1/(k_b u^b)_2$. We can think of this as a *time-dilatation* effect: if a proper time interval $\mathrm{d}t$ is observed to elapse between particular signals (such as pulses emitted at unit time intervals by one of the observers), then $\mathrm{d}t_2/\mathrm{d}t_1 = (k_a u^a)_1/(k_b u^b)_2$. In particular, the observed frequencies ν of light or radio waves are related by

$$\frac{\nu_1}{\nu_2} = \frac{(k_a u^a)_1}{(k_b u^b)_2}\,.$$

The *red-shift* z of a source as measured by an observer is defined in terms of wavelengths by

$$z := \frac{\lambda_{\text{observed}} - \lambda_{\text{emitted}}}{\lambda_{\text{emitted}}} =: \frac{\Delta\lambda}{\lambda_{\text{emitted}}}\,. \tag{6.10a}$$

We therefore find that

$$1 + z = \frac{\lambda_{\text{observed}}}{\lambda_{\text{emitted}}} = \frac{\nu_{\text{emitted}}}{\nu_{\text{observed}}}\,,$$

so

$$1 + z = \frac{(u^a k_a)_{\text{emitter}}}{(u^b k_b)_{\text{observer}}} \tag{6.10b}$$

determines the red-shift from the 4-velocity vectors $u^a|_{\text{observer}}$, $u^a|_{\text{emitter}}$ and from the tangent vector k^a to the null geodesic. This relation is true no matter what the separation of emitter and observer, and holds independent of an interpretation of the red-shift as a « Doppler » or « gravitational » red-shift. as an example, if both observer and emitter are at the « same » point (*) and

(*) In the cosmological context, at distances up to, say, 10 Mpc.

the emitter moves radially away from the observer, we can write

$$u_a = \cosh\beta u_a + \sinh\beta e_a\,, \qquad e_a e^a = 1\,, \qquad e^a u_a = 0\,, \tag{6.11a}$$

as the 4-velocity of the emitter, where u^a is the observer's 4-velocity, e_a is the direction of motion of the emitter in the observer's rest frame, and $V = \operatorname{tgh}\beta$ is the velocity of relative motion. Then a null ray representing a signal from the emitter to the observer is $k^a = k(u^a - e^a)$; from (6.10*b*) we immediately find

$$1 + z = \exp[-\beta] = \sqrt{\frac{1+V}{1-V}}\,, \tag{6.11b}$$

the standard result for the red-shift due to radial motion of the source relative to the observer when both are at the same space-time point.

For later use, we introduce the following decomposition of k^a. Consider an observer with 4-velocity u^a and let v be an affine parameter along the null geodesic with tangent vector k^a; so $k^a = \mathrm{d}x^a/\mathrm{d}v$. If n^a is the unit vector in the direction of the projection of k^a into the rest space of the observer, then

$$k^a = (-u_b k^b)(u^a + n^a)\,, \qquad n^a n_a = 1\,, \qquad n^a u_a = 0\,. \tag{6.12}$$

Thus a small increment $\mathrm{d}v$ in the affine parameter will be considered by the observer to correspond to a time difference $\mathrm{d}t$ and a spatial displacement $\mathrm{d}l$, with

$$|\mathrm{d}t| = |\mathrm{d}l| = (-k^a u_a)\,\mathrm{d}v\,. \tag{6.13}$$

6'2.1. The linear red-shift relation in a cosmological model. In a given cosmological model, the emitter and observer coincide with particular galaxies moving with the unique fluid velocity u^a. The change in $(u^a k_a)$ occurring in a parameter distance $\mathrm{d}v$ along the null geodesic is

$$\mathrm{d}(u^a k_a) = (u^a k_a)_{;b} k^b\,\mathrm{d}v = (u_{a;b} k^a k^b)\,\mathrm{d}v + u_a(k^a{}_{;b} k^b)\,\mathrm{d}v\,.$$

The second term vanishes by (6.9). Substituting from (2.10) and (6.12),

$$\mathrm{d}(u^a k_a) = (\theta_{ab} n^a n^b + \dot u_a n^a)(u^c k_c)^2\,\mathrm{d}v\,.$$

As (6.10) implies that the change $\mathrm{d}\lambda$ in any wavelength λ in the parameter distance $\mathrm{d}v$ is given by

$$\frac{\mathrm{d}\lambda}{\lambda} = -\frac{\mathrm{d}(u_a k^a)}{(u_b k^b)}\,,$$

the change of red-shift along the null geodesic is

(6.14*a*) $$\frac{d\lambda}{\lambda} = (\theta_{ab} n^a n^b + \dot{u}^a n^a)\, dl\,.$$

Using (2.11), this can be written (EHLERS [22])

(6.14*b*) $$\frac{d\lambda}{\lambda} = (dl)^{\bullet} + (\dot{u}^a n_a)\, dl\,.$$

Thus the red-shift has been split (because we have a unique 4-velocity u^a determined at each point) into a radial « Doppler » part (the first term, equivalent to (6.11*b*) for small distances when (2.11) is valid, since the latter then implies the motion is slow) and a « gravitational » part (the second term). Further, we can see how this red-shift-distance relation varies with direction in the sky; since the angular dependence of the terms due to θ, σ_{ab} and $\dot{u}_a$ are different, we can in principle determine these quantities directly from the linear red-shift-distance relation around that point, estimating the distances from the observed brightness of the sources (see Subsect. **6**'4).

6'2.2. Spherically symmetric cosmological models. To illustrate these relations, we consider a Robertson-Walker universe in the co-ordinates of Subsect. **5**'1.3. By the homogeneity and isotropy of these space-times all future-directed null geodesics are equivalent, so it suffices to consider future-directed radial null geodesics through the origin of co-ordinates. The corresponding solution of the geodesic equation is

(6.15) $$k^a = \frac{1}{R}\left(1, \frac{1}{R}, 0, 0\right) \Leftrightarrow k_a = \frac{1}{R}(-1,\ R, 0, 0)\,.$$

Since the fluid velocity vector is $u^a = (1, 0, 0, 0)$, we find

(6.16*a*) $$-u^a k_a = \frac{1}{R}\,.$$

Thus in these models,

(6.16*b*) $$1 + z = \frac{R_{\text{observer}}}{R_{\text{emitter}}}\,.$$

This result can also be obtained by direct integration of (6.14), or by simple geometric methods (cf. NARLIKAR and DAVIDSON [9], and the interesting discussion by SCHRÖDINGER [71]).

6'3. *Polarization.* – It follows from (6.5*b*) that the state of polarization of the light is completely unaffected by the curvature of space-time. More precisely, any numerical parameters describing the polarization are unchanged along the null geodesics, while any directions associated with the polarization are parallely propagated along the null geodesics.

6'4. *Luminosity.* – The quantiy $\mathfrak{f}$ in eq. (6.8) is the instantaneous flux of the radiation; the rate of change of $\mathfrak{f}$ along the null geodesics is determined by the equation

$$(A^2)_{;a}k^a = -A^2 k^a{}_{;a} \tag{6.17}$$

which follows from (6.5*b*). However what is measured in practice is not $\mathfrak{f}$, but a time-average of $\mathfrak{f}$ over a fairly large number of high-frequency oscillations. Thus the observed flux (*) is $A^2(G(\varphi))^2(k^a u_a)^2$ where $G(\varphi)$ is a suitable average of $g'(\varphi)$, and is a slowly varying function of φ. As $G(\varphi)$ is constant along the null geodesics, it can be absorbed into A without affecting (6.17). The measured flux F can therefore be written in the form

$$F = A^2(k_a u^a)^2 \,. \tag{6.18}$$

6'4.1. The area law. Consider a bundle of null geodesics diverging from a radiation source, with cross-sectional area $\mathrm{d}S$ perpendicular to the propagation vector $\boldsymbol{k}$ at a point with affine parameter value v. We quote two geometrical results describing the geometry of such a bundle of null geodesics (see SACHS [70] or PIRANI [37] for proofs):

a) the measurement $\mathrm{d}S$ is independent of the 4-velocity of the observer measuring $\mathrm{d}S$;

b) the change of $\mathrm{d}S$ along the null geodesics is determined by

$$\frac{\mathrm{d}}{\mathrm{d}v}(\mathrm{d}S) = (\mathrm{d}S)_{;a}k^a = \mathrm{d}S\, k^a{}_{;a} \,. \tag{6.19}$$

Combining (6.19) with (6.17) shows that $(A^2\mathrm{d}S)$ is constant along the geodesics. Thus

$$A^2\mathrm{d}S|_1 = A^2\mathrm{d}S|_2 \tag{6.20}$$

describes the change of A along the bundle of null geodesics. Combining

(*) Rate at which radiation crosses unit area per unit time.

(6.18) and (6.20) shows that

$$F \propto \left[\frac{(k^a u_a)_{\text{observer}}}{(k^b u_b)_{\text{emitter}}}\right]^2 \frac{1}{\mathrm{d}S}\,,$$

(the factor $(k_b u^b)_{\text{emitter}}$ being constant along the geodesic), so from (6.10) we find

$$F = \frac{\text{constant}}{(1+z)^2\,\mathrm{d}S}\,, \tag{6.21}$$

gives the measured flux at any point along the bundle of null geodesics. The $(1+z)$ factors may be understood as arising from i) the loss of energy suffered by each photon due to the red-shift, and ii) the lower measured rate of arrival of photons due to the time dilatation. Apart from these factors, the flux is proportional to $(\mathrm{d}S)^{-1}$; this expresses the conservation of photons along the bundle of null geodesics. When the energy conditions (3.8) are fulfilled, the space-time curvature tends to cause the bundle of null geodesics to con-

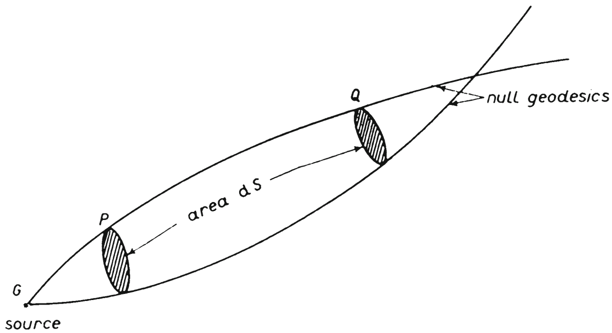


Fig. 5. – The refocussing of a bundle of null geodesics emanating from a source G; the cross-sectional areas of the bundle at P and Q are the same. As a result, the source will appear to be anomalously bright, and to have an anomalously large angular diameter, to an observer at Q.

verge (*) (Sachs [70], Penrose [72], Bertotti [73]). If there is sufficient matter present to reconverge the null geodesics, so that the cross-sectional area $\mathrm{d}S$ is the same at two points (P and Q in Fig. 5), then the factor A^2 will be the same at these two points. Thus the source will seem anomalously bright to an observer at Q; if he and an observer at P both adjust their velocities so as to see the same red-shift, they both measure the same flux of radiation

(*) This is true even if space-time is empty.

from the source (in practice the red-shift factors will often nullify this effect). Near a point where the null geodesics are refocussed, this gravitational lens effect tends to produce very high fluxes (cf. Salpeter's seminar).

6'4.2. Relation to the source luminosity. The constant in eq. (6.21) has still to be related to the source characteristics. The luminosity of the source is defined as the total rate of emission of radiant energy by the source. In principle, the luminosity of the source at some instant t_0 would

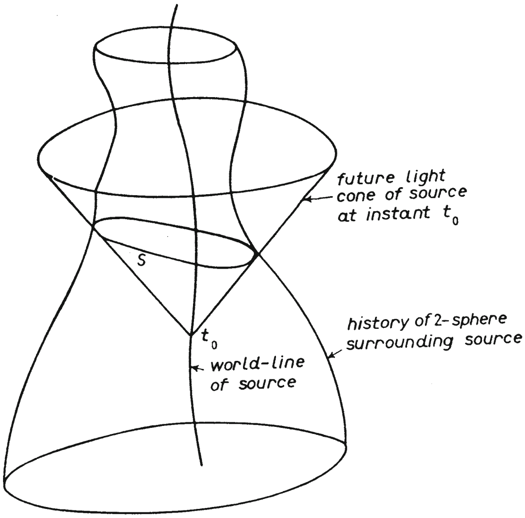


Fig. 6. – A space-time diagram of the intersection S of the future light cone of the source at an instant t_0, with the history of a 2-sphere surrounding the source. The total light emitted by the source at instant t_0 can be measured on S.

have to be measured by enclosing it in a 2-sphere and measuring the rate at which radiation emitted at time t_0 crosses each surface elements $\mathrm{d}S$ (*) of the sphere (in the space-time diagram of this situation, Fig. 6, we measure the radiation on the 2-sphere S). Then we form the integral

$$L := \int_S (1+z)^2 F \,\mathrm{d}S \,. \tag{6.22}$$

By (6.21), this is a constant, independent of the choice of the 2-sphere and of its motion; this constant is the source luminosity L.

(*) Remember that $\mathrm{d}S$ is defined as the area projected perpendicular to k^a.

In practice, we observe the flux from the source along some bundle of geodesics which subtends a small angle $d\Omega_G$ at the source and has cross-sectional area dS_G at the observer. (Subscript G denotes the bundle of geodesics diverging from the galaxy.) Consider a unit sphere lying in the locally Euclidean space-time near the source, and centered on the source (this implies that its 4-velocity is the same as that of the source, so $(1+z)=1$ on this sphere). Let the value of F on this sphere be denoted by F_G. Then, assuming the source radiates spherically symmetrically,

(6.23) $$L = \int_{\text{2-sphere}} F_G \, dS = 4\pi F_G$$

relates the source luminosity to F_G. Now (6.21) applied to the bundle of null rays shows that

(6.24) $$((1+z)^2 F \, dS_G) = \text{const} = F_G \, d\Omega_G \,,$$

where we have evaluated the constant on the unit 2-sphere. Therefore if we define the *galaxy area distance* r_G of the source from the galaxy by

(6.25) $$dS_G = r_G{}^2 \, d\Omega_G \,,$$

eqs. (6.23), (6.24) show that

(6.26*a*) $$F = \frac{L}{4\pi} \frac{1}{(r_G)^2(1+z)^2} \,,$$

is the observed flux from the source in terms of the galaxy luminosity L, red-shift z and area distance r_G. The effects of the curvature of space-time are contained in the factor $(1/r_G)^2$, as definition (6.25) determines r_G as a function of the affine parameter (and so, in a cosmological model, of the red-shift). In the context of astronomical measurement, this flux is called the *observed luminosity* of the source; the *apparent magnitude* m of the source is defined by

$$m = -2.5 \log_{10} F + \text{const}$$

so we can rewrite (6.26*a*) as

(6.26*b*) $$m = -2.5 \log_{10} L + 5 \log_{10} r_G(1+z) + \text{const} \,.$$

6.4.3. Distance definitions. Since we cannot measure the solid angle $d\Omega_G$, the galaxy area distance r_G is not a measurable quantity. However we can define an analogous quantity, the *observer area distance* r_0, which is, in principle, observable. Let $\mathrm{d}\Omega_0$ be the solid angle subtended by a bundle of null geodesics diverging from the observer, and let $\mathrm{d}S_0$ be the cross-sectional area of this bundle at some point. Then the observer area distance r_0 of this point from the observer is defined by

$$\mathrm{d}S_0 = r_0{}^2\,\mathrm{d}\Omega_0\,. \tag{6.27}$$

Thus we can find r_0 if we can measure the solid angle subtended by some object (e.g. a galaxy of given type, or H_{II} regions in a galaxy) whose cross-sectional area can be found from astrophysical considerations. When the distortion effect is not large (if it is large, we should be able to detect it), we can determine the area distance to reasonable accuracy from the observed angular diameter α of some object whose linear dimension d perpendicular to the line of sight can be estimated; then $d \simeq r_0\alpha$.

If we consider a given galaxy and observer, there are defined two seemingly independent area distances (r_0 and r_G) between them. An important geometrical result, discovered by ETHERINGTON in 1933 and rediscovered by SACHS and PENROSE in 1966, is that these area distances are essentially equivalent. More precisely, we have the

Reciprocity Theorem. The area distances r_0 and r_G are related by

$$(r_G)^2 = (r_0)^2(1+z)^2\,. \tag{6.28a}$$

Proof. Let a bundle of null geodesics diverging from G with solid angle $\mathrm{d}\Omega_G$ have tangent vector k^a, and let a bundle of null geodesics converging to O with solid angle $\mathrm{d}\Omega_0$ have tangent vector k'^a, where OG is a null geodesic common to both bundles (see Fig. 7); k^a and k'^a are to coincide on OG. Let v, v' be affine parameters and p^a, p'^a be connecting vectors for k^a, k'^a respectively. Then $k^a = \partial x^a/\partial v$, $k^a{}_{;b}k^b = 0$, $\mathrm{D}p^a/\mathrm{D}v := p^a{}_{;b}k^b = k^a{}_{;b}p^b$ hold, together with the corresponding primed equations (the proof of the last equation is essentially identical to the proof of eq. (2.5)). These equations imply the *geodesic deviation equation* $\mathrm{D}^2p^a/\mathrm{D}v^2 = -R^a{}_{bcd}k^bp^ck^d$, and the corresponding primed equation. Since $R^a{}_{bcd}k^bk^d = R^a{}_{bcd}k'^bk'^d$ and $v = v'$ along OG, these equations and the Riemann tensor symmetries (4.1) show

$$p'^a\frac{\mathrm{D}^2p_a}{\mathrm{D}v^2} - p^a\frac{\mathrm{D}^2p'_a}{\mathrm{D}v^2} = 0 \text{ along } OG\,.$$

Therefore

$$p'^a\frac{\mathrm{D}p_a}{\mathrm{D}v} - p^a\frac{\mathrm{D}p'_a}{\mathrm{D}v} = \text{constant along } OG.$$

Evaluating this constant at O (where $p'^a = 0$) and at G (where $p^a = 0$), we find

$$p^a \Big|_O \frac{\mathrm{d}p'_a}{\mathrm{d}v}\Big|_O = -p'_a\Big|_G \frac{\mathrm{d}p^a}{\mathrm{d}v}\Big|_G . \tag{6.29}$$

To completely determine the set of connecting vectors we consider, we specify that p^a will be orthogonal to the observers 4-velocity at O and p'^a will be orthogonal to the galaxies 4-velocity at G; then the set of connecting vectors p^a, p'^a satisfying these conditions is 2-dimensional at each point on OG. We write the magnitudes of $\boldsymbol{p}, \boldsymbol{p}'$ as p, p'.

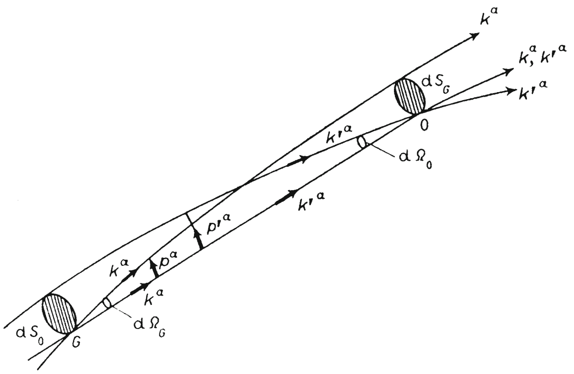


Fig. 7. – Bundles of null geodesics diverging from a source G and converging to an observer O, with a null geodesic OG common to both bundles. If O moves so as to see no redshift in the source spectrum, then equal areas $\mathrm{d}S_O$, $\mathrm{d}S_G$ subtend equal solid angles $\mathrm{d}\Omega_O$, $\mathrm{d}\Omega_G$.

We wish to use relation (6.29) to relate the areas $\mathrm{d}S_O$, $\mathrm{d}S_G$ of the bundles of geodesics at G, O respectively. To do so, we choose from the connecting vectors $\boldsymbol{p}$ a pair $\boldsymbol{p}_1, \boldsymbol{p}_2$ such that

$$\frac{\mathrm{d}p_1{}^a}{\mathrm{d}v}\Big|_G \frac{\mathrm{d}p_{2a}}{\mathrm{d}v}\Big|_G = 0 . \tag{6.30a}$$

In general, these vectors will not be orthogonal at O; let the angle between them at O be $\pi/2 + \psi$. If we rotate the pair $\boldsymbol{p}_1, \boldsymbol{p}_2$ at G through an angle $\pi/2$, this maps $\boldsymbol{p}_1$ to $\boldsymbol{p}_2$ and $\boldsymbol{p}_2$ to $-\boldsymbol{p}_1$, and so changes the angle at O to $\pi/2 - \psi$. Thus if we rotate the pair through all angles up to $\pi/2$ at G, the angle between them at O will certainly take all values between $\pi/2 + \psi$ and $\pi/2 - \psi$. We can therefore find a particular pair of connecting vectors $\boldsymbol{p}_1, \boldsymbol{p}_2$ which satisfy

(6.30a) and also

(6.30b) $$p_1{}^a|_O p_{2a}|_O = 0\,.$$

We now choose from the connecting vectors p' a pair p'_1, p'_2 determined by the condition

$$p'_{1a}\bigg|_G \frac{\mathrm{d}p_{2a}}{\mathrm{d}v}\bigg|_G = 0\,, \qquad p'_{2a}\bigg|_G \frac{\mathrm{d}p_1{}^a}{\mathrm{d}v}\bigg|_G = 0\,.$$

Then (6.29) shows

$$p_2{}^a\bigg|_O \frac{\mathrm{d}p_1{}'_a}{\mathrm{d}v}\bigg|_O = 0\,, \qquad p_{1a}\bigg|_O \frac{\mathrm{d}p_2{}'^a}{\mathrm{d}v}\bigg|_O = 0\,,$$

i.e. we have (because of (6.30a), (6.30b)) the conditions

(6.30c) $$p_1{}'^a|_G p_2{}'_a|_G = 0\,, \qquad \frac{\mathrm{d}p_1{}'^a}{\mathrm{d}v}\bigg|_O \frac{\mathrm{d}p_2{}'_a}{\mathrm{d}v}\bigg|_O = 0\,.$$

With this choice of connecting vectors,

$$\mathrm{d}S_G = p_1|_O p_2|_O\,, \qquad \mathrm{d}S_O = p'_1|_G p'_2|_G\,, \qquad \mathrm{d}\Omega_O = \frac{\mathrm{d}p'_1}{\mathrm{d}l}\bigg|_O \frac{\mathrm{d}p'_2}{\mathrm{d}l}\bigg|_O\,,$$

$$\mathrm{d}\Omega_G = \frac{\mathrm{d}p_1}{\mathrm{d}l}\bigg|_G \frac{\mathrm{d}p_2}{\mathrm{d}l}\bigg|_G\,.$$

However (6.29) shows

$$(p_1 p_2)\bigg|_O \left(\frac{\mathrm{d}p'_1}{\mathrm{d}v}\frac{\mathrm{d}p'_2}{\mathrm{d}v}\right)\bigg|_O = (p'_1 p'_2)\bigg|_G \left(\frac{\mathrm{d}p_1}{\mathrm{d}v}\frac{\mathrm{d}p_2}{\mathrm{d}v}\right)\bigg|_G\,.$$

Using (6.13) this relation is

$$\mathrm{d}S_G\,\mathrm{d}\Omega_O(k^a u_a)^2|_O = \mathrm{d}S_O\,\mathrm{d}\Omega_G(k_a u^a)^2|_G\,,$$

which is relation (6.28) in view of (6.25), (6.27) and (6.10).

This result is simply a consequence of the geodesic deviation equation. It tells us that if $z = 0$, then equal surface elements $\mathrm{d}S_O$, $\mathrm{d}S_G$ subtend equal solid angles $\mathrm{d}\Omega_O$, $\mathrm{d}\Omega_G$, irrespective of the curvature of space-time (The factor $(1+z)^2$ is simply the special relativistic correction to solid-angle measurements.) Note that if there is a gravitational lens effect leading to anomalously large source brightness, this is accompanied by an anomalously large source solid angle. In fact, if we have the situation (see Fig. 5) that light is refocussed, the angular diameter of a given object decreases to a minimum and then starts increasing again as that object is moved further down the past light cone of the observer.

The reciprocity theorem allows us to extend (6.26*a*) to

$$F = \frac{F_G}{(r_G)^2(1+z)^2} = \frac{F_G}{(r_0)^2(1+z)^4}, \tag{6.26c}$$

so we can rewrite (6.26*b*) as

$$m = -2.5 \log_{10} L + 5 \log_{10} r_0(1+z)^2 + \text{const}. \tag{6.26d}$$

Another way of stating the result of the theorem is to define a ***corrected luminosity distance*** r by

$$r^2 = \frac{F_G}{F(1+z)^4}.$$

(see KRISTIAN and SACHS [17]); then the theorem is

$$r = r_0. \tag{6.28b}$$

Since we can in principle find the corrected luminosity distance r for any source with known intrinsic luminosity by measuring the flux from the source, we can (in principle) verify eq. (6.28*b*) experimentally. However in practice (see Subsect. 6˙6.2) we are usually unable to measure r, r_0 independently.

In the literature, r, r_0 and D (defined by $D^2 := F_G/F$) have all been called luminosity distances; this has caused some confusion about what the correct red-shift factors in various formulae are (*). A further distance defined by the null geodesics is the ***parallax distance*** r_p, defined by $r_p = 2/k^a{}_{;a}|_G$, which is equivalent to the usual definition of parallax distance when there is no distortion (JORDAN, EHLERS and SACHS [74]); however this distance is clearly not measurable in the cosmological context.

These distances are all (in principle) directly measurable quantities, which reduce to the usual special relativistic distance *a*) for slowly moving nearby sources, and *b*) in a static situation in a flat space-time. However in the situation in which there is refocussing and so a minimum angular diameter for any given object, some or all of these distances will be double-valued. Even the cosmological red-shift, which is often a single-valued directly observable di-

(*) When we refer to a « luminosity distance » it will be the uncorrected distance D, and a « corrected luminosity distance » will be $r = r_0$.

stance indicator, may be double-valued in certain circumstances (for example, in a Robertson-Walker universe ($\Lambda > 0$, $K > 0$) which contracts from an infinite to a finite radius and then re-expands to an infinite radius).

6.4.4. Spherically symmetric cosmological models. As an example, consider a Robertson-Walker universe with the metric given in Subsect. 5.1.3, and the source taken at the centre of the spatial co-ordinates; we use (6.15) to give the components of k^a. Let the time at the observev be t_0 and the time at the source be t_G.

A small affine parameter change $\mathrm{d}v$ along the null geodesics corresponds to co-ordinate changes $\mathrm{d}t$, $\mathrm{d}r$ where

$$\mathrm{d}v = R\,\mathrm{d}t = R^2\,\mathrm{d}r\,, \tag{6.31}$$

(cf. eq. (6.13)). The divergence of k^a is

$$k^a{}_{;a} = \frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^a}(\sqrt{-g}\,k^a) = \frac{2}{R^2}\left(R^{\boldsymbol{\cdot}} + \frac{f'(r)}{f(r)}\right).$$

Since (6.19) implies $[\log \mathrm{d}S]_G^O = \int_G^O k^a{}_{;a}\,\mathrm{d}v$, we can use (6.31) together with this equation and the definition of $\mathrm{d}\Omega_G$ to find

$$\log(\mathrm{d}S_G/\mathrm{d}\Omega_G) = \int_G^O \frac{2R^{\boldsymbol{\cdot}}}{R^2}\cdot R\,\mathrm{d}t + \int_G^O \frac{2f'(r)}{R^2 f(r)}\cdot R^2\,\mathrm{d}r = \log\left(R^2 f^2(r)\right)\Big|_G\,, \tag{6.32}$$

since

$$\lim_{r\to 0}\left(\frac{\mathrm{d}S}{R^2 f^2(r)}\right)\Big|_G = \mathrm{d}\Omega_G\,.$$

Here r is the radial co-ordinate difference between the source and the observer; to express the result in a manner independent of the spatial co-ordinates used, we note that

$$r = \int_G^O \mathrm{d}r = \int_{t_G}^{t_0} \frac{\mathrm{d}t}{R(t)}\,,$$

the equivalence holding because the light moves on a null geodesic (see (6.31)). Thus if we define

$$u := \int_{t_G}^{t_0} \frac{\mathrm{d}t}{R(t)}\,, \tag{6.33}$$

then (6.32) is equivalent to

$$r_G^2 = R^2|_O f^2(u)\,, \tag{6.34a}$$

where

$$f(u) = \begin{cases} \sin u \\ u \\ \sinh u \end{cases} \quad \text{if} \quad k = \begin{cases} +1 \\ 0\,. \\ -1 \end{cases}$$

Similarly we can show that

$$r_O^2 = R^2|_G f^2(u) = (1+z)^{-2} R^2|_O f^2(u)\,, \tag{6.34b}$$

which is consistent with (6.16) (*). Whenever $\dot{R} \neq 0$ we can express u in the form

$$u = \int_{R_G}^{R_O} \frac{\mathrm{d}R}{R\dot{R}} = \int_O^z \frac{\mathrm{d}z}{(1+z)\dot{R}}\,,$$

and then substitute from the Friedmann equation to obtain $u(z)$ and so $r_O(z)$. For example when pressure-free matter is the dominant energy component in the universe (as it is at recent times) we can use the Friedmann equation in the form (5.4c) and numerically integrate to obtain $r_O(z)$. (See, for example, Refsdal, Stabell and de Lange [53].) In the particular case $\Lambda = 0$ we can integrate analytically to find (using (5.5b) to evaluate R_0 when $k \neq 0$) that

$$r_0 = \frac{1}{H_0 q_0^2 (1+z)^2}\left\{q_0 z + (q_0 - 1)\left((1 + 2q_0 z)^{\frac{1}{2}} - 1\right)\right\}, \tag{6.35a}$$

when $q_0 \neq 0$, and

$$r_0 = \frac{1}{2H_0}\left(1 - \frac{1}{(1+z)^2}\right), \tag{6.35b}$$

when $q_0 = 0$. (cf. Mattig [75], Sandage [76]). Combining these with (6.26d) gives the usual m-z relations.

We see that when $q_0 \neq 0$, there is a maximum value of r_0 for some value of z, and thereafter r_0 decreases as z increases; *i.e.* we *do* have refocussing of

(*) We can also derive these equations direct from the Robertson-Walker geometry, (cf. Davidson and Narlikar [9]).

null geodesics in these models. This is still true if we include a pressure term in the Friedmann equation, so it is true in all nonempty Robertson-Walker universes which expand from a singularity (we can always ignore a Λ-term at early enough times). However the (uncorrected) luminosity distance $D = r_0(1+z)^2$ always decreases as z increases; *i.e.* although the area factor in (6.26*c*) tends to increase F after a certain point, the cosmological red-shift factors suffice to ensure that the observed flux from identical sources always decreases as distance increases.

Equations (6.35*a*), (6.35*b*) imply a simple expression for z as a function of D; in fact

$$1 + z = q_0(1 + DH_0) - (q_0 - 1)\sqrt{1 + 2DH_0}\,. \tag{6.35c}$$

6'5. *Number counts in a cosmological model.* – Consider a small affine parameter displacement $\mathrm{d}v$ at a point A on a bundle of past null geodesics subtending a solid angle $\mathrm{d}\Omega_0$ at 0; this corresponds to a distance $\mathrm{d}l = (-k^a u_a)\,\mathrm{d}v$

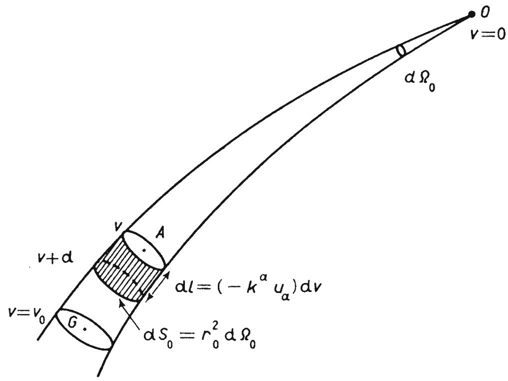


Fig. 8. – A section $(v, v + dv)$ of a bundle of null geodesics which subtends a solid angle $\mathrm{d}\Omega_0$ at 0.

in the rest-frame of a galaxy at A. The cross-section area of the bundle is $\mathrm{d}S_0 = r_0^2\,\mathrm{d}\Omega_0$, so if the number-density of radiation sources is n per unit proper volume, then the number of sources in this section of the bundle (see Fig. 8) is

$$\mathrm{d}N = r_0^2\,\mathrm{d}\Omega_0(n(-k^a u_a))_A\,\mathrm{d}v\,. \tag{6.36}$$

Integrating this expression with respect to v gives the number $N(v)$ of sources seen by O in the solid angle $\mathrm{d}\Omega_0$ which lie at affine parameter distances less than v. We can combine $N(v)$ with the functions $r_0(v)$, $z(v)$ to obtain $N(F)$, the number of sources with intrinsic luminosity seen in $d\Omega_0$ with observed flux greater than F, or $N(z)$, the number of sources seen in $d\Omega_0$ with red-shift less than z.

6.5.1. Spherically symmetric models. In a Robertson-Walker universe, if the number of sources is conserved then $n = n_0(R_0^3/R^3)$, and (6.36) becomes

$$dN = \frac{n_0 R_0^3}{R^3} . d\Omega_0 . R^2 f^2(u) \frac{1}{R} . R^2 du ,$$

on using (6.16a), (6.31), (6.34b). Thus

$$N(u) = 4\pi n_0 R_0^3 \int_0^u f^2(u')\, du' = 4\pi n_0 R_0^3 \begin{cases} \frac{1}{2}(u - \sin u \cos u), & \text{if } k = +1, \\ u^3/3 & \text{if } k = 0, \\ \frac{1}{2}(\sinh u \cosh u - u), & \text{if } k = -1, \end{cases}$$

is the number of sources in all directions at distances up to that characterized by u. Using (6.34) and (5.5b) we can write this as a function of $r_0(1+z)$, (*i.e.* of r_G):

$$\text{(6.37}a\text{)} \qquad N = \begin{cases} \dfrac{2\pi n_0}{(H_0^2(2q_0-1)+\Lambda)^{\frac{3}{2}}} \left(\arcsin f \mp f\sqrt{1-f^2}\right), & \text{if } k = +1, \\ \dfrac{4\pi}{3} n_0\big((1+z) r_0\big)^3, & \text{if } k = 0, \\ \dfrac{2\pi n_0}{(H_0^2(1-2q_0)-\Lambda)^{\frac{3}{2}}} \left(f\sqrt{1+f^2} - \arg\sinh f\right), & \text{if } k = -1, \end{cases}$$

where

$$\text{(6.37}b\text{)} \qquad f = \big(k(H_0^2(2q_0-1)+\Lambda)\big)^{\frac{1}{2}} \cdot r_0(1+z), \qquad \text{when } k \neq 0.$$

(Care has to be taken in choosing the correct sign $\mp$ when $k = +1$ (Roeder and McVittie [77]); note also that N is monotonic with affine distance although r_0 is not.) Combining this expression with $r_0(z)$, we obtain $N(z)$; alternatively we can eliminate z to obtain N as a function of one of the area distances. In particular if $\Lambda = 0$ in a dust-filled universe, (6.35c) shows we can obtain the simple expresssion

$$r_0(1+z) = \frac{D}{q_0(1+DH_0) - (q_0-1)\sqrt{1+2DH_0}},$$

which, combined with (6.37) expresses $N(D)$ in terms of simple analytic functions. (cf. MATTING [78], SANDAGE [76].)

6'6. *Observed specific intensity.* – The theoretical relations we have derived so far need to be modified in three ways to correspond more closely to what is actually observed.

6'6.1. Specific flux. So far, we have considered a total (bolometric) luminosity L and flux F. However we actually observe in very restricted wavelength ranges (in radio observations, nearly at one frequency: in optical observations usually in the U, B, V bands). To allow for this, we represent the source spectrum by a function $\mathscr{I}(\nu)$, where $L\mathscr{I}(\nu)\,\mathrm{d}\nu$ is the rate at which radiation is emitted by the source at frequencies between ν and $\nu+\mathrm{d}\nu$; $\mathscr{I}(\nu)$ is normalized by the condition $\int_0^\infty \mathscr{I}(\nu)\,\mathrm{d}\nu = 1$. The frequency ν_o measured by some observer is related to the frequency ν_G of the same radiation in the rest-frame of the emitting galaxy by $\nu_o = \nu_G/(1+z)$ (cf. Subsect. 6'2), which implies $\mathrm{d}\nu_o = \mathrm{d}\nu_G/(1+z)$. Equation (6.26) can therefore be written

$$F = \frac{L}{4\pi}\,\frac{\int_0^\infty \mathscr{I}(\nu_G)\,\mathrm{d}\nu_G}{r_0^2(1+z)^4} = \frac{L}{4\pi}\,\frac{\int_0^\infty \mathscr{I}\big(\nu_o(1+z)\big)\,\mathrm{d}\nu_o}{r_0^2(1+z)^3}\,.$$

Thus the flux measured in the frequency range ν, $\nu+\mathrm{d}\nu$ by the observer is

$$F_\nu\,\mathrm{d}\nu = \frac{L}{4\pi}\,\frac{\mathscr{I}\big(\nu(1+z)\big)\,\mathrm{d}\nu}{r_0^2(1+z)^3} = \frac{L}{4\pi}\,\frac{\mathscr{I}\big(\nu(1+z)\big)\,\mathrm{d}\nu}{r_G^2(1+z)}\,. \tag{6.38a}$$

We call F_ν the *specific flux* (in the case of radio sources, it is called the *flux density*) of the radiation. We can often assume $\mathscr{I}(\nu)$ has the form $\mathscr{I}(\nu) = \mathrm{const}\,\nu^{-\alpha}$, where α is a constant (the « spectral index »); (this is a good approximation for many radio sources when $0.7 \leqslant \alpha \leqslant 0.9$ and for many optical sources at wavelengths longer than 5000 Å when $\alpha \simeq 2$). Then we find

$$F_\nu = \frac{F_G\mathscr{I}(\nu)}{r_0^2(1+z)^{3+\alpha}} = \frac{F_G\mathscr{I}(\nu)}{r_G^2(1+z)^{1+\alpha}}\,. \tag{6.38b}$$

An alternative way of allowing for the effect of the source spectrum is to introduce a correction term, the « K-correction » (SANDAGE [79]) which represents the difference between F_ν (given by (6.38a)) and F (given by (6.26)).

6'6.2. Intensity. So far we have implicitly assumed the sources observed are point sources. In practice we usually observe extended sources, and instead of measuring the flux from the source, our direct measurements tell us the flux per unit solid angle from the source (HOYLE [80]), that is, the *intensity* of radiation from the source. Considering a source (or part of a source) of area A, we find from (6.26*c*) (6.27) that the intensity is

$$I := \frac{F}{\mathrm{d}\Omega_o} = \frac{I_G}{(1+z)^4}, \tag{6.39}$$

where the factor $I_G := F_G/A$ depends simply on the source characteristics. For a given source, the intensity is independent of the area distance r_o and depends only on the red-shift z; if $z = 0$, I has the value I_G which is the *surface brightness* of the source. Thus, for example, the density of marking of an image on a photographic plate is independent of the space-time curvature, depending simply on the source surface brightness and observed red-shift; as we can find z independent of I, we can in principle find the surface brightness of the source direct from our measurements.

To find the flux F from an extended source, we have to measure I and then integrate over the image to obtain F; therefore we have in fact (except in the case of quasars or very distant galaxies), explicitly or implicitly, to estimate the solid angle subtended by the source before we can deduce F from our direct observations (see HOYLE [80] for further discussion). It is for this reason that we are in practice unable to measure the corrected luminosity distance r and area distance r_o independently (see Subsect. 6.4.3). We may in fact lose useful information if we consider only the flux F (the « magnitude » of the source) rather than the intensity and solid angle information which is combined to give F.

We may note that it is expression (6.39) which is involved in Olber's paradox (BONDI [1]; see also WHITROW and YALLOP [81], HARRISON [82] and references cited there); this expression shows that we have the alternatives of assuming I_G is very low for the sources along almost all null geodesics from us (perhaps because of the source evolution which must occur in view of the finite source lifetime), or that the red-shift increases indefinitely along almost all null geodesics.

6.6.3. Specific intensity. Combining the effects considered in Sections 6.6.1 and 6.6.2, we see that what we usually measure directly is the flux per unit solid angle in some frequency range $(\nu, \nu + \mathrm{d}\nu)$. That is, we measure the *specific intensity* I_ν of radiation from the source; (6.27), (6.38) show

$$I_\nu \mathrm{d}\nu := \frac{F_\nu \mathrm{d}\nu}{\mathrm{d}\Omega} = I_G \cdot \frac{\mathscr{I}(\nu(1+z))\,\mathrm{d}\nu}{(1+z)^3}, \tag{6.40}$$

which for a given source depends simply on the redshift of the source ($I_G \cdot \mathscr{I}(\nu)$ is the surface brightness of the source at frequency ν). To estimate the specific flux (*e.g.* to estimate the visual magnitude m_ν) from an extended source, we have to integrate the observed specific intensity, which is what is actually measured by the photographic plate or radio receiver, over the image of the source.

As an application of (6.40), we consider measurement of radiation from a source emitting black-body radiation. Defining the function $g(\nu)$ by $g(\nu) = I_G \mathscr{I}(\nu)/\nu^3$, we can rewrite eq. (6.40) as

(6.41*a*) $$I_\nu \,\mathrm{d}\nu = g(\nu(1+z))\,\nu^3 \,\mathrm{d}\nu\,,$$

which gives the observed specific intensity at each frequency ν for any observer who measures the source red-shift as z. On the other hand,

$$(I_\nu \mathrm{d}\nu)|_G = f\left(\frac{\nu}{T_e}\right)\nu^3 \mathrm{d}\nu\,,$$

is the specific intensity of radiation at the source (*), where $f(\nu/T_e)$ is the Planck function for black-body radiation at temperature T_e. Comparing these expressions at the source shows $g(\nu) = f(\nu/T_e)$, which implies

$$g(\nu(1+z)) = f\left(\frac{\nu(1+z)}{T_e}\right).$$

Hence we can rewrite eq. (6.41*a*) as

(6.41*b*) $$I_\nu \mathrm{d}\nu = f\left(\frac{\nu}{T}\right)\nu^3 \mathrm{d}\nu\,,$$

where we have defined

(6.42) $$T := \frac{T_e}{1+z}\,;$$

i.e. an observer who measures the source red-shift as z, sees the radiation as black-body radiation at temperature T where T is defined by (6.42). Thus black-body radiation propagates with an unchanged spectrum through curved space-time, the observed intensity of radiation depending only on the observer's 4-velocity. (cf. Ehler's lectures for a different derivation of this result). We may note that in the Robertson-Walker models, (6.42) enables us to find the change of black-body temperature along the null geodesics, (3.23) enables us to find it along timelike geodesics, and (6.16) shows these results are consistent.

(*) This form follows direct from Wien's law.

This calculation is equally valid if we regard it as relating to observations at one point. Suppose an observer S moving with 4-velocity u^a measures the temperature of black-body radiation, which has a propagation vector k^a, as T, while an observer S' at the same point moving with 4-velocity u'^a measures the temperature of the same radiation as T_0. Then

$$T = \frac{T_0}{1+z} = T_0 \frac{(k^a u_a)}{(k^b u'_b)},$$

where z is the red-shift S measures for radiation from a source moving with 4-velocity u'^a. If the relative velocity of S and S' has magnitude v and the angle measured by S between the direction of $\boldsymbol{k}$ and the direction of motion of S' is θ, we can directly evaluate this equation by using (6.11a) and (6.12). (Note that $e^a n_a = \cos\theta$). We find (*)

$$T = \frac{T_0}{\cosh\beta - \sinh\beta\cos\theta} = T_0 \frac{(1-v^2)^{\frac{1}{2}}}{(1-v\cos\theta)}, \tag{6.43}$$

where $v = \operatorname{tgh}\beta$. In particular, if S' observes an isotropic black-body radiation field with temperature T_0, then S observes a black-body radiation field with temperature distribution in the sky given by (6.43). Conversely, if S sees such a radiation field, then there exists an observer S' who sees an isotropic radiation field. S can find the 4-velocity of S' by noting the maximum and minimum black-body temperatures T_{max}, T_{min} on his celestial sphere; then the velocity of S' relative to S is in the direction in which S observes T_{min} and has magnitude

$$v = \frac{T_{max} - T_{min}}{T_{max} + T_{min}},$$

while the isotropic radiation temperature measured by S' is $T_0 = \sqrt{T_{max} T_{min}}$. (Remember we are using units in which the speed of light is 1.)

6.6.4. **Absorption and emission in a cosmological model.** Finally, we should allow for absorption and emission of radiation along the line of sight from the source G to the observer O. Let $I_\nu(v)$ be the specific intensity of radiation travelling along a bundle of null geodesics from G to O as measured by an observer moving with the average velocity u^a at a point A (affine parameter distance v from O) in a cosmological model. We consider the change in I_ν as v increases to $v + dv$ (see Fig. 8). By eq. (6.40), we can represent the change in I_ν due to geometrical and red-shift effects alone by the differential

(*) This is just the generalization of (6.11 b) to include the transverse Doppler shift.

equation

$$\frac{dI_{\nu'}}{dv} = \frac{3}{1+z} I_{\nu'} \frac{dz}{dv},$$

where $\nu' := \nu(1+z)$ is the frequency of radiation at A which, when red-shifted to O, is observed at frequency ν. Let $S(v, \nu)\,d\nu$ be the rate of emission of radiation by each source at A per unit solid angle in the frequency range ν to $\nu + d\nu$, let $n_s(v)$ be the number density of sources at A, let $n_a(v)$ be the number density of particles scattering or absorbing radiation at A, and let $\sigma(v, \nu)$ be the interaction cross-section of these particles at frequency ν (we may allow for re-emission by suitable additions to either σ or S; similarly we may, if we wish, represent absorption processes by a negative contribution to S). Allowing for these processes in the volume $dl\,dS_0 = (-u_a k^a)\,dv\,dS_0$ at A, the change in I'_ν along the geodesic can be represented by the differential equation

$$\frac{dI_{\nu'}}{dv}(v) - \frac{3}{1+z} I_{\nu'}(v) \frac{dz}{dv} - n_a(v)\,\sigma(v, \nu')\,I_{\nu'}(v).(-k_a u^a)(v) = \\ = -n_s(v)\,S(v, \nu').(-k_a u^a)(v),$$

(remembering that when dv is positive, S acts as a *negative* term and σ as a *positive* term). Integrating this equation along the geodesic from $O(v = 0)$ to $G(v = v_*)$ we find the specific intensity at 0 is

$$(6.44a) \qquad I_\nu = \int_0^{v_*} \frac{n_s(v)\,S\big(v, \nu(1+z)\big)}{(1+z)^3} .\exp[-p(v, \nu)](-k_a u^a)(v)\,dv + \\ + \frac{I_{\nu(1+z_*)}(v_*)}{(1+z_*)^3} .\exp[-p(v_*, \nu)],$$

where the optical depth $p(v, \nu)$ between A and O for radiation observed at O at frequency ν is

$$(6.44b) \qquad p(v, \nu) := \int_0^v n_a(v')\,\sigma\big(v', \nu(1+z')\big)(-k_a u^a)(v')\,dv';$$

in these expressions, z and $(k^a u_a)$ are regarded as known functions of v.

This equation determines the specific intensity of radiation we observe in any direction in the sky; the second term represents radiation propagating according to eq. (6.40) from G but attenuated by absorption, while the first term represents the integrated emission from sources between O and G, again attenuated by absorption. This equation implies a very similar equation, which can also be derived directly from (6.39), determining the (integrated) intensity I; the only essential differences are that the factors $1/(1+z)^3$ are replaced by factors $1/(1+z)^4$, and the arguments containing the frequency ν are omitted (*).

6.6.5. Spherically symmetric models. To illustrate the use of these formulae, we will apply (6.44) to a Robertson-Walker universe, taking the matter content to be dust and ignoring the effects of radiation on the dynamics of the universe. Then, since (6.13) shows

$$(-k_a u^a)\,\mathrm{d}v = -\mathrm{d}t = -\frac{1}{R^{\bullet}}\frac{\mathrm{d}R}{\mathrm{d}z}\,\mathrm{d}z\,.$$

we find from (6.16), (5.4c) and (5.5b) that

(6.45)
$$(-k_a u^a)\,\mathrm{d}v = \frac{\mathrm{d}z}{(1+z)^2\{H_0^2(1+2q_0 z) + (\Lambda/3)(2z-1+(1+z)^{-2})\}^{\frac{1}{2}}}\,.$$

(We can easily allow for noninteracting radiation in (6.45) (cf. the comments following (5.4a)) if integrated emission from early times is important; if other field equations than those of general relativity are used, one can allow for this by using a suitable replacement of the Friedmann equation to determine a new form for (6.45).) For simplicity we shall further take $\Lambda = 0$ and assume that the radiation sources and absorbing particles are conserved; then

$$n(z) = n(0)(R_0^3/R^3) = n(0)(1+z)^3.$$

With these assumptions, the contribution to I_ν from sources up to a red-shift z_*, ignoring absorption, is

(6.46)
$$I_\nu = \frac{n_s(0)}{H_0}\int_0^{z_*}\frac{S(z,\nu(1+z))\,\mathrm{d}z}{(1+z)^2(1+2q_0 z)^{\frac{1}{2}}}\,.$$

We might further assume that the source emission can be described by a source spectrum $\mathcal{I}(\nu)$ which is independent of z, and an amplitude $\bar{S}(z)$; then

(6.47a)
$$S(z,\nu) = \bar{S}(z)\,\mathcal{I}(\nu)\,.$$

If in particular the emission is that of a line spectrum at frequency ν_e we can set $\mathcal{I}(\nu) = \delta(\nu - \nu_e)$ where δ is the Dirac delta function; the integrated emis-

(*) For a derivation of an equivalent expression in the Robertson-Walker case, see Davidson and Narlikar [9].

sion $\mathscr{G}_\nu(\nu_e)$ in this case is

$$\mathscr{G}_\nu(\nu_e) = \begin{cases} \dfrac{n_s(0)}{H_0 \nu_e} \dfrac{\bar{S}(z_\nu)}{(1+z_\nu)(1+2q_0 z_\nu)^{\frac{1}{2}}}, & \text{if } z_\nu \leqslant z_*, \\ 0, & \text{if } z_\nu > z^*, \end{cases}$$

where $z_\nu := (\nu_e/\nu) - 1$. (Line emission from galaxies may be important in the infrared (Gould and Sciama [83]), microwave (Petrosian, Bahcall and Salpeter [84]), X-ray (Gould and G. R. Burbidge [85]), and γ-ray (Clayton and Silk [86]), regions, while line emission from an intergalactic gas could be important in the radio (Penzias and Wilson [87]) and ultra-violet and infrared (Weymann [88]) regions.) Regarding a source of form (6.47a) as built up by such line emission spectra, we can re-express (6.46) in the form

(6.47b)
$$I_\nu = \int_0^\infty \mathscr{I}(\nu_e)\, \mathscr{G}_\nu(\nu_e)\, \mathrm{d}\nu_e .$$

While in general we have to integrate by machine, for simple spectra we can sometimes obtain analytic expressions for I_ν. For example if the sources have a spectral index α and their amplitude varies as $(R(t))^m$, then

$$S(t, \nu(1+z)) = S(t_0, \nu)(1+z)^{m-\alpha};$$

so in an Einstein-de Sitter universe ($q_0 = \frac{1}{2}$) we find from (6.46)

$$I_\nu = \frac{n_s(0)\, S(\nu, t_0)}{H_0(\alpha + \frac{3}{2} - m)} \left\{1 - \frac{1}{(1+z_*)^{\frac{3}{2}+\alpha-m}}\right\}, \qquad \text{when } m \neq \alpha + \tfrac{3}{2},$$

and in a Milne universe ($q_0 = 0$) we find

$$I_\nu = \frac{n_s(0)\, S(\nu, t_0)}{H_0(\alpha + 1 - m)} \left\{1 - \frac{1}{(1+z_*)^{1+\alpha-m}}\right\}, \qquad \text{when } m \neq \alpha + 1;$$

in the exceptional cases $q_0 = \frac{1}{2}$, $m = \alpha + \frac{3}{2}$ and $q_0 = 0$, $m = \alpha + 1$ we find

$$I_\nu = \frac{n_s(0)\, S(\nu, t_0)}{H_0} \log_e (1 + z_*) .$$

Thus if the source brightness increases faster than $(1+z)^{\alpha+\frac{3}{2}}$ in an Einstein-de Sitter universe, or than $(1+z)^{\alpha+1}$ in a Milne universe, there must be a cut-off in the source evolution before some value of z which can be determined by comparing these expressions with the observed values of I_ν. The source of radiation might be galaxies or other discrete sources (see, for example, REES and SETTI [89], and discussions by DAVIDSON and NARLIKAR [9] of the radio, LOW and TUCKER [90] of the infra-red, WOLFE and BURBIDGE [91] of the microwave, and GOULD [92] of the ultra-violet) or an intergalactic gas (see, for example, FIELD and HENRY [93] and WEYMANN [88]).

With the assumptions discussed at the beginning of this Section, (6.45) shows that the optical depth up to a red-shift z in a Robertson-Walker universe is

$$p(z, \nu) = \int_0^z \frac{n_a(0)}{H_0} \sigma(z', \nu(1+z')) \frac{(1+z')\,\mathrm{d}z'}{(1+2q_0 z')^{\frac{1}{2}}} . \tag{6.48}$$

Considering some particular scattering or absorption process, it is often reasonable to assume

$$\sigma(z, \nu) = \bar{\sigma}(z)\,\mathscr{A}(\nu) . \tag{6.49a}$$

In particular, we can represent line absorption at some frequency ν_a by setting $\mathscr{A}(\nu) = \delta(\nu - \nu_a)$; the resulting optical depth $A_\nu(z, \nu_a)$ is

$$A_\nu(z, \nu_a) = \begin{cases} \dfrac{n_a(0)}{H_0 \nu_a} \bar{\sigma}(z_\nu) \dfrac{(1+z)^2}{(1+2q_0 z_\nu)^{\frac{1}{2}}} , & \text{for } z_\nu < z_* , \\ 0 , & \text{for } z_\nu \geqslant z_* , \end{cases}$$

where $z_\nu := (\nu_a/\nu) - 1$. (Lack of observed 21 cm (FIELD [94]; PENZIAS and SCOTT [95]) and Ly α (GUNN and PETERSON [96]; BAHCALL and SALPETER [97]) absorption gives useful limits (see *e.g.* SCIAMA [98]) on the intergalactic medium; Ly α absorption could be important in Q.S.O. observations (REES [99]).) Regarding an absorption process with cross-section (6.49*a*) as built up of such line absorptions, we can re-express (6.48) in the form

$$p(z, \nu) = \int_0^\infty A_\nu(z, \nu_a)\,\mathscr{A}(\nu_a)\,\mathrm{d}\nu_a . \tag{6.49b}$$

While in general we have to integrate these equations numerically, in special cases we can integrate them analytically. In particular, when $\sigma(z, \nu) = \sigma_0 = \text{const}$, we can integrate (6.48) to show

$$p(z, \nu) = \frac{\sigma_0 n_a(0)}{H_0} \cdot \frac{1}{3q_0^2} \{(3q_0 + q_0 z - 1)(1+2q_0 z)^{\frac{1}{2}} - (3q_0 - 1)\} , \tag{6.50a}$$

when $q_0 \neq 0$, and

$$p(z, \nu) = \frac{\sigma_0 n_a(0)}{H_0} \cdot \frac{1}{2}(1+z)^2 , \tag{6.50b}$$

when $q_0 = 0$. Equation (6.50*a*) would be applicable to the Thomson scattering arising if an intergalactic medium were ionized (Bahcall and Salpeter [97]; Bahcall and May [100]); for a fully ionized gas in an Einstein-de Sitter universe with $n_a(0) \simeq 10^{-5}$ the optical depth would be unity at a red-shift of about 7, while in a low-density universe with $n_a(0) \simeq 10^{-7}$, the optical depth would be unity at a red-shift of about 30. We can also use this expression to calculate the probability of absorption effects arising from galaxies (Wagoner [101]; Roeder and Verreault [102]) or other randomly distributed matter (Bahcall and Peebles, [103]) intervening between an observer and some galaxy or quasi-stellar object in a Robertson-Walker universe.

In some situations, one would have to consider absorption and emission processes together (see, for example, studies of the spectrum of the background radiation by Weymann [88, 104], Zel'dovich and Sunyaev [105], and of absorption and emission by an intergalactic gas (Peebles [106])). Finally we note that while we have been considering (6.44) in the context of a cosmological model, one can apply it in other situations when a fluid approximation is appropriate in astrophysics; (for example, to determine the propagation of radiation in a star).

6·7. ***Null cone observations in a cosmological model.*** – In any given cosmological model, we can deduce the following quantities:

i) the proper motion in the sky observed for distant galaxies (see eq. (2.12) for first-order effect) and the distortion of images caused by the curved space-time, both of which can be found by integrating the geodesic equation (6.9);

ii) the red-shift observed for any distant sources in the model, given by (6.10*b*) (see eq. (6.14) for the first-order effect);

iii) the area distance r_0 of any source, obtained from $k^a{}_{;a}$ by integrating (6.17) and using the definition (6.27);

iv) the number N of sources of any given type observed up to some affine parameter distance v in any direction, obtained by integrating (6.36);

v) the specific intensity of radiation from any given source, and the specific intensity of background radiation in the model as a function of position in the sky (given by (6.44));

and find relations between them. For example, in a Robertson-Walker model filled with dust and with $\Lambda = 0$, there is no proper motion or distortion effect;

the area distance r_0 is related to the red-shift by (6.35) and the number density to the area distance and red-shift (and so to the luminosity distance D) by (6.37); eq. (6.44) together with (6.45) can be used to determine the observed spectrum from any source at red-shift z_* and to find the spectrum of background radiation (taking z_* to correspond to an optical depth of unity).

Observations of the black-body radiation in certain homogeneous anisotropic models have been discussed by THORNE [59], MISNER [28], REES [107] and HAWKING [108], and in perturbations of a Robertson-Walker model by SACHS and WOLFE [109], REES and SCIAMA [110], and others. The observational relations for discrete sources have been derived in certain homogeneous anisotropic models by SAUNDERS [111, 112], TOMITA [113] and MACCALLUM and ELLIS [114]; in an exact study of some inhomogeneities in Robertson-Walker models by KANTOWSKI [115]; and in perturbations of a Robertson-Walker model by BERTOTTI [73], ZIPOY [116], GUNN [117], PETROSIAN and SALPETER [118], and others. They may be obtained in the form of a power-series expansion about « here and now » in any space-time as discussed in a very interesting paper by KRISTIAN and SACHS [17] (cf. also MCCREA [119]). The divergence $k^a{}_{;a}$ is needed before r_0 can be found (*) in a general model, so it is often useful to know that $k^a{}_{;a}$ obeys an equation very similar to Raychaudhuri's equation (see for example SACHS [70], PENROSE [72]); in fact there is a close correspondence between equations obeyed by the kinematic quantities defined for a timelike congruence (see Sect. **4** above) and equations obeyed by analogous quantities (« optical scalars ») defined for a congruence of null geodesics; see JORDAN, EHLERS and SACHS [74], SACHS [70], NEWMAN and PENROSE [120]. (This correspondence is based on the fact that both timelike and null geodesics obey the geodesic deviation equation.)

In principle we can determine whether any of these space-times could be good models of the universe or not by comparing these theoretical relations directly with observation of classes of sources with known physical characteristics, and with observations of the background radiation. The essential problem in carrying out this programme is that we do not know *a priori* the properties of the particular sources we observe (cf. the discussions in this volume by E. M. BURBIDGE and G. R. BURBIDGE). Apart from statistical scatter in the properties of sources of a given type and selection effects which are exacerbated by the fact that we often have only a poor idea of what kind of object we are observing, there are two systematic effects we shall briefly mention. Firstly, the average luminosity L, spectrum $\mathscr{I}(\nu)$ and cross-section area A of a given class of sources will in general change in time, and so be functions of the red-shift z; we have to determine this evolution somehow. Secondly, we can determine r_0 either directly by measuring the angular diameters of

(*) There are other methods of obtaining r_0 which are essentially equivalent.

sources of known diameters, or indirectly by measuring the specific flux of radiation from the source and using eq. (6.38*a*). However the sources we observe in a cosmological context seldom have sharp edges: we therefore measure the angular diameter of the source up to some isophote (SANDAGE [79]) rather than measuring a metric diameter. (We can obtain angular measurements corresponding to a metric diameter by measuring the angles observed between centres of galaxies in a cluster, or between radio emission centres in a multiple radio source, but then we have considerable difficulties in estimating the dimensions of the system.) Correspondingly a measurement of the flux from the source may omit a contribution from the outer regions which are lost in the general background. Thus we can only measure the area distance r_0 of an extended source accurately if we know the surface brightness of the source as a function of frequency, time, and position on the surface of the source (cf. STOCK and SCHÜCKING [121]). If we could measure I_ν separately from r_0, this might give some guide as to the evolution taking place in the source.

7. – The observable universe.

In this Section we aim to give a brief qualitative description of the observable universe, adopting conventional interpretations of the observations. (For an alternative viewpoint, see Hoyle's Bakerian Lecture [122].)

7·1. *Causality.* – In special relativity we assume that no signal or particle can travel faster than the speed of light; it follows that an event can be causally influenced only by events within or on its past light cone, and can causally influence only events in or on its future light cone. As the past and future light cones never intersect in flat space-time, these sets of events are disjoint. In a general curved space-time, this need not be so (see KRONHEIMER and PENROSE [123] and Geroch's lectures in this volume for very general discussions). Further, imposing the field equations is not sufficient to prevent causality violations, as Gödel's solution (see Subsect. 5·2) contains closed timelike lines (GÖDEL [56]); an observer in this space-time can travel into, and influence, his own history.

We will regard solutions in which this can happen as physically unreasonable. We shall therefore assume that space-time obeys the *strong causality condition*: every point in space-time is contained in a small open neighbourhood such that every timelike (and null) curve that leaves this neighbourhood never re-enters it.

It follows from this postulate that the past light cone of any point is part of the boundary of the past of that point (HAWKING [124]); so the total region of space-time which can have a causal influence on any observer lies within (*) or on his past light cone.

7·2. *Isotropy of the microwave radiation.* – The microwave background radiation is known (see Sciama's course) to be extremely isotropic about us, so the Copernican principle (Sect. 1) leads us to believe the same is true for any observer moving with the average velocity u^a. If this isotropy were exact, it would follow from a theorem of EHLERS, GEREN and SACHS [125] that the universe was exactly a Robertson-Walker universe. This theorem states that if freely-propagating radiation is isotropic with respect to some velocity field u^a everywhere in an open set in space-time, then the shear σ_{ab} of u^a vanishes in that open set. However we can assume that in recent times $\theta > 0$ and $\dot{u}_a = 0$ (see eq. (3.16)); it then follows from the work of EHLERS, GEREN and SACHS, or from the theorem quoted in Subsect. 5·3.1, that $\omega_a = 0$ also. But these conditions ensure that space-time is a Robertson Walker universe (cf. Subsect. 5·1).

In view of this result, we shall regard the very high degree of isotropy of the radiation as evidence that the universe has been very nearly a Robertson-Walker universe since the time of last scattering of the microwave radiation.

We can in fact use the limits on the anisotropy of the radiation to place limits on the anisotropy and inhomogeneity of the universe (see THORNE [59], SACHS and WOLFE [109], REES and SCIAMA [110], HAWKING [126], WOLFE [127]) and on our motion relative to the mean velocity u^a (SCIAMA [128], STEWART and SCIAMA [129]; and cf. Subsect. 6·6.3); in particular, the measurements may be taken to indicate limits

$$\omega_0 < 10^{-3}\theta_0 , \qquad \sigma_0 < 10^{-3}\theta_0$$

on the present shear and vorticity of the universe.

The isotropy of radio source counts (**) (HUGHES and LONGAIR [130]) and of background radiation at other wavelengths seems to confirm this picture.

7·3. *The existence of singularities.* – If the universe were exactly a Robertson-Walker universe, with the energy conditions (3.8) satisfied, then the condition $\Lambda \leqslant 0$ would imply there had been a singularity a finite time t_0 ago in the past; if $q_0 > 0$ the same conclusion would hold irrespective of the sign of Λ (cf. Subsection 5·1.1). Even if q_0 were not positive it would be plausible that the same

(*) GEROCH has pointed out that if there are « wormholes » in the universe, part of the « interior » of the light cone can reappear *outside* the light cone.

(**) Isotropy of the number counts may also be used to put limits on the vorticity, cf. GÖDEL [43].

result were true, because it is plausible to assume that the microwave radiation is black-body radiation (cf. Sciama's course) which has been thermalized by being in equilibrium with ionized matter. Decoupling of the primeval fireball of matter and radiation would take place when $R/R_0 \sim 1/1000$ (cf. Sect. **2**), but if a dense intergalactic gas had been ionized by reheating, the last effective scattering of radiation could occur as late as a red-shift of about 7 (corresponding to an optical depth unity, see Sect. **6**). However the reheated gas could not thermalize the radiation to its observed temperature, so we could assume that the existence of black-body radiation implied that at some time, $R/R_0 \leqslant 1/1000$. But at that time $\mu \geqslant 10^9 \mu_0$, so the Λ-term was totally negligible (cf. estimate (4.14*c*)). In fact even if the reheated matter were responsible for the thermalization we would find $R/R_0 \leqslant \frac{1}{8}$ at some time; but then $\mu \geqslant 512 \mu_0$, so use of estimate (4.14*c*) of Λ in Raychaudhuri's equation would show that $l^{\bullet\bullet} < 0$ at that and all earlier times.

Thus any of the conditions i) $\Lambda \leqslant 0$, ii) $q_0 > 0$, or iii) the existence of black-body radiation indicating a previous time when matter and radiation were in equilibrium, would imply that the radius function increased monotonically from a singularity a finite time previously. Two consequences that would follow are

a) there would be a minimum angular diameter observed for sources of the same metric size (this follows from the results of Subsect. **6**˙4.4), and

b) at an early state there would be a ionized plasma filling the universe (cf. Sect. **3**).

The existence of the plasma implies a cut-off to optical and radio observations; this must occur by a red-shift ~ 1000, but could possibly happen by as low a red-shift as 7, in a high-density universe. Similarly, at early times the universe would be opaque to neutrinos; in fact there would be a maximum red-shift for observations by each other kind of signal.

The existence of minimum angular diameters implies that there are 2-surfaces on which the past null geodesics generating the observer's light cone are reconverging (see Fig. 9); these 2-surfaces are « closed trapped surfaces » in the sense defined by Penrose [131]. In fact it can be shown that the reconvergence will occur before the optical depth is unity (Hawking and Ellis [131*a*]).

The real universe is not exactly a Robertson-Walker universe. Since however we can assume it is very like a Robertson-Walker universe at least up to the time of last scattering, we expect the qualitative features *a*), *b*) to remain true in the real universe. Figure 9 illustrates these features; if $q_0 > 0$ we can estimate the time to the surface of last scattering as being of the order of, but less than, $1/H_0$ (cf. Subsect. **5**˙1.1).

Perhaps the most important feature is the existence of the closed trapped surface. Hawking and Penrose [132] have shown i) the existence of this surface, ii) the strong causality condition (Subsect. 7'1. above), iii) the energy condition: for every timelike vector V^a at each point, $R_{ab} V^a V^b \geq 0$ (in the case of a perfect fluid and $\Lambda = 0$, this is equivalent to $\mu + p \geq 0$, $\mu + 3p \geq 0$, cf. (3.8)), together with a « generality condition » which would certainly be satisfied in the real universe (*), imply that there must be a singularity in the universe. We can therefore deduce, on the basis of reasonable physical assumptions, that the inhomogeneity and anisotropy of the real universe is una-

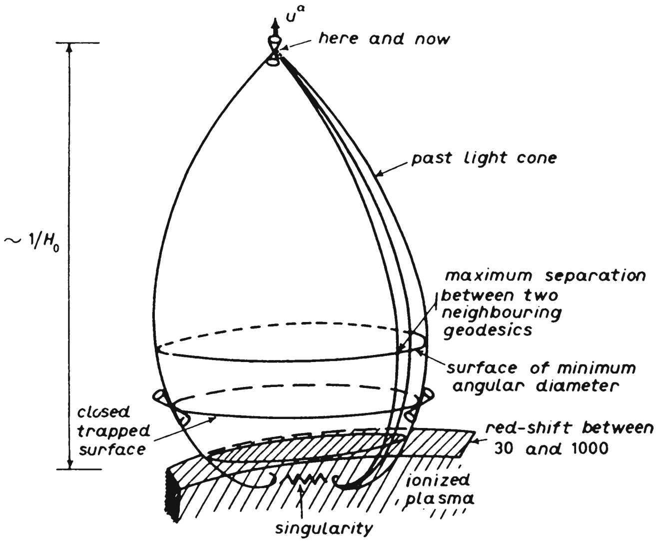


Fig. 9. – A closed, trapped surface in our past light cone. We cannot observe by optical or radio telescopes beyond the surface of ionization; this surface lies beyond the surface of minimum angular diameter. Numbers are for a low-density universe with $\Lambda = 0$.

ble to prevent the existence of a singularity in our past (**) if $\Lambda \leq 0$, and unlikely to prevent it even if $\Lambda > 0$. To postulate a « bounce » in an oscillating universe means abandoning either general relativity (and similar theories such as the Brans-Dicke theory, in which gravity is always attractive) or the energy condition (which we can do only by introducing large negative pressures or densities, if $\Lambda \leq 0$) or the causality condition (but in general, a violation of causality cannot prevent a singularity occurring (HAWKING [133])).

(*) For example, it is fulfilled if the universe is filled with a fluid with $\mu > 0$.
(**) To show it is in the *past*, we have to show that all the past timelike geodesics from the observer also reconverge in a compact region; this we can do (HAWKING and ELLIS [131*a*]). (Then we no longer need the « generality condition ».)

7·4. *The early universe.* – While we have concluded that the universe is very like a Robertson-Walker universe since the decoupling of matter and radiation, we have no direct observational evidence about earlier times, and so cannot derive the same conclusion for these times. In fact we may suspect that the universe was *not* very like a Robertson-Walker universe at early times, for the following reasons.

a) There must have been fluctuations larger than purely thermal fluctuations present in the universe at times soon after the decoupling of matter and radiation, in order that galaxy formation could have taken place (cf. the seminars by SALPETER and REES). Allowing for the damping effect of viscosity when photons decoupled and earlier when neutrinos decoupled, there must have been very large fluctuations early on in order to provide sufficiently large fluctuations later.

b) There are particle horizons (RINDLER [134]) in nonempty Robertson-Walker models which develop from a singularity (since this is true for all such models with $\mu > 0$, $p \geqslant 0$). This means that at early times in these models each particle was causally connected only to a small number of other particles; we can now observe radiation from parts of the universe which, if the early universe was very like a Robertson-Walker universe, were not causally connected to each other at the time of emission of the radiation. However both the physical state and the physical parameters of the matter in these regions appear to be very similar, despite their having been unable to interact with each other. This seems a rather implausible situation (*). By contrast, MISNER [28, 29] has suggested that the early universe could have been highly anisotropic and inhomogeneous; dissipative processes would tend to smooth out the anisotropy and inhomogeneity, thus providing a mechanism to explain the present high degree of homogeneity and isotropy. If the universe were highly anisotropic or inhomogeneous, the horizon structure could be very different from that in a Robertson-Walker universe (MISNER [135]) and so the previous problem need not arise.

c) One can argue (HAWKING and ELLIS [131*a*]) that solutions of the general-relativity field equations in which singularities have a spacelike character (as is the case in those Robertson-Walker universes in which particle horizons occur, see PENROSE [136]) are rather special; it may be that more general solutions have singularities with a timelike character (as in the Reissner-Nordström and Kerr solutions). This suggests that a realistic solution would be unlike a Robertson-Walker solution in the vicinity of its singularities (incidentally implying a different horizon structure, cf. *b*)).

(*) Unless one can characterize in detail a plausible creation process that must of necessity proceed in a uniform way.

d) Observations of a fairly uniform distribution of ^{4}He (about 25% by weight) suggests that many parts of the universe evolved through the helium formation phase with time scales similar to those in the Robertson-Walker universes (see PEEBLES [31], THORNE [59], HAWKING and TAYLER [137], WAGONER, FOWLER and HOYLE [33]). However very low helium abundances observed in the atmospheres of some halo stars (cf. BURBIDGE and SCIAMA in this volume) might indicate inhomogeneous element production resulting from large density inhomogeneities in the early universe.

7·5. *The nature of the singularity.* – The existence of a singularity in our past implies a breakdown of ordinary physical laws. However the theory has been extrapolated to circumstances where we would expect some quantized theory of gravitation to be a more appropriate description of space-time than general relativity; it is not yet clear whether a real physical singularity would be predicted by such a theory or not (cf. the discussion by MISNER [138]). Thus although the singularity theorem does not necessarily imply the existence of a real physical singularity (an end to space-time) in the universe, it may be taken to imply that conditions in our past were so extreme that general relativity was no longer a valid theory.

Although it appears that a singularity is predicted in every sufficiently large co-moving volume in the universe (HAWKING and ELLIS [131*a*]) this does not necessarily imply that all the matter in the universe experienced indefinitely high densities in the past. In fact if the singularities are timelike (cf. *c*) above), it is conceivable that a contracting phase in the universe could have changed to an expanding phase with most of the matter in the universe passing between isolated singularities. A consequence of this would be that no Cauchy surface would exist: complete knowledge of the state of the universe on any one space-section would be insufficient to determine its complete time development.

In any case, it is clear that the existence of a « singularity » in our past implies a breakdown of our ability to predict, at least on the basis of present theory; perhaps we should regard this as the essential feature of the singularity.

7·6. *The observable universe.* – Our optical and radio observations give us information about the distribution of matter and the curvature of space-time on our past light cone, back to the time when the universe was opaque. In principle we can obtain information about conditions on our past light cone at even earlier times (for example, by neutrino telescopes). Using the field equations, we can extrapolate inwards to find the geometry of space-time inside our light cone; in practice our observations of distant sources give us relatively little information so we can only extrapolate a short way in for these sources, while we can reasonably extrapolate the history of nearby sources for a rather greater time.

We can also use astrophysical, geophysical, and geological observations to give evidence about the history of the universe near the world line of our galaxy up to very early times. (For example, both the theory of element formation combined with observations of element abundances and the theory of stellar evolution based on observations of relatively nearby stars give restrictions on the physical conditions and the thermal evolution of the early universe near our galaxy.) Thus the part of the universe about which we can obtain reasonably good evidence is the shaded region in Fig. 10 (cf. HOYLE [80]).

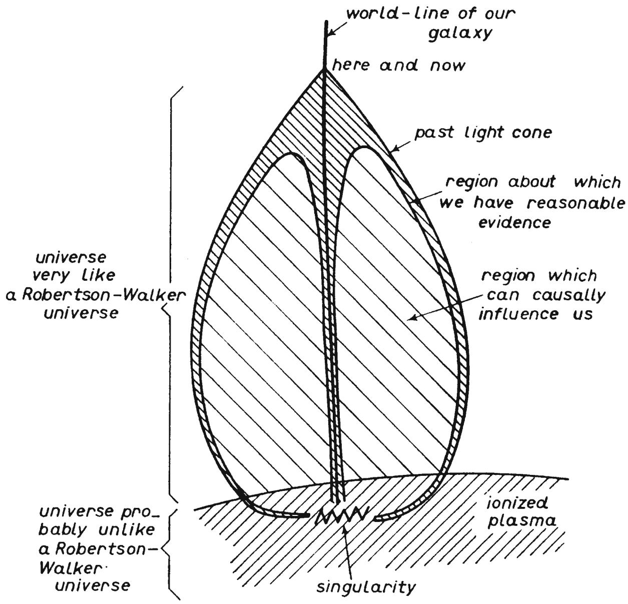


Fig. 10. – The observable universe, and the part of the universe we can observe in some detail. The part of our past since decoupling of matter and radiation is very like a Robertson-Walker universe; the hidden part is probably rather inhomogeneous in structure.

Using the Copernican principle we have concluded that the region of our past light cone since decoupling is very like a Robertson-Walker universe; we should therefore like to find a particular Robertson-Walker universe which is a good model for this space-time region by comparing the theoretical relations (Sect. 6) with the observations. (The model would be determined by μ_0 and $H_0^2 q_0$, and the epoch at which we observe it by H_0.) We might further aim to describe in detail the differences between the actual universe and the Robertson-Walker universe by comparing observations with theoretical relations in a perturbed Robertson-Walker universe, and to find the detailed distribution of matter and radiation in this perturbed universe. Finally we may aim to determine something of the early universe by indirect arguments based on the astrophysical evidence (see, for example Subsect. 7·3, 7·4, 7·5 above); in doing so we find useful powerful general theorems (such as those developed by PENROSE and HAWKING) and detailed studies of anisotropic and inhomogeneous cosmological models (cf. Sect. 5).

Having collated this information about the universe within our past light cone, the Copernican principle enables us to assume that conditions for other observers would be much the same. However we must use this principle with caution. Although we cannot necessarily assume particle horizons exist in the universe (since the early universe, about which we have no direct observational information, may be very different from a Robertson-Walker space-time), there will in general be *effective* particle horizons limiting the galaxies we can have observed since the universe became transparent to light and radiation (cf. SATO [139]) (but in a Lemaitre model ($\Lambda >$, $K > 0$) one might be able to see round the universe since decoupling occurred (*).) Further there are space-time events we have not yet observed and which no astronomer will be able to observe within, say, the next 10 000 years (although if there were no effective particle horizons, we could in principle obtain sufficient information to predict conditions at these events).

While we may in principle obtain *some* information about the space-time outside our light-cone by measuring the gravitational « Coulomb » field due to matter outside, it is as yet unclear precisely what we might be able to determine in this way. Thus statements about distant events outside our light cone are, in a fundamental way, unverifiable statements, and may be very misleading. (Consider for example the situation of an astronomer in a high-density region in a Lemaitre universe, the high density region recollapsing to a second singularity while most of the universe expands indefinitely.) Thus if we postulate a « cosmological principle » on philosophic grounds it may be impossible to disprove it; in this case, its scientific status is rather obscure.

(*) Personally, I support the view that $\Lambda = 0$ because otherwise it is a field which acts on everything but is not acted on by anything, which seems unreasonable (cf. RINDLER [5]).

Appendix

For a general matter tensor the left-hand sides of eqs. (4.21) are unchanged; the right-hand sides are:

$$(4.21a)\qquad := \tfrac{1}{3} h^{tb}\mu_{,b} - \tfrac{1}{2} h^t{}_c \pi^{cb}{}_{;b} - \tfrac{3}{2}\omega^t{}_b q^b + \tfrac{1}{2}\sigma^t{}_b q^b + \tfrac{1}{2}\pi^t{}_a \dot{u}^a - \tfrac{1}{3}\theta q^t ,$$

$$(4.21b)\qquad := \tfrac{1}{2}\sigma^{(t}{}_e \eta^{m)bef} u_b q_f - \tfrac{1}{2} h_c{}^{(t}\eta^{m)bef} u_b \pi^c{}_{e;f} + \tfrac{1}{2}(h^{mt}\omega_c q^c - 3\omega^{(m} q^{t)}) ,$$

$$(4.21c)\qquad := (\mu + p)\omega^t + \tfrac{1}{2}\eta^{tbef} u_b q_{[e;f]} + \tfrac{1}{2}\eta^{tbef} u_b \pi_{ec}(\omega^c{}_f + \sigma^c{}_f) ,$$

$$(4.21d)\qquad := -\tfrac{1}{2}(\mu + p)\sigma^{tm} - \dot{u}^{(t} q^{m)} - \tfrac{1}{2} h^t{}_a h^m{}_c q^{(a;c)} -$$

$$- \tfrac{1}{2} h^t{}_a h^m{}_c \dot{\pi}^{ac} - \tfrac{1}{2}\pi^{b(m}\omega_b{}^{t)} - \tfrac{1}{2}\pi^{b(m}\sigma_b{}^{t)} - \tfrac{1}{6}\pi^{tm}\theta + \tfrac{1}{6}(q^a{}_{;a} + \dot{u}_a q^a + \pi^{ab}\sigma_{ab}) h^{mt} .$$

References

[1] H. Bondi: *Cosmology* (Cambridge, 1960).
[2] G. C. McVittie: *Handbuch der Physik*, Vol. **53**, edited by S. Flügge (Berlin, 1959), p. 445.
[3] A. Einstein: *Sitzberichte der Preuss. Akad. d. Wissens.* (1917); see *The Principle of Relativity* (New York, 1923).
[4] J. Schouten: *Ricci Calculus* (Berlin, 1954).
[5] W. Rindler: *Essential Relativity* (New York, 1969).
[6] O. Heckmann and E. Schücking: *Handbuch der Physik*, edited by S. Flügge, Vol. **53** (Berlin, 1959), p. 489.
[7] O. Heckmann and E. Schücking: *Gravitation*, edited by L. Witten (New York, 1962), p. 438.
[8] Ya. B. Zel'dovich: *Adv. in Ast. and Ast.*, **3**, 242 (1965).
[9] W. Davidson and J. Narlikar: *Rep. Progr. Phys.*, **29**, 539 (1966).
[10] W. Rindler: *Phys. Today*, **20**, 23 (1967).
[11] E. Schücking: *Lectures in Applied Mathematics*, edited by J. Ehlers, Vol. **8**, (Providence, R.I., 1967), p. 217.
[12] A. R. Sandage: *Observatory*, **88**, 91 (1968).
[13] O. Heckmann and E. Schücking: *Zeits. f. Astrophys.*, **38**, 95 (1955).
[14] A. Trautman: *Perspectives in Geometry and Relativity* (Hlavaty Festschrift), edited by B. Hoffmann (1966), p. 413.
[15] J. L. Synge: *Relativity, the General Theory* (Amsterdam, 1960).
[16] A. Trautman: *Lectures on General Relativity* (Brandeis Summer Institute in Theoretical Physics, Vol. **1**), edited by S. Deser and K. Ford (1964), p. 7.
[17] J. Kristian and R. K. Sachs: *Astrophys. Journ.*, **143**, 379 (1966).
[18] K. C. Freeman: duplicated lecture notes, Cambridge University (1967).

[19] L. Woltjer: *Lectures in Applied Mathematics*, edited by J. Ehlers, Vol. **9** (Providence, R. I., 1967), p. 1.
[20] J. Serrin: *Handbuch der Physik*, edited by S. Flügge, Vol. **8/1** (Berlin, 1959), p. 125.
[21] G. K. Batchelor: *An Introduction to Fluid Dynamics* (Cambridge, 1967).
[22] J. Ehlers: *Abh. Akad. Wiss. und Lit. Mainz, Math. Nat. Kl.*, No. 11 (1961).
[23] J. M. Stewart: *Month. Not. Roy. Astr. Soc.*, **145**, 347 (1969).
[24] R. P. Geroch: Ph.D. Thesis (Princeton University, 1967).
[25] A. R. Curtis: *Proc. Roy. Soc.*, A **200**, 248 (1950).
[26] B. K. Harrison, K. S. Thorne, M. Wakano and J. A. Wheeler: *Gravitation Theory and Gravitational Collapse* (Chicago, 1965).
[27] R. F. Tooper: *Astrophys. Journ.*, **142**, 1541 (1965).
[28] C. W. Misner: *Astrophys. Journ.*, **151**, 431 (1968).
[29] C. W. Misner: *Nature*, **214**, 40 (1967).
[30] R. H. Dicke, P. J. E. Peebles, P. G. Roll and D. T. Wilkinson: *Astrophys. Journ.*, **142**, 414 (1965).
[31] P. J. E. Peebles: *Astrophys. Journ.*, **142**, 542 (1966).
[32] R. A. Alpher, G. Gamow and R. C. Herman: *Proc. Nat. Acad. Sci.*, **58**, 2179 (1967).
[33] R. Wagoner, W. Fowler and F. Hoyle: *Astrophys. Journ.*, **148**, 3 (1967).
[34] E. R. Harrison: *Physics Today*, **21**, 31 (1968).
[35] P. Jordan, J. Ehlers and W. Kundt: *Abh. Akad. Wiss. und Lit. Mainz, Mat. Nat. Kl.*, No. 2 (1960).
[36] O. Heckmann and E. Schücking: *Zeits. f. Astrophys.*, **40**, 81 (1956).
[37] F. A. E. Pirani: *Lectures in General Relativity* (Brandeis Summer Institute in Theoretical Physics, Vol. 1) edited by S. Deser and K. Ford (1964), p. 249.
[38] A. Raychaudhuri: *Phys. Rev.*, **98**, 1123 (1955).
[39] A. Raychaudhuri: *Zeits. f. Astrophys.*, **43**, 161 (1957).
[40] M. Ruderman: *Rep. Progr. Phys.*, **28**, 411 (1965).
[41] G. B. Field, M. J. Rees and D. W. Sciama: *Comments Astrophys. Space Sci.*, **1**, 187 (1969).
[42] R. B. Partridge and P. J. E. Peebles: *Astrophys. Journ.*, **147**, 868 (1967).
[43] K. Gödel: *Proc. Int. Cong. Math. (Camb., Mass.)*, **1**, 175 (1952).
[44] J. L. Synge: *Proc. Lond. Math. Soc.*, **43**, 376 (1937).
[45] W. Kundt and M. Trümper: *Abh. Akad. Wiss. und Lit. Mainz., Mat. Nat. Kl.*, No. 12 (1962).
[46] M. Trümper: *Zeits. f. Astrophys.*, **66**, 215 (1967).
[47] J. M. Anderson: *Principles of Relativity Physics* (New York, 1967).
[48] Y. Choquet-Bruhat: *Gravitation*, edited by L. Witten (New York, 1962), p. 130.
[49] L. C. Shepley and A. Taub: *Commun. Math. Phys.*, **5**, 237 (1967).
[50] H. P. Robertson: *Rev. Mod. Phys.*, **5**, 62 (1933).
[51] H. P. Robertson: *Pub. Astr. Soc. Pac.*, 82 (1955).
[52] R. Stabell and S. Refsdel: *Month. Mot. Roy. Astr. Soc.*, **132**, 379 (1966).
[53] S. Refsdal, R. Stabell and G. de Lange: *Mem. Roy. Astr. Soc.*, **71**, part 3 (1967).
[54] G. Tauber: *Journ. Math. Phys.*, **8**, 118 (1967).
[55] W. Rindler: *Astrophys. Journ.*, **157**, L148 (1969).
[56] K. Gödel: *Rev. Mod. Phys.*, **21**, 447 (1949).
[57] G. F. R. Ellis: *Journ. Math. Phys.*, **8**, 1171 (1967).

[58] G. F. R. Ellis and M. A. H. MacCallum: *Commun. Math. Phys.*, **12**, 108 (1969).
[59] K. S. Thorne: *Astrophys. Journ.*, **148**, 51 (1967).
[60] K. C. Jacobs: *Astrophys. Journ.*, **153**, 661 (1968).
[61] K. C. Jacobs: *Astrophys. Journ.*, **155**, 379 (1969).
[62] S. W. Hawking: *Astrophys. Journ.*, **145**, 544 (1966).
[63] P. Szekeres: *Journ. Math. Phys.*, **7**, 751 (1966).
[64] M. Trümper: *Journ. Math. Phys.*, **6**, 584 (1965).
[65] O. Heckmann: *Astronom. Journ.*, **66**, 205 (1961).
[66] J. V. Narlikar: *Month. Not. Roy. Astr. Soc.*, **126**, 203 (1963).
[67] J. M. Stewart and G. F. R. Ellis: *Journ. Math. Phys.*, **9**, 1072 (1968).
[68] J. Ehlers: *Zeits. f. Naturforsch.*, **22** a, 1328 (1967).
[69] A. Trautman: *Recent Developments in General Relativity* (Infeld Festschrift) (Warsaw, 1962), p. 459.
[70] R. Sachs: *Proc. Roy. Soc.*, A **264**, 309 (1961).
[71] E. Schrödinger: *Expanding Universes* (Cambridge, 1956).
[72] R. Penrose: *Perspectives in Geometry and Relativity* (Hlavaty Festschrift), edited by B. Hoffman (Bloomington, 1966), p. 259.
[73] B. Bertotti: *Proc. Roy. Soc.*, A **294**, 195 (1966).
[74] P. Jordan, J. Ehlers and R. Sachs: *Abh. Akad. Wiss. und Lit. Mainz, Mat. Nat. Kl.*, No. 1 (1961).
[75] W. Mattig: *Astr. Nach.*, **284**, 109 (1958).
[76] A. R. Sandage: *Journ. Soc. Indust. Appl. Math.*, **10**, 781 (1962).
[77] R. G. Roeder and G. C. McVittie: *Astrophys. Journ.*, **138**, 899 (1963).
[78] W. Mattig: *Astr. Nach.*, **285**, 1 (1959).
[79] A. R. Sandage: *Astrophys. Journ.*, **133**, 355 (1961).
[80] F. Hoyle: *Rendiconti S.I F.*, XX Corso (New York, 1960), p. 141.
[81] G. J. Whitrow and B. D. Yallop: *Month. Not. Roy. Astr. Soc.*, **127**, 301 (1964).
[82] E. R. Harrison: *Month. Not. Roy. Astr. Soc.*, **131**, 1 (1965).
[83] R. J. Gould and D. W. Sciama: *Astrophys. Journ.*, **140**, 1634 (1964).
[84] V. Petrosian, J. Bahcall and E. Salpeter: *Astrophys. Journ.*, **155**, L57 (1969).
[85] R. J. Gould and G. R. Burbidge: *Astrophys. Journ.*, **138**, 969 (1963).
[86] D. Clayton and J. Silk: *Astrophys. Journ.*, **158**, L43 (1969).
[87] A. Penzias and R. Wilson: *Astrophys. Journ.*, **156**, 799 (1969).
[88] R. Weymann: *Astrophys. Journ.*, **147**, 887 (1967).
[89] M. J. Rees and G. Setti: *Nature*, **219**, 127 (1968).
[90] F. J. Low and W. H. Tucker: *Phys. Rev. Lett.*, **21**, 1538 (1968).
[91] A. Wolfe and G. R. Burbidge: *Astrophys. Journ.*, **156**, 345 (1969).
[92] R. J. Gould: *Ann. Rev. Ast. and Ast.*, **6**, 195 (1968).
[93] G. Field and R. C. Henry: *Astrophys. Journ.*, **140**, 1002 (1964).
[94] G. Field: *Astrophys. Journ.*, **129**, 525 (1959).
[95] A. Penzias and E. Scott: *Astrophys. Journ.*, **153**, L7 (1969).
[96] J. Gunn and B. A. Peterson: *Astrophys. Journ.*, **142**, 1633 (1965).
[97] J. Bahcall and E. Salpeter: *Astrophys. Journ.*, **142**, 1677 (1965).
[98] D. W. Sciama: *Rendiconti S.I.F.*, XXV Corso (New York, 1966), p. 418.
[99] M. J. Rees: *Astrophys. Lett.*, **4**, 61 (1969).
[100] J. Bahcall and R. M. May: *Astrophys. Journ.*, **152**, 37 (1968).
[101] R. Wagoner: *Astrophys. Journ.*, **149**, 465 (1967).
[102] R. G. Roeder and R. Verreault: *Astrophys. Journ.*, **155**, 1047 (1969).
[103] J. Bahcall and P. J. E. Peebles: *Astrophys. Journ.*, **156**, L7 (1969).
[104] R. Weymann: *Astrophys. Journ.*, **145**, 560 (1966).

[105] Ya. B. Zel'dovich and R. Sunyaev: *Ast. and Space Sci.*, **4**, 301 (1969).
[106] P. J. E. Peebles: *Astrophys. Journ.*, **157**, 45 (1969).
[107] M. J. Rees: *Astrophys. Journ.*, **153**, L1 (1968).
[108] S. W. Hawking: *Month. Not. Roy. Astr. Soc.*, **142**, 129 (1969).
[109] R. Sachs and A. Wolfe: *Astrophys. Journ.*, **147**, 73 (1967).
[110] M. J. Rees and D. W. Sciama: *Nature*, **217**, 511 (1968).
[111] P. T. Saunders: *Month. Not. Roy. Astr. Soc.*, **141**, 427 (1968).
[112] P. T. Saunders: *Month. Not. Roy. Astr. Soc.*, **142**, 213 (1969).
[113] K. Tomita: *Progr. Theor. Phys.*, **40**, 264 (1968).
[114] M. A. H. MacCallum and G. F. R. Ellis: *Commun. Math. Phys.*, **19**, 31 (1970).
[115] R. Kantowski: *Astrophys. Journ.*, **155**, 89 (1969).
[116] D. Zipoy: *Phys. Rev.*, **142**, 825 (1966).
[117] J. Gunn: *Astrophys. Journ.*, **150**, 737 (1967).
[118] V. Petrosian and E. Salpeter: *Astrophys. Journ.*, **151**, 411 (1968).
[119] W. H. McCrea: *Zeits. f. Astrophys.*, **18**, 98 (1939).
[120] E. Newman and R. Penrose: *Journ. Math. Phys.*, **3**, 566 (1962).
[121] J. Stock and E. Schücking: *Astronom. Journ.*, **62**, 98 (1957).
[122] F. Hoyle: *Proc. Roy. Soc.*, A **308**, 1 (1969).
[123] E. H. Kronheimer and R. Penrose: *Proc. Camb. Phil. Soc.*, **63**, 481 (1967).
[124] S. W. Hawking: *Proc. Roy. Soc.*, A **294**, 511 (1966).
[125] J. Ehlers, P. Geren and R. K. Sachs: *Journ. Math. Phys.*, **9**, 1344 (1968).
[126] S. W. Hawking: *Month. Not. Roy. Astr. Soc.*, **142**, 129 (1969).
[127] A. Wolfe: *Astrophys. Journ.*, **156**, 803 (1969).
[128] D. W. Sciama: *Phys. Rev. Lett.*, **18**, 1065 (1967).
[129] J. M. Stewart and D. W. Sciama: *Nature*, **216**, 748 (1967).
[130] R. G. Hughes and M. S. Longair: *Month. Not. Roy. Astr. Soc.*, **135**, 131 (1967).
[131] R. Penrose: *Phys. Rev. Lett.*, **14**, 57 (1965).
[131*a*] S. W. Hawking and G. F. R. Ellis: *Astrophys. Journ.*, **152**, 25 (1968).
[132] S. W. Hawking and R. Penrose: *Proc. Roy. Soc.*, **314**, 529 (1970).
[133] S. W. Hawking: *Proc. Roy. Soc.*, A **300**, 187 (1967).
[134] W. Rindler: *Month. Not. Roy. Astr. Soc.*, **116**, 662 (1956).
[135] C. W. Misner: *Phys. Rev. Lett.*, **22**, 1071 (1969).
[136] R. Penrose: *Relativity, Groups and Topology* (Les Houches Summer School, 1963), edited by C. de Witt and B. de Witt (New York, 1964), p. 563.
[137] S. W. Hawking and R. J. Tayler: *Nature*, **209**, 1278 (1966).
[138] C. W. Misner: *Phys. Rev.*, **186**, 1319 (1969).
[139] H. Sato: *Progr. Theor. Phys.*, **40**, 781 (1968).