Relativistic Cosmology

H. P. Robertson, *Princeton University*

Table of Contents

I. General Theory

1. Introduction. Statement of the problem

THIS report deals with the attempts which have been made during the past decade and a half to solve the general problem of the structure of the universe as a whole. The recrudescence of such cosmological speculations is due to Einstein, who through his general theory of relativity advanced the view that the structure of the space-time continuum is determined causally by its material and energetic content, and who took the first step toward a solution. That the general theory of relativity is acting within its domain in attacking this problem can be doubted by no one who accepts it in its original form, perhaps on the basis of its well-known successes within the solar and galactic systems, but it must be borne in mind that in applying it or its later modifications to cosmology the choice of a particular one among the multitude of possibilities which it offers must be on the basis of additional facts supplemented, if these alone do not suffice, by such hypotheses as appeal to the general or philosophical predilections of the investigator. The observations which lend themselves to this purpose are meager indeed, consisting in the main in a relationship which has recently been found to exist between the distances and apparent radial velocities of extra-galactic nebulae, and in conclusions concerning their distribution within that portion of the universe which can be surveyed with the most powerful telescopes. Nevertheless, we hope to show that with their aid, under the guidance of a few seemingly natural assumptions and extrapolations, we can arrive at an intrinsically reasonable system of relativistic cosmology which is not in serious conflict with modern astrophysics; this, in brief, is the purpose of the present report.

The plan of the report is to present in the text a systematic account of the main results of these investigations in a manner in keeping with the purpose of these *Reviews*, avoiding as far as possible unfamiliar mathematical methods. But since the entire subject is based on certain aspects of the mathematical discipline known as differential geometry, some understanding of which is essential to a deeper appreciation of the problem, we have ventured to supplement the text by notes of a more mathematical nature in which the elements of the appropriate mathematical tools are presented. Also, in order not to interrupt the development of what may be called the *mechanics* of the universe, we have placed among the notes an account of a most promising line of attack, which has been initiated above all by Tolman, on the problem of the *thermodynamics* of the universe. The bibliography at the end contains only those articles which have been drawn upon heavily in the preparation of the report—indeed, the report attempts to present a unified digest of the results contained in these sources.

We therefore begin this report on relativistic cosmology with a very brief summary of the general theory of relativity, of which it is a natural offshoot. According to this theory the four-dimensional space-time of experience is a manifold in which space-like and time-like intervals are measured, on introducing general coordinates x^μ ($\mu=0, 1, 2, 3$), by an invariant Riemannian metric

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu, \qquad (1.1)$$

which measures the four-dimensional interval ds between the two events whose coordinates are x^μ and $x^\mu+dx^\mu$, respectively; repeated greek indices μ, ν, $\cdots$ imply throughout the summation of the term in which they occur over their range. The contingency of this metric on the physical content of space-time is expressed by means of the ten field equations

$$R_{\mu\nu} - g_{\mu\nu}(R/2-\lambda) = -\kappa T_{\mu\nu} \qquad (1.2)$$

for its ten coefficients $g_{\mu\nu}(x)$. The expression on the left, involving the contracted Riemann-Christoffel tensor $R_{\mu\nu}$ and its associated scalar R, is, except for the (so far arbitrary) cosmological constant λ, a function only of the $g_{\mu\nu}$ and their first and second derivatives with respect to the coordinates x^μ; this tensor has the extremely important property of having a divergence which vanishes identically.[1] (Einstein's constant of gravitation $\kappa = 2.07\times10^{-48}$ c.g.s. units is related to the Newtonian constant $G=6.66\times10^{-8}$ by the equation $\kappa = 8\pi G/c^4$, where $c=3.00\times10^{10}$ cm/sec. is the velocity of light *in vacuo*.) The stress-energy tensor $T_{\mu\nu}$ of the matter-energy field is, on the other hand, determined by the actual physical content of space-time, although it will in general contain the coefficients $g_{\mu\nu}$ as well; the vanishing of the four components

$$(1/g^{\frac{1}{2}})(\partial/\partial x^\nu)(g^{\frac{1}{2}}T_\mu^\nu) - \tfrac{1}{2}(\partial g_{\nu\rho}/\partial x^\mu)T^{\nu\rho} = 0 \qquad (1.3)$$

of its divergence yields the so-called conservation laws of energy and momentum. The theory of relativity allows a precise formulation of *Mach's principle*, according to which the inertial field is determined solely by the distribution of matter in the universe; in its more modern form it appears here in the weaker statement that the metric field (1.1) is causally determined to within a possible transformation of coordinates by the stress-energy tensor through the fundamental field equations (1.2).

The equations of motion of matter are contained implicitly in the field equations. In particular, it can be shown with the aid of (1.3) that the path of an isolated neutral test particle is a geodesic in space-time; its equations are, in terms of an arbitrary parameter σ,

$$\frac{d^2x^\mu}{d\sigma^2} + \left\{ {\mu \atop \nu\rho} \right\} \frac{dx^\nu}{d\sigma}\frac{dx^\rho}{d\sigma} = \frac{dx^\mu}{d\sigma}\frac{d^2s}{d\sigma^2}\Big/\frac{ds}{d\sigma}, \qquad (1.4)$$

where

$$(ds/d\sigma)^2 = g_{\mu\nu}(dx^\mu/d\sigma)(dx^\nu/d\sigma) \qquad (1.5)$$

and $\left\{ {\mu \atop \nu\rho} \right\}$ is the Christoffel symbol of the second kind formed from the $g_{\mu\nu}$ and their first derivatives.[2] These equations may be considerably simplified by choosing the parameter σ proportional to the proper distance s along the path, but we have written them in this more general form for later purposes and because in this form they are more readily applicable in determining the path of a beam of light, which is a geodesic for which

$$ds = 0. \qquad (1.6)$$

Again, the equations of motion of an ideal

[1] See note A, p. 83, for explicit definitions of these quantities. The notation is adapted from W. Pauli's article, *Relativitätstheorie*, in Encykl. Math. Wiss. V19, pp. 539–775 (Teubner, Leipsic 1921), and from L. P. Eisenhart's *Riemannian Geometry* (Princeton, 1926); the reader is referred to these treatises for more detail concerning the physical and mathematical aspects of the general theory of relativity and its underlying geometry.

[2] See note A, p. 83.

hydrodynamic fluid may be obtained from the divergence relations satisfied by the stress-energy tensor

$$T^{\mu\nu}=(\rho_{\rm m}+p_{\rm m}/c^2)v^\mu v^\nu-p_{\rm m}g^{\mu\nu} \tag{1.7}$$

for such a fluid, where the two scalars $\rho_{\rm m}$ and $p_{\rm m}$ are interpretable in terms of density and pressure, and the vector $v^\mu=dx^\mu/d\tau$—in which $d\tau=ds/c$ is the proper time interval measured along the world line of the fluid—defines the velocity field. This tensor (1.7) may also be made to describe the stress-energy tensor of isotropic radiation of density u by replacing $\rho_{\rm m}$, $p_{\rm m}$ by

$$\rho_{\rm r}=u/c^2, \quad p_{\rm r}=u/3, \quad v^\mu=\delta_0^\mu, \tag{1.8}$$

where $\delta_\nu^\mu=1$ if $\mu=\nu$ and $=0$ otherwise.

The three principal observational predictions of the general theory of relativity—the advance in the perihelion of mercury, the deflection of light passing through the field of the sun, and the *mass* red shift in the spectra of the sun and of the companion of Sirius—are obtained from a single rigorous solution of the field equations, Schwarzschild's solution for the field of a spherically symmetric body. Very considerable difficulties stand in the way of obtaining further rigorous solutions of astronomical importance—even the problem of two bodies is as yet unsolved. Approximations to the theory of n bodies have been given, but they are restricted in application by the limits of observational accuracy.

The problem with which we deal in this account is of a quite different nature; relativistic cosmology is an attempt to examine the structure of the universe *as a whole*, neglecting the local irregularities due to the agglomeration of matter into stars, or even into stellar systems. This rawest of all possible approximations may be considered as an attempt to set up an ideal structural background on which are to be superposed the local irregularities due to the actual distribution of matter and energy in the actual world. The empirical justification for such a treatment is to be found in the fact that recent astronomical research on extra-galactic nebulae, the most distant objects observed, indicates that although they are often observed to occur in large groups or clusters, yet on a still larger scale the density of distribution of these objects in space—the analogue on a tremendous scale of the density of molecules of a gas—is uniform throughout the observable universe, and that the relative peculiar motions of the nebulae in a given region, insofar as they are known, are, at most, of the order of a few hundreds of kilometers per second. Hubble[3] remarks that "to apparent magnitude about 16.7 . . . the number of nebulae to various limits of total magnitude vary directly with the volumes of space represented by the limits" and concludes that the density of nebulae is 9×10^{-18} nebulae per cubic parsec or, on introducing an estimate for the mass of an average nebula, 1.5×10^{-31} grams per cubic centimeter. In a more recent estimate he places the density of the luminous matter contained in nebulae at the somewhat higher figure

$$\rho_0=5\times10^{-31}\ \text{g/cc}. \tag{1.9}$$

The fact that more nebulae per square degree are observed at higher galactic latitudes than in the neighborhood of the equator is attributed to the obscuring effect of matter in our own galaxy.

These observational results—the constant density of observable matter in the large and the fact that known velocities can be considered as relatively small deviations from the mean for the region under consideration—lead us to consider that approximation in which the entire space is filled with a homogeneous and isotropic distribution of matter. But in order to specify more precisely this *spatial* uniformity we must first set up a space-time framework in terms of which we can express it.

2. Cosmic time. A priori specification of the line element

That we require the spatial distribution of matter in our highly idealized universe to be uniform implies the existence of a significant simultaneity, and would at first seem contrary to the postulates of the theory of relativity, according to which each observer refers the world to his own proper space and time. But this difficulty disappears on introducing, in accordance with the observed facts, the assumption that

[3] See bibliography, 1926.1, p. 321. See also Hubble and Humason, 1931.3, and Shapley and Ames, 1932.7.

there exists in each region of cosmic space-time a *mean* motion which represents the actual motions to within relatively small and unsystematic deviations. That some assumption concerning the natural state of motion of the matter in the universe is required in order to account for the facts has been emphasized above all by Weyl[4]; we take the above as the expression for the case in hand of his assumption (for the de Sitter universe) that the world lines of all matter belong to a pencil of geodesics which converges toward the past—the universe is a coherent whole rather than the fortuitous superposition of two or more incoherent parts.

We now set up a coordinate framework for space-time which satisfies the following, rather loosely expressed, conditions: (a) The lines of parameter $x_0 = t$ shall be geodesics, along which t measures proper time, so chosen that for a given time t_0 each shall represent, as closely as consistent with regularity and the condition below, the mean motion of the matter in its neighborhood (where, by this latter, we mean a region whose linear dimensions are large compared with the mean distance between nebulae), and (b) the space $t=$const., the coordinates of which are x^α $(\alpha = 1, 2, 3)$, shall be orthogonal to this (normal) congruence of geodesics. The possibility of thus introducing in a natural and significant way this *cosmic time* t we consider as guaranteed by Weyl's postulate, which is in turn a permissible extrapolation from the astronomical observations. The effect on the line element (1.1) of this resolution of space-time into space and time is expressed by the fact that it may now be written in the form

$$ds^2 = c^2dt^2 + g_{\alpha\beta}dx^\alpha dx^\beta, \qquad (2.1)$$
$$(\alpha, \beta \text{ summed over } 1, 2, 3),$$

where the $g_{\alpha\beta}(t, x)$ are such that $ds^2 < 0$ for any two neighboring points in the same space $t=$const. This partial reinstatement of absolute simultaneity into the actual world allows us to give a relatively precise formulation of the assumption that our ideal approximation to the actual world is spatially uniform. We demand (c) that to any observer (test body) in the idealized universe all purely *spatial* directions shall be fully equivalent in the sense that he shall be unable to distinguish between them by any intrinsic property of space-time and (d) that he shall similarly be unable to distinguish between his own observations and those of any *contemporary* observer. These two assumptions concerning the *spatial isotropy* and the *spatial homogeneity* of space-time are, in virtue of Mach's principle, fully equivalent to the corresponding assumptions concerning its material content which were obtained in the previous section by extrapolation of Hubble's data. Now the appropriate mathematical tool for the further specification of the line element (2.1) in accordance with these uniformity assumptions is the theory of groups of motions, as developed by Lie, Killing, Fubini and others[5]; with its aid we are, in fact, able to show that the line element must be of the form[6]

$$ds^2 = c^2dt^2 - R^2(t)du^2, \qquad (2.2)$$

where $R(t)$ is an arbitrary function and

$$du^2 = h_{\alpha\beta}(x_1x_2x_3)dx^\alpha dx^\beta \qquad (2.3)$$

defines a space of constant Riemannian curvature[7] k in which $du^2 > 0$ for any two distinct neighboring points. We can, without loss of generality, take $k = +1$, 0 or -1 according as it is >0, $=0$ or <0, respectively, for we need merely replace R^2 by R^2k in the first case and by $-R^2k$ in the third. The first of these cases, which will be found to be of most interest in our discussion of cosmology, characteristizes a space-time in which the three-spaces $t=$const. have a finite volume which is proportional to $R^3(t)$, and is in particular equal to $2\pi^2R^3$ or π^2R^3 according as (2.3) is interpreted as defining "spherical" or "elliptic" space, respectively.[8]

We consider in passing the further restriction placed on (2.2) by the *a priori* requirement that the structure of space-time be stationary. That an observer shall be unable to distinguish by any intrinsic property of space-time between its state

[4] See bibliography, 1923.2; 1930.10.

[5] For a brief account of the elements of this discipline and for references see note C, p. 84.

[6] See bibliography, Robertson, 1929.2. Also see note C, p. 85.

[7] For an elementary discussion of this concept see note B, p. 83.

[8] See note B, p. 84.

at any two arbitrarily chosen times t_1, t_2 is readily shown to imply one of the two possibilities[9]

$$k \text{ arbitrary}, \; R = \text{const.}, \qquad (2.4)$$

or

$$k=0, \quad R(t) = R_0 e^{ct/a}. \qquad (2.5)$$

The plausibility of this result will be more apparent in the sequel, for we shall find that the first of these two possibilities leads to the Einstein universe (§5) in which all matter is at rest, and the second to the empty de Sitter universe (§6).

In order to discuss the properties of the space-time universe as a whole or to set up the analogue of the Minkowski diagram for such a universe it is often convenient to consider it as a four-dimensional surface imbedded in a flat space of higher dimensionality. It can be shown that from this point of view the space defined locally by (2.2) has the same intrinsic geometry as a general hyper-surface of revolution in a flat five-space.[10] We shall in fact have occasion to resort to this representation in discussing the stationary universes of Einstein and of de Sitter in §§5, 6 below.

3. Physical interpretation of the line element

The above considerations led to the conclusion that any relativistic cosmology which considers the actual world as approximated by a spatially uniform background must be based on a line element of the form (2.2), involving an arbitrary function $R(t)$ and a constant k which may assume any of the three values $+1$, 0 or -1. We have now to examine what further restrictions are placed on the line element by the requirement that it satisfy the field equations (1.2) for some suitable choice of the stress-energy tensor $T_{\mu\nu}$. To do this we begin by computing the $T_{\mu\nu}$ defined by these field equations when the values of $g_{\mu\nu}$ given by (2.2) are substituted into the left-hand side; we find that the resulting stress-energy tensor may be written in the form

$$T^{\mu\nu} = (\rho + p/c^2)\delta_0^{\mu}\delta_0^{\nu} - pg^{\mu\nu}, \qquad (3.1)$$

where

$$\begin{aligned} \kappa\rho c^2 &= -\lambda + 3(k + R'^2/c^2)/R^2, \\ \kappa p &= \lambda - 2R''/Rc^2 - (k + R'^2/c^2)/R^2, \end{aligned} \qquad (3.2)$$

and the prime indicates differentiation with respect to t. The conservation equations (1.3), which are of course automatically satisfied by (3.1), reduce to the single relation

$$R\rho' + 3R'(\rho + p/c^2) = 0 \qquad (3.3)$$

between the quantities (3.2).

On comparing (3.1) with (1.7) we see that (2.2) may be interpreted as defining the field due to matter which is at rest ($v^{\mu} = \delta_0^{\mu}$) and whose density and pressure are given by (3.2); this result is immediately understandable in view of the fact that we introduced our coordinates x^{μ} in such a way that the material content of the universe is on the whole at rest with respect to them. We may, however, consider (3.1) as containing a contribution due to a field of isotropic radiation; if u is the energy density of this radiation field we have, on taking (1.8) into account,

$$\rho = \rho_m + u/c^2, \quad p = p_m + u/3, \qquad (3.4)$$

where ρ_m and p_m are the density and pressure of matter. Of these quantities certainly ρ_m and u are inherently non-negative, and we shall for the most of our report make the reasonable assumption that $p \geqq 0$.

The matter contained within a closed surface S, whose equation $S(x^{\alpha}) = 0$ is independent of t, can never cross over the boundary, and the rate at which radiation is escaping through S is just balanced by the rate at which it is entering. Denoting by $V(t)$ the volume in the space $t = \text{const.}$ enclosed by a surface S whose volume content is unity with respect to the auxiliary metric (2.3), and by $M(t)$ and $E(t)$ the total mass and energy (which includes the energy content of matter in accordance with the law $E_m = Mc^2$ of the theory of relativity) contained within it, we have

$$V = R^3, \quad M = \rho_m V, \quad E = \rho c^2 V = Mc^2 + uV. \qquad (3.5)$$

The conservation Eq. (3.3) now assumes the significant form

$$dE + pdV = 0. \qquad (3.6)$$

The general line element (2.2) is thus found to satisfy the field equations for a spatially homogeneous distribution of isotropic radiation and matter at rest, without restriction on the three arbitrary elements k, $R(t)$ and λ, except for such inequalities, such as $\rho \geqq 0$, as we may impose in

[9] See bibliography, Robertson, 1929.2, p. 825. See note C, p. 85, for a sketch of the proof.

[10] See bibliography, Robertson, 1929.2, p. 826. See note D, p. 86, for the explicit result.

order to insure a sensible interpretation. The further specification of the line element must therefore rest on assumptions concerning the finiteness of space, the density or pressure of matter and radiation, and the cosmological constant λ; the precise specification of the arbitrary function involved will in the main be accomplished by imposing assumptions concerning E or M, or both. For this purpose it will be found convenient to consider E as a function of R which is single-valued over a sufficiently restricted range of R; the first of Eqs. (3.2) may then be written, with the aid of (3.5),

$$c(t-t_0)=\pm\int_{R_0}^{R}\left(\frac{3x}{D(x,\lambda)}\right)dx, \qquad (3.7)$$

where

$$D(x,\lambda)=\kappa E(x)-3kx+\lambda x^3. \qquad (3.8)$$

The determination of R as a function of t is thus reduced to the quadrature (3.7) and the subsequent inversion of the resulting equation $t=t(R)$. But even in cases where E is not known explicitly as a function of R these results will be of considerable value in discussing the behavior of R as a function of t, as will be shown in §7 below.

The treatment which we have here given the general foundations of relativistic cosmology—which is based only on translating, with the aid of Mach's principle, into mathematical restrictions on the line element observational results concerning the uniform distribution of matter in large and their extrapolation—follows comparatively recent work of the author, as indicated in the references, and deviates from the historical development in the following particulars. The first treatment of the subject, that of Einstein,[11] dealt only with the stationary case (2.4) in which $k=+1$, $R=\text{const.}$, $p=0$ (cf. §5) and was followed shortly by de Sitter's treatment[12] of the alternative stationary possibility (2.5) in which $k=0$, $R=e^{ct/a}$, $p=0$ (§6). The first general treatment is that of Friedmann,[13] who arrived at the line element (2.2) for $k=+1$ by a combination of *a priori* homogeneity assumptions *and* the requirement that the matter-energy tensor be of the form (3.1) (for $p=0$), but his homogeneity requirements are weaker than those imposed in the above and as a result his derivation of (2.2) is incorrect. Friedmann discussed in some detail the various cases arising for $k=\pm1$ under the assumption $p=0$ (cf. §8). Subsequently Lemaître[14] wrote down the line element (2.2) as representing an Einstein world of variable radius, threw the conservation equation into the form (3.6), obtained $E(R)$ explicitly under the assumption $M=\text{const.}$ (see Eq. (9.1)) and discussed in detail a particular case (see Eq. (8.4)) contained in Friedmann's survey, suggesting it as an alternative to the de Sitter-Weyl explanation of the red shift in light from distant nebulae. The next general treatment was that outlined in the above account, into which is incorporated Weyl's coherency assumption, and which exhibits explicitly the full range of possibilities. This was followed shortly by an alternative derivation by Tolman[15] of the case $k=+1$, in which *a priori* homogeneity assumptions are supplemented by the requirement that the matter-energy tensor be of the form (1.7) where ρ and p are functions of t alone. Tolman[16] has also initiated an extension of the subject into the domain of thermodynamics, but in order not to interrupt the development on the basis of the hitherto accepted principles of the relativity theory, which are capable of dealing only with the mechanical aspects of cosmology, we have chosen to give a brief account of this most interesting extension in note F, p. 87.

4. Motion of particles and of light

Before considering the various special assumptions leading to the more precise specification of space-time we develop certain aspects of the geometry and kinematics of the general case. The Eqs. (1.4) of the geodesics, which represent the paths of particles and of beams of radiant energy, assume a particularly simple and useful form on choosing as the parameter σ the distance u measured along the projection of the geodesics in the auxiliary space (2.3). In the first place, three of the four Eqs. (1.4) become[17]

[11] See bibliography, 1917.1.

[12] See bibliography, 1917.2, 3.

[13] See bibliography, 1922.1; 1924.1.

[14] See bibliography, 1927.1.

[15] See bibliography, 1930.6.

[16] See bibliography, 1931.8–10; 1932.5–6.

[17] See note E, p. 87, for a derivation of these results.

$$\frac{d^2x^\alpha}{du^2}+\begin{Bmatrix}\alpha\\ \beta\gamma\end{Bmatrix}^*\frac{dx^\beta}{du}\frac{dx^\gamma}{du}=0,\quad (\alpha,\beta,\gamma=1,2,3),\quad (4.1)$$

where the asterisk on the Christoffel symbol indicates that it is computed from the coefficients of the auxiliary metric (2.3); one of these three equations may of course be replaced by the first integral

$$h_{\alpha\beta}(dx^\alpha/du)(dx^\beta/du)=1.\quad (4.2)$$

The remaining Eq. (1.4) (for $\mu=0$) may be replaced by the first integral

$$(R(du/dt))^2=c^2/(1+\gamma^2R^2);\quad (4.3)$$

the constant of integration γ has the value 0 for light and ∞ for a particle at rest. Eqs. (4.3) and (1.5) then determine the spatial distance u and the space-time interval s as functions of t by the quadratures

$$u=\pm c\int\frac{dt}{R(1+\gamma^2R^2)^{\frac{1}{2}}},\qquad s=\pm\gamma c\int\frac{Rdt}{(1+\gamma^2R^2)^{\frac{1}{2}}}.\quad (4.4)$$

These results show that the spatial projection of a geodesic is itself a geodesic of the space $t=$ const.—or, what amounts to the same, of the auxiliary space (2.3). The function $R(t)$ enters into the determination of the geodesics only by the first of the Eqs. (4.4), which tells us at what point u of its path the particle is to be found at time t.

In the idealization contemplated in relativistic cosmology the extra-galactic nebulae are represented at rest in the coordinate system t, x^α; hence if u is the (constant) parameter distance, measured with respect to the metric (2.3), between two such objects at P_0 and P_1, the space-time interval between them at time t is described by the "distance"

$$l(t)=R(t)u.\quad (4.5)$$

This distance is therefore in general a function of time, and we may speak of its rate of change

$$v=dl/dt=(1/R)(dR/dt)l\quad (4.6)$$

as the "apparent velocity" of P_1 relative to P_0; this apparent velocity of such an object is consequently proportional to its distance. We have now to examine the possibility of determining this velocity by means of observations on light emitted by the nebula at P_1 and received by an observer at P_0; this we do with the aid of a simple, practically corpuscular, theory which is justified in its essential predictions by a more rigorous investigation of von Laue[18] based on wave optics. The light emitted from P_1 at time t_1 is, in accordance with (4.4) for $\gamma=0$, received at P_0 at the time t_0 defined by

$$\int_{t_1}^{t_0}\frac{dt}{R(t)}=\frac{u}{c}.\quad (4.7)$$

Consequently the light which is emitted at P_1 in the time interval Δt_1 is received at P_0 in the interval $\Delta t_0=\Delta t_1\cdot R_0/R_1$, where R_1 and R_0 are the values of $R(t)$ at times t_1 and t_0; this will cause a change $\Delta\nu$ defined by

$$-\Delta\nu/\nu=(R_0-R_1)/R_0\quad (4.8)$$

in the frequency ν of the light. On expanding $R(t)$ and R_1 in a Taylor series about $t=t_0$ we find, with the aid of (4.5)–(4.8), that the shift $\Delta\lambda$ in the wave-length λ is given by

$$c(\Delta\lambda/\lambda)=v_0-(1/2c)(v_0^2+v_0'l)+\cdots,\quad (4.9)$$

correct to terms of order higher than the second. Hence to a first approximation, the velocity of recession attributed to the spiral nebulae because of the red shift exhibited in their spectra may in fact be equated to the apparent velocity (4.6) if we wish to interpret the observed shifts as due to this kinematical effect (but see §11 below for the logical position of this assumption). The work of Hubble and Humason[19] on nebulae with radial velocities ranging up to more than 20,000 km/sec. shows that the observed dependence of the Doppler velocity $c\Delta\lambda/\lambda$ on distance l is, in fact, linear within this range, making allowance for small deviations as peculiar motions, and allows us to take

$$h=(R'/R)_0=1.8\times10^{-17}\text{ sec.}^{-1}.\quad (4.10)$$

It is to be noted in passing that this approximate linearity of velocity of recession with distance

[18] See bibliography, 1931.4. Our analysis follows Lemaître, 1927.1.

[19] See bibliography, 1931.3. For critical discussion see bibliography, Oort, 1931.12.

was first predicted by H. Weyl[20] in 1923 on the basis of the more restricted cosmology of de Sitter, to be considered in §6 below. We are not as yet in a position to conclude anything from the observational data concerning the value of the second order term in (4.9).[21]

The above treatment of the red shift tacitly assumes that (4.7) has a solution—i.e., that the value of the integral for $t_0=\infty$ is greater than u/c. But this need not be the case, and we are thus led to the conclusion that there may exist events $P_1(t_1, x_1)$ of which an observer at $P_0(x_0)$ can never be aware; we will in this case be able to speak of the "radius $\bar{l}(t)$ of the *observable* universe at time t," which is defined by

$$\bar{l}(t)=R(t)\bar{u}(t) \quad \text{where} \quad \bar{u}(t)=c\int_t^\infty \frac{dt}{R(t)}. \qquad (4.11)$$

All events at time t of which a stationary observer can ever be aware must at that time lie at a distance $l<\bar{l}(t)$. An additional conclusion which may be drawn from this result is that if we are considering the case $k=+1$ in which space is spherical (elliptic), a beam of light or a particle of sufficiently high velocity starting from 0 at time t_1 will be able to circumnavigate space and return to 0 at time t_0 defined by (4.7) in which u is replaced by $2\pi(\pi)$, provided $\bar{u}(t_1)>2\pi(\pi)$.

A definition which is in closer accord with astronomical practice than (4.5) has been suggested by Tolman and by Whittaker.[22] Since the distances of extra-galactic objects are estimated with the aid of intensity measurements and it is assumed in practice that the intensity falls off inversely with the square of the distance, these authors propose as a measure of the distance a quantity which is proportional to the square root of the intensity of a standard source. For the case of spherical or elliptic space this quantity, which we denote by d, is found to be

$$d=R(t)\sin u, \qquad (4.12)$$

but this expression differs from (4.5) only in terms of third and higher orders in u and will therefore lead to an expression for $\Delta\lambda/\lambda$ which differs from (4.9) only in terms of order higher than the second.

II. Stationary Universes

5. The static Einstein universe

We consider first the stationary cosmology (2.4) in which $R=$const. (referred to in the following as of type E). This case is of considerable historical interest, as it was obtained by Einstein[23] in conjunction with his attempt to avoid the difficulties inherent in an infinite space filled with matter by introducing the cosmological constant λ into the field equations (1.2). We shall not review the arguments which led Einstein to this hypothesis, as he no longer considers them valid.[24]

The physical specification of this space-time is given by the matter-energy tensor (3.1), which is expressed in terms of the density ρ and the pressure p by

$$\kappa\rho c^2=-\lambda+3kR^2, \quad \kappa p=\lambda-k/R^2. \qquad (5.1)$$

If we demand that $p\geqq 0$, $\rho>0$ we see that, of the three possibilities for k, the only one which is permissible is $k=+1$, whence

$$\lambda=\kappa p+1/R^2, \quad \kappa\rho c^2=-\kappa p+2/R^2. \qquad (5.2)$$

Einstein further assumed that the density of radiant energy is negligible compared with the energy density of matter and that matter exerts no pressure; we then have

$$\lambda=1/R^2, \quad \kappa\rho c^2=2/R^2 \qquad (5.3)$$

and the total mass contained in the (spherical) universe is

$$2\pi^2 M=4\pi^2 R/\kappa c^2. \qquad (5.4)$$

Hubble's value (1.9) for the mean density of

[20] See bibliography, 1923.2.

[21] See bibliography, Tolman, 1930.7.

[22] See bibliography, Tolman, 1930.8; Whittaker, 1931.11. An explicit application of this definition to general cosmological spaces, leading to (4.12) for $k=1$, is given in note E, p. 87.

[23] See bibliography, 1917.1.

[24] See bibliography, Einstein, 1931.1, p. 236, where he characterizes the λ term as "theoretically unsatisfying"; however, this view does not seem to be shared by all relativists—cf. e.g., A. S. Eddington, *The Expanding Universe*, Proc. Phys. Soc. London **44**, 1–16 (1932), in particular p. 5. For an exhaustive review of the older arguments concerning the structure and physical content of the world as a whole see F. Selety, *Beiträge zum kosmologischen Problem*, Ann. d. Physik [4] **68**, 281–334 (1922).

luminous matter contained in the nebulae leads to a radius

$$R = 4.6 \times 10^{28} \text{ cm}; \qquad (5.5)$$

if we multiply this value of the density by a factor α (between 1 and 1000?) to allow for nonluminous matter R is decreased in the ratio $1 : \alpha^{\frac{1}{2}}$.

The Einstein universe has the amusing property of allowing light to circumnavigate space—in $2\pi R/c = 9.6 \times 10^{18}$ sec. or 3×10^{11} yrs.! But such a static universe is unsatisfactory for two reasons: In the first place it does not enable us to explain the motion of recession of the spiral nebulae, and in the second it is unstable, as was pointed out explicitly by Eddington[26]—a slight disturbance will lead to a continued contraction or expansion, as discussed more fully in §§7 and 10.

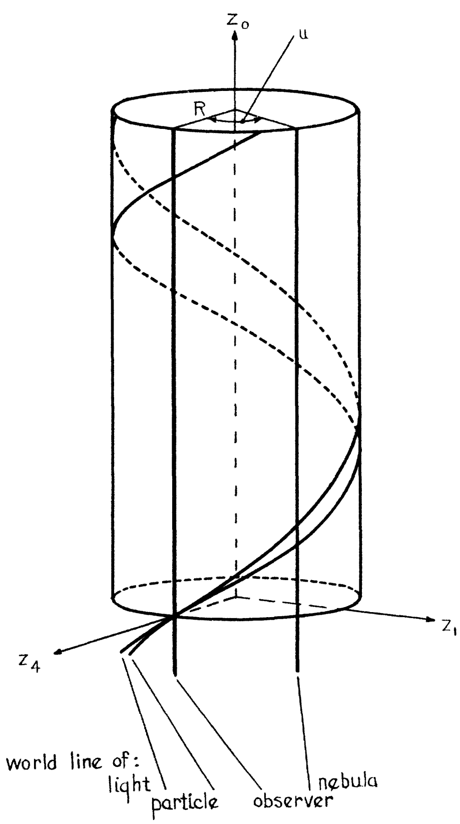

FIG. 1. The Einstein universe.[25]

[25] For the preparation of the figures, I am gratefully indebted to Dr. Edwin M. McMillan.

As shown in note D, the Einstein universe may be considered as a four-dimensional cylinder of radius R in a five-dimensional flat space, the axis of which is along the direction of cosmic time; this fact enables us to obtain a graphic representation of it as an ordinary cylinder of radius R in three-space on disregarding two of the spatial dimensions. This cylinder, which is the analogue of the Minkowski diagram of the special theory of relativity and which may be obtained from it on rolling it up into a cylinder of radius R in such a way that the world line of the observer becomes a generator of the cylinder, is illustrated in Fig. 1, showing the paths of the observer, particles at rest and in free motion, and light. On allowing $R \rightarrow \infty$ the Einstein universe degenerates into the space of special relativity (type SE) and the cylinder into the familiar Minkowski diagram.

6. The de Sitter universe

The second of the two stationary possibilities (which we refer to as of type S), suggested by Ehrenfest and investigated by de Sitter,[27] offers considerably more promise than Einstein's. On suitable choice of coordinates we may take the line element of this de Sitter universe in the form[28]

$$ds^2 = c^2dt^2 - e^{2ct/a}(dx_1^2 + dx_2^2 + dx_3^2). \qquad (6.1)$$

We then find from (3.2) that

$$\kappa\rho c^2 = -\lambda + 3/a^2, \quad \kappa p = \lambda - 3/a^2 \qquad (6.2)$$

and on requiring that neither ρ nor p be negative it is evident that they must both vanish and that $\lambda = 3/a^2$. We are thus dealing with that approximation in which the effect of matter on the underlying metric is neglected; the nebulae are to be considered as test particles whose existence

[26] See bibliography, 1930.1.

[27] See bibliography, 1917.2–3.

[28] See bibliography, Lemaître, 1925.1; Robertson, 1928.1; As shown by Weyl, 1930.10, these coordinates are the explicit analytical expression of the geometrical derivation in Weyl, 1923.2.

has no effect on the structure of the universe. Eq. (6.1) defines a four-space of constant Riemannian curvature $-1/a^2$; this case has therefore often been called the "spherical" universe.

Eqs. (4.7), (4.8) lead to the *rigorous* law

$$-\Delta\nu/\nu = l/a \qquad (6.3)$$

for the red shift, and on applying Hubble's and Humason's value (4.10) for $h = R'/R = c/a$ we find

$$a = 1.7 \times 10^{27} \text{ cm.} \qquad (6.4)$$

It is to be emphasized that this possibility of obtaining a unique Doppler shift in the de Sitter universe is directly attributable to Weyl's assumption that the world lines of matter constitute a pencil which converges toward the past; the many (impossibly low) determinations of the "radius" of the de Sitter universe made by Silberstein[29] between 1924 and the present are obtained from data on red *and violet* shifts in the spectra of stars, globular clusters and nebulae without the aid of any such assumption, and are therefore without the spirit of this report.

In the de Sitter universe *space* is unlimited, but objects at a distance greater than the radius $\bar{l} = a$ (4.11) of the observable universe will be completely isolated from an observer at O, as light from such objects can never reach O (as may also be seen immediately from (6.3)); in this sense space has an *apparent* constant radius a.

As shown explicitly in note D, de Sitter's spherical universe can be represented by a portion of a certain four-dimensional surface, which we may call a "pseudo-sphere," immersed in a flat five-space. On suppressing two of the spatial dimensions the situation is given graphically by Fig. 2, in which space-time is represented as that portion of an hyperboloid of one sheet which is bounded by two parallel generators L —the world lines of light which was emitted at time $t = -\infty$ from the vertex of Weyl's diverging pencil of geodesics. The observable universe of an observer O consists entirely of that portion of the diagram which lies within the two parallel generators L′ which represent the paths of light which approach the world line of O asymptotically in the future. The world line N of a nebula at rest in the t, x^α coordinate system is asymptotic to that of O in the past, and will at some time in the future cross one of the generators L′ which represent the limits of O's observable universe; this, together with the fact that de Sitter's original coordinates covers only this observable universe in such a way that his time coordinate has the value $+\infty$ on the lines L′, accounts for the illusion of a "mass horizon" which once

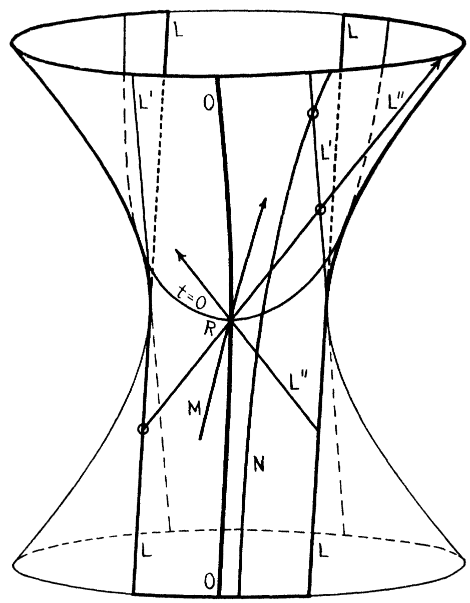

FIG. 2. The de Sitter universe.

seemed so puzzling. The "spaces" $t = \text{const.}$ are the parabolas in which the hyperboloid is cut by planes parallel to the plane determined by the two generators L. The two generators, such as L″, through any point P represent the paths of a beam of light sent out from or received at that point, and the world line M of any particle passing through P must lie between the positive directions of the two null-lines L″. The fact that all such generators or world lines seem, on following them backward in time, to have entered the universe merely means that they must at

[29] For an account of this work see L. Silberstein, *The Size of the Universe* (Oxford, 1930). But see review and criticism by the present author, Am. Math. Monthly **39**, 600–603 (1932).

some previous point have suffered an interaction with other matter which threw them off their natural course; as we follow them still further back they must, in accordance with Weyl's coherency postulate, approach asymptotically the world line of O.

We should, of course, expect that any universe which expands without limit will approach the empty de Sitter case, and that its ultimate fate is a state in which each physical unit—perhaps each nebula or intimate group of nebulae—is the only thing which exists within its own observable universe. In this connection it is to be noted, however, that the stationary form (6.1) of the de Sitter universe is not the only line element suitable for relativistic cosmology which describes an empty universe for $\lambda>0$, although in accordance with Mach's principle it must be possible to transform any other such line element into this stationary form.[30] In order to discuss all such possibilities we need only evaluate the integral (3.7) on setting $E=0$; by writing $\lambda=3/a^2$ as above, R is defined (for an expanding universe) by the equation

$$c(t-t_0)=a\int_{R_0}^{R}\frac{dx}{(x^2-ka^2)^{\frac{1}{2}}}. \tag{6.5}$$

The stationary form considered above is, of course, that in which $k=0$. The case $k=+1$ leads to the solution

$$R=a\cosh\left[c(t-\bar{t})/a\right], \tag{6.6}$$

which was found some years ago by Lanczos,[31] and the case $k=-1$ to the solution

$$R=a\sinh\left[c(t-\bar{t})/a\right]. \tag{6.7}$$

The existence of these alternative forms for the de Sitter universe emphasizes the necessity of augmenting the field equations by some assumption, such as Weyl's, concerning the natural state of matter; as remarked above, these line elements are mathematically equivalent to (6.1), but the physical universes which they describe are not stationary if t is interpreted as cosmic time—for example, the time required for light to travel a given proper distance l depends on the time at which it starts. Obviously both of these universes pass into the de Sitter universe as $t\rightarrow\infty$—for a formal proof we need merely replace $t-\bar{t}$ by $t+(a/c)\log(2A/a)$ and x^α by x^α/A, and allow $A\rightarrow\infty$.

The remaining possibilities for an empty universe lead only to the Minkowski world (for $\lambda=0$, $k=\pm1$) and to the form

$$R=b\sin\left[c(t-\bar{t})/b\right] \tag{6.8}$$

for $0>\lambda=-3/b^2$, $k=-1$.

[30] That this is in fact the case, is readily seen from the formulae (A. 7) for the Riemann-Christoffel tensor for a space in which both ρ and p vanish, for according to them all such spaces are of constant Riemannian curvature $-\lambda/3$ and may therefore be transformed into one another (cf. Eisenhart, *Riemannian Geometry*, p. 86).

[31] See bibliography, 1922.2.

III. Nonstationary Universes

7. Types of nonstationary universes in which $p\geqq0$

Preliminary to a more detailed discussion of the nonstationary solutions of the Eqs. (3.6), (3.7), which have received special attention at the hands of various authors, we examine the range of types which may arise under the physically plausible assumption that the total pressure p is never negative. This classification of types, which follows closely the scheme employed by the Russian mathematician, Friedmann, for the cases in which $p=0$ (cf. §8 below), is accomplished in terms of the three constants k, λ and R_0, the initial (i.e., present) value of the radius $R(t)$.

We first note that under the assumption $p\geqq0$ Eq. (3.6) implies that E is a monotonically decreasing (increasing) function of t during the entire duration of a phase of expansion (contraction); we here employ the term "monotonically decreasing" in the weakened sense of "never increasing," thus allowing for the possibility that E remains constant over any part of the range. Since furthermore $E\geqq0$ over the whole range, $E(R)$ must approach an asymptotic value $E_\infty\geqq0$ as $R\rightarrow\infty$, and in such a way that the integral

$$\int_{R_0}^{\infty}p(x)x^2dx$$

exists. On the other hand, as R decreases $E(R)$ may increase without limit to a singularity at some value of R; we take explicit account here only of those cases in which this singularity, if it exists, is at $R=0$, as the behavior of $R(t)$ in all other cases can be inferred immediately from its behavior in these. We may, of course, assume that $E(0)>0$, as $E(0)=0$ leads to the de Sitter universe considered in the previous section.

Eq. (3.7) or

$$R'^2=c^2D(R,\lambda)/3R \qquad (7.1)$$

tells us that the only admissible values of R for a given λ are those for which $D(R,\lambda)\geqq 0$, and that the only points at which R' may change sign are the zeros of $D(R,\lambda)$. We are thus led to examine the roots of the equation $D(R,\lambda)=0$, which is most conveniently done by considering the locus of these points (R,λ) in the $(R\geqq 0, \lambda)$ half-plane; since the behavior of the curve differs essentially in the two cases $k=+1$ and $k=0$ or -1 we treat these cases separately.

Case $k=+1$. Those cases in which space is finite offer by far the most varied and most interesting range of possibilities. The critical curve $D(R,\lambda)=0$ or

$$\lambda=3/R^2-\kappa E(R)/R^3 \qquad (7.2)$$

has as λ-extremals those points R for which

$$d\lambda/dR=-6/R^3+3\kappa E/R^4+3\kappa p/R=0, \qquad (7.3)$$

and the value of λ at such a point is

$$\lambda=1/R^2+\kappa p \quad (>0). \qquad (7.4)$$

From these results and the assumptions $E\geqq 0$, $p\geqq 0$ we can conclude the following points concerning the behavior of the critical curve, as indicated in *full line* in the accompanying Fig. 3. (a) The curve is asymptotic to the negative λ-axis and as R increases from zero it rises steadily until it crosses the R-axis, for by (7.4) it cannot have a horizontal tangent for any point at which $\lambda\leqq 0$; (b) it has at least one maximum $M(R_e, \lambda_e)$; and (c) it approaches the R-axis asymptotically as $3/R^2$ as $R\to\infty$.

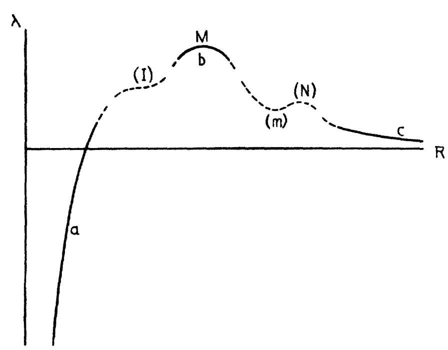

FIG. 3. Behavior of $D(R,\lambda)=0$ for $k=1$ on assuming $p\geqq 0$ (in full line; dotted line illustrates hypothetical possibilities considered in text).

The only (R,λ) points which can represent a state of the universe are those which lie above or on this critical curve. From the facts we have stated above we are able to predict without ambiguity the general behavior of a universe for which λ is greater than the maximum value λ_e assumed on the curve, or for which it is less than or equal to zero. The behavior for values of λ greater than zero but not greater than λ_e depends on the nature of the non-negative function $p(R)$, but we are, nevertheless, able to survey the various types which may here arise—and, as we shall see shortly, the introduction of a physically plausible assumption concerning $p(R)$ enables us to remove the ambiguity in this intermediate region as well. In describing the various possibilities which may arise we shall in general confine our attention to those which are in the expanding phase at the initial time $t=t_0$.

$\boxed{\lambda>\lambda_e}$ In this case the line $\lambda=$const. does not cut the critical curve, and consequently the point (R,λ) representing the universe moves out along the line with a velocity which approaches $(\lambda/3)^{\frac{1}{2}}cR$; the universe approaches the stationary de Sitter state (6.1) through the nonstationary Lanczos form (6.6). Looking backward in time, we might say that the universe "began" at that finite time

$$t_0=-\frac{1}{c}\int_0^{R_0}\left(\frac{3x}{D(x,\lambda)}\right)^{\frac{1}{2}}dx \qquad (7.5)$$

in the past when its radius was 0; we should of course not expect our idealization to hold back to such a time, but we might consider a singularity of $E(R)$ at some point $R>0$ as the "beginning." If we wish to follow its course back through such a singular state we find that at all times still further in the past the radius was decreasing. Such a universe, which expands monotonically in the future and which was in such a singular state as $R=0$ or that state R at which $E(R)$ has a singularity, we refer to as a *monotonic world of the first kind* and denote it by M_1.

$\boxed{\lambda \leqq 0}$ Here the point (R, λ) representing the state of the universe must be on the line $\lambda=$ const. between $R=0$ and that point R_1 in which this line cuts the critical curve $D(R, \lambda)=0$. The fate of such a universe is to expand until its radius attains the critical value R_1, at which point it begins to contract and continues to do so until it reaches the singular state $R=0$. If we care to follow it further we should expect to find that it will again expand, only to repeat the oscillation—which is strictly periodic in time if $E(R)$ is a single-valued function of R. Such a universe we refer to as *oscillating* and denote it by O; that all universes for which $k=+1$, $\lambda \leqq 0$ are of this oscillating type has been proved by Tolman and Ward.[32]

As stated above, the general features of the critical curve for $0<\lambda \leqq \lambda_e$ are not uniquely specified by the assumptions imposed so far in this section—the essential features for our present purpose being the number and distribution of horizontal tangents to the critical curve $D(R, \lambda)=0$. Let us consider briefly the conditions under which horizontal tangents in addition to that one at $M(R_e, \lambda_e)$ may exist. Obviously, the existence of any minimum such as m on the dotted line in Fig. 3 (which may, however, occur on either side of M) implies the existence of a subsequent maximum N; it is then readily seen from (7.4) that this situation can only exist if $p_N > p_m$. Again, the existence of a point of inflection, such as I in Fig. 3, or of horizontal tangents (including $\lambda=\lambda_e$) possessing any higher order of tangency to the curve, implies, as can be seen by differentiating (7.3), that at such a point $dp/dR>0$. Hence if we impose the additional physically plausible assumption that the pressure can never increase with the radius we automatically exclude the possibility of any horizontal tangent other than the simple tangent $\lambda=\lambda_e$ through M, and the curve is of the general form indicated in Fig. 4. We therefore consider in detail the behavior of any universe in which $dp/dR \leqq 0$ before returning to a brief survey of the more general situation.

[32] See bibliography, 1932.6.

$\boxed{0<\lambda<\lambda_e,\ dp/dR \leqq 0}$ Here $D(R, \lambda)$ has two positive roots $R_1(\lambda) < R_2(\lambda)$, between which D is negative; two cases may therefore arise for each value of λ, according as $R_0 < R_1$ or $R_0 > R_2$. The first of these possibilities obviously leads to a universe of the type O which expands until $R=R_1$ and subsequently contracts to the singular state; whether it passes through the singular state only to repeat the process as in those cases for which $\lambda \leqq 0$ (as it will if E is a single-valued function of R) or whether it then becomes a universe of type M_1 (or A_1 as described below) depends on whether λ is less than the new λ_e or not. The second case, in which $R_0 > R_2$, continues to expand monotonically as those of type M_1. The essential difference between this type, which we refer to as a *monotonic world of the second kind* M_2, and M_1, is that at a finite time

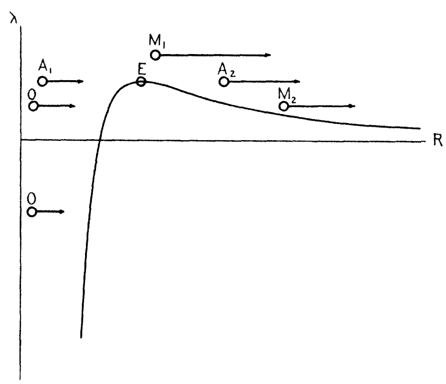

FIG. 4. Types of universes for $k=1$ in which $dp/dR \leqq 0$.

$$t_0 = \int_{R_2}^{R_0} \left(\frac{3x}{D(x, \lambda)} \right)^{\frac{1}{2}} dx \qquad (7.6)$$

in the past it assumed a minimum radius $R=R_2$; on following it back through this state we find that it was at all previous time monotonically contracting.

$\boxed{\lambda=\lambda_e,\ dp/dR \leqq 0}$ The case in which λ is just equal to the maximum value λ_e assumed on the critical curve is the most interesting of all. $D(R, \lambda_e)$ has a double root at $R=R_e$ and the radius may therefore remain at this critical value; this is, in fact, the general form (5.2) of the static Einstein universe discussed in §5 above. Such an equilibrium is of course unstable,[33] as is readily seen from Fig. 4. If $R_0 < R_e$ the radius increases monotonically, approaching the critical value $R = R_e$ in such a way that as $t \to \infty$, $R_e - R \to 0$ asymptotically as e^{-Ct}; we refer to such a universe as an *asymptotic world of the first kind* and denote it by A_1. Finally, if $R_0 > R_e$ the universe continues to expand monotonically without limit into the de Sitter universe as in types M_1 and M_2, but it differs from them markedly with respect to its behavior in the past, for it has been expanding forever; its deviation $R - R_0$ from the equilibrium radius is in the earlier stages of the form e^{Ct}, from which time may be said to become "logarithmically infinite" in the past, of which more anon. This unique type A_2, the *asymptotic world of the second kind*, may be considered as having originated in an Einstein universe which began to expand because of perturbations (cf. §10 below), and is the nonstationary transition stage between the two stationary types E and S.

Returning to the general case by dropping the assumption that p is a monotonically decreasing function of R, we see that for all values of λ greater than zero but not exceeding λ_e there exists a universe which is either of type A_1 or O, depending on whether $\lambda =$ const. is tangent to the critical curve at its first contact with it or not. Similarly there exists for each such value a universe which is either of type A_2 or M_2, depending on whether $\lambda =$ const. is tangent to the curve at its last point of contact or not. If in these asymptotic cases the tangency is of order $n > 2$ the deviation from the (unstable) equilibrium state is measured by $(\pm t)^{2/2-n}$ instead of by the exponential $e^{\mp Ct}$ as in the case of simple tangency $n = 2$, thus replacing the "logarithmic infinity" of t by an infinity of higher order. If the critical curve possesses a true minimum (as at m in Fig. 3) there exists a stable Einstein universe E corresponding to this minimum value; for each value of λ not too much greater there exists an oscillating universe neither limit of which is the singular state and which is strictly periodic if $E(R)$ is a single-valued function of its argument.[34] As λ increases still more this oscillating universe goes over into one which is of the asymptotic type at either or both of its limits. What happens as it increases beyond this value depends on the nature of the curve in more distant regions, but the total range of possibilities is completely covered by the preceding analysis.

Cases $k = 0, -1$. The cases in which space is infinite are treated together, for the types of behavior which such universes may exhibit and their dependence on λ are the same in both. Subject only to the assumption $p \geqq 0$, the curves $D(R, \lambda) = 0$, or

$$\lambda = 3k/R^2 - \kappa E/R^3, \tag{7.7}$$

are asymptotic to the negative λ-axis and to the positive R-axis, and have throughout a positive

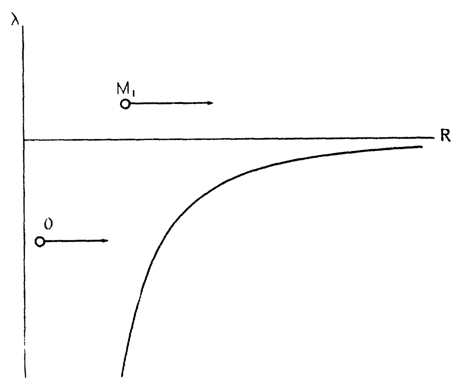

FIG. 5. $D(R, \lambda) = 0$ for $k = 0, -1$.

slope (Fig. 5). Hence but two types of universes arise, depending on the sign of λ.

$\boxed{\lambda \geqq 0}$ The line $\lambda =$ const. does not cut the critical curve, and consequently R increases monotonically without limit, giving rise to a universe of the type M_1 which approaches the stationary de Sitter world S—directly if $k = 0$ and through the intermediate nonstationary form (6.7) if $k = -1$. Note, however, that in case $\lambda = 0$ this ultimate form is that of Minkowski—a de Sitter world in which $a = \infty$.

[33] An explicit analytical proof of this result has been given by Eddington, 1930.1 (see bibliography), for the Lemaître case of $p = 0$.

[34] As shown by Tolman, 1931.10, p. 1764 (see bibliography), the pressure at the upper limit of such a universe exceeds that at its (nonsingular) lower limit.

$\boxed{\lambda<0}$ Here the line $\lambda=$const. cuts the critical curve in a point R_1, leading to an oscillating world (type O).

Our survey of all cosmologies in which $p\geqq 0$ leads to the conclusion that the only type (A_2) in which the universe has been expanding for all time in the past is possible only if $k=+1$, $\lambda=\lambda_e>0$; while we are hardly prepared to maintain that it is the only type fulfilling the requirements for a system of relativistic cosmology, we can say that unless we wish to go beyond the idealization contemplated so far in this report it would seem to be the only one which is consistent with the spirit of Weyl's postulate concerning the coherency of all matter in the actual universe—unless we wish to assume that the world "began" at a definite time some 10^{10} years in the past. We shall also find additional evidence for this view in the fact that the time scales to which the other possibilities lead in the more specific models considered in the following section, seem difficult to reconcile with modern astronomical theories; as we shall see in the following sections even this case has been objected to because of its time scale. We have collected in Table I the results of this section

TABLE I. *Types of universes in which* $E>0$, $p\geqq 0$, $dp/dR\leqq 0$.

$\lambda \backslash k$		1	0, −1
>0	$>\lambda_e$	M_1	M_1
	λ_e	A_1 E A_2	
	$<\lambda_e$	O M_2	
0		O	M_1
<0		O	O

for universes in which p is a monotonically decreasing non-negative function of R; the additional possibilities obtained on the weaker assumption that $p\geqq 0$, upon which we touched briefly in the above, affect only the compartments $k=1$, $0<\lambda\leqq\lambda_e$. It is to be noted that Einstein's recent proposal to set $\lambda=0$ would restrict us to the oscillating universes (type O) and the monotonic universes of the first kind (type M_1), both of which originate in the singular state $R=0$.

8. Universes in which energy is conserved

Having set forth in the preceding section a general survey of the possible types of universes we return to a more nearly historical account of the specific models which have been proposed. Soon after the appearance of Einstein's and de Sitter's solutions it was generally conceded that no further stationary universes could exist, although no satisfactory proof of this fact was given until much later.[35] Accordingly, in 1922, Friedmann[36] set himself the problem of deriving the most general line element, regardless of the condition that it be stationary, and although his derivation was faulty he succeeded in obtaining the fundamental form (2.2) for the cases $k=\pm 1$; he apparently did not entertain the possibility $k=0$, but we shall nevertheless consider it as falling under his general class of solutions in the following discussion.

From the beginning Friedmann restricted himself to universes in which the total pressure p vanishes; in accordance with (3.6) this is equivalent to the assumption that the total energy contained in a volume whose boundaries are fixed with respect to the spatial coordinates x^α is rigorously conserved. Support for the contention that this represents a good approximation to the actual world is found in the fact that the density u of radiant energy, which we consider largely responsible for the pressure p, is estimated to be but a small fraction of the energy density $\rho_m c^2$ due to matter. Friedmann then analyzed the various possibilities which could arise under these conditions; the essential points of his analysis are in fact used as a basis for the more general treatment given in the preceding section. Since $dp/dR=0$ the critical curve is of the form shown in Fig. 4 (which was in fact drawn for just this case $E=$const.); it cuts the R-axis at $R=\kappa E/3$ and is maximal at

$$R_e=\kappa E/2,\quad \lambda_e=1/R_e^2=(2/\kappa E)^2. \qquad (8.1)$$

For general values of λ the radius R is expressible in terms of t by means of elliptic functions, but

[35] See bibliography, Robertson, 1929.2, p. 825. Friedmann, 1922.1, and Tolman, 1929.3, had previously published proofs of the fact that no additional *static* solutions satisfying the general requirements for a system of cosmology could exist.

[36] See bibliography, 1922.1; 1924.1.

for some particular values of λ these reduce to elementary functions. Because of the interest which some of these latter have attracted of late we devote the remainder of this section to an explicit determination of them; their place among the other types of energy-conserving universes is shown in Table II.

TABLE II. *Universes in which energy is conserved.*

k	E	λ	R_0	Type	Remarks
$+1$	0	>0	Arb.	S	Lanczos' form (6.6) of S.
	>0	$>\lambda_e$	Arb.	M_1	$\rightarrow$(6.6)$\rightarrow$(6.1).
		λ_e	$<R_e$	A_1	$\rightarrow$E(5.3).
			R_e	E	Einstein universe [§5].
			$>R_e$	A_2	Lemaître's case (8.4).
		$<\lambda_e$ >0	$>R_2$	M_2	$\rightarrow$(6.6)$\rightarrow$(6.1).
		$<\lambda_e$	$<R_1$	O	$\lambda=0$ Einstein's case (8.6).
0	0	>0	Arb.	S	Stationary form (6.1) of de Sitter universe.
		0	Arb.	SE	Minkowski universe.
	>0	>0	Arb.	M_1	$\rightarrow$(6.1).
		0	Arb.	M_1	Einstein's and de Sitter's case (8.8).
		<0	$<R_1$	O	
-1	0	>0	Arb.	S	Form (6.7) of S.
		0	Arb.	SE	Minkowski space $R=ct$.
		<0	$<R_1$	S	Form (6.8).
	>0	>0	Arb.	M_1	$\rightarrow$(6.7).
		0	Arb.	M_1	$\rightarrow$Minkowski space SE.
		<0	$<R_1$	O	

Lemaître's case $k=1$, $\lambda=\lambda_e$. Of all Friedmann's worlds that one of type A_2, for which $k=+1$, $\lambda=\lambda_e$, $R_0>R_e$, is of most interest, for, as indicated by Friedmann, it is only when we allow λ to approach this critical value we get a universe with a limitless time scale. Five years after the appearance of Friedmann's first paper the Belgian mathematician, Lemaître,[37] set up the line element (2.2) for $k=1$ as "an Einstein universe where the radius of space or of the universe is allowed to vary in an arbitrary way" and, after drawing certain general conclusions which are discussed in other sections of this report, confined himself to the discussion of this most interesting case.

[37] See bibliography, 1927.1.

The cubic $D(R, \lambda_e)$ has a double root at $R=R_e$, and the integral (3.7) or

$$c(t-t_0)=R_e\int_{R_0}^{R}\left(\frac{3x}{x+2R_e}\right)^{\frac{1}{2}}\frac{dx}{x-R_e} \tag{8.2}$$

is therefore expressible in terms of elementary functions; the indefinite integral is

$$\begin{aligned}F(x)=2\cdot 3^{\frac{1}{2}}\log\left[x^{\frac{1}{2}}+(x+2R_e)^{\frac{1}{2}}\right]\\ +\log\left[(3x)^{\frac{1}{2}}-(x+2R_e)^{\frac{1}{2}}\right]\\ -\log\left[(3x)^{\frac{1}{2}}+(x+2R_e)^{\frac{1}{2}}\right]\end{aligned} \tag{8.3}$$

and consequently

$$c(t-t_0)=R_e[F(R)-F(R_0)]. \tag{8.4}$$

In order to apply this idealization to the actual universe it is only necessary to determine two constants, the equilibrium radius R_e and the present radius R_0, from empirical data. This we are able to do from Hubble's value (4.11) for the constant of proportionality h in the velocity-distance relation and from his estimate (1.9) for the density of the luminous matter which is contained in nebulae, by means of the equations

$$\begin{aligned}\left(\frac{R'}{R}\right)_0=\frac{c(R_0-R_e)}{R_0R_e}\left(\frac{R_0+2R_e}{3R_0}\right)^{\frac{1}{2}},\\ R_e=\kappa E/2=\kappa\rho c^2R_0^3/2\end{aligned} \tag{8.5}$$

obtained from (7.1) or (8.2), (3.5) and (8.1). Since, as Hubble points out, ρ may be considerably in excess of the value 5×10^{-31} because of the existence of nonluminous or scattered matter, we have set $\rho=\alpha 5\times10^{-31}$ and computed the resulting values of R_e, R_0 for various values of α ranging up to one thousand; the results of this computation are collected in Table III. The equilibrium radius R_e, which is

TABLE III. *R_e and R_0 for various densities $\rho=\alpha 5\times10^{-31}$.*

α	$R_e\times10^{-26}$	$R_0\times10^{-26}$	R_0/R_e
0.1	9.60	274	28.6
1	9.54	127	13.3
10	9.29	58.4	6.29
100	8.37	26.2	3.13
1000	6.15	11.0	1.78

of the order of 10^{27} cm, is found to be surprisingly insensitive to density in the range considered, whereas the present radius R_0 shows a considerable variation. The dependence of R/R_e on ct/R_e, which is independent of the particular constants involved, is represented in Fig. 6; the

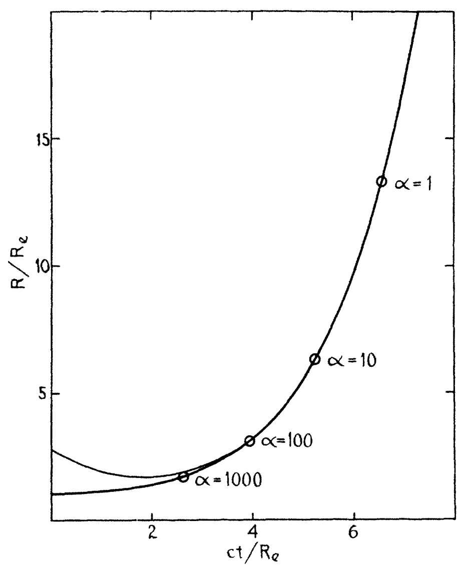

FIG. 6. Lemaître's case A_2 of Friedmann's universes.

points on the curve show the present state of the universe for the various assumed values of α from one to one thousand. The lighter curve shows Lanczos' nonstationary form (6.6) of the de Sitter universe, to which Lemaître's case of Friedmann's worlds is asymptotic; the two curves agree to within less than one percent at all points above $R/R_e=3$.

Einstein's case $k=1$, $\lambda=0$. This periodic universe, to which Friedmann referred explicitly in order to obtain some estimate of orders of magnitude, has recently been investigated in detail by Einstein.[38] The fundamental Eq. (3.7) is in this case readily integrated; in parametric form the solution is the cycloid

$$2ct=R_1(\varphi-\sin\varphi),\quad 2R=R_1(1-\cos\varphi) \tag{8.6}$$

[38] See bibliography, 1931.1.

generated by a circle whose radius is one-half the maximum radius $R_1=\kappa E/3$ of the universe. On applying the observational data on red shift and density as in the case considered above we find that the present and maximum radii are

$$R_0=\frac{10^{28}}{(0.031\alpha-36)^{\frac{1}{2}}}\quad R_1=\frac{0.031\alpha 10^{28}}{(0.031\alpha-36)^{3/2}}, \tag{8.7}$$

whence the solution is inapplicable unless the density of matter in the actual universe is in excess of one thousand times that estimated by Hubble for matter observed in nebulae. The present radius R_0, or R_1 and hence the period $T=\pi R_1/c$, is very sensitive to variations in the density; following Einstein in assuming that R_1-R_0 is of the same order as R_0 we may take $\alpha\sim 2000$ ($\rho\sim 10^{-27}$ g/cc), $R_0\sim 2\times 10^{27}$ cm and $T\sim 5\times 10^{17}$ sec. or 1.7×10^{10} yrs. But in the present state of our knowledge such extremely high densities seem difficult to justify—and the higher the density we must assume, the less important are some of the objections which have been raised against the case considered by Lemaître.

Einstein's and de Sitter's case $k=0$, $\lambda=0$. As the last case of a world in which energy is conserved we consider that extremely simple one of type M_1 which has been discussed recently by Einstein and de Sitter jointly.[39] The dependence of R on t is found immediately from (3.7) to be

$$R=(0.75\kappa E)^{1/3}(ct)^{2/3} \tag{8.8}$$

where t is measured from the time at which $R=0$. Here the predictions concerning red shift and density of matter are not independent, for we have

$$R'/R=c(\kappa c^2\rho/3)^{\frac{1}{2}}; \tag{8.9}$$

on setting this equal to 1.8×10^{-17} from (4.10) we find $\rho=5.8\times 10^{-28}$, which is just on the upper limit of the densities considered in the foregoing.

[39] See bibliography, 1932.1. These authors, as well as Tolman-Ward, 1932.6, and de Sitter, 1932.4, attribute the possibilities $k=-1$, 0 to Heckmann, 1931.2, although the former had been discussed explicitly by Friedmann, 1924.1, the latter had been used by Lemaître, 1925.1, and the present author, 1928.1, in their discussions of the stationary form (6.1) of the de Sitter universe, and both had been derived from the general standpoint of this report by the present author 1929.2!

9. Universes in which matter is conserved. Effect of annihilation on the line element

The next extensive class of models of the universe to receive detailed attention is that in which matter is rigorously conserved and exerts no pressure. The explicit form of the dependence of E on R was found by Lemaître,[40] and a systematic discussion of the dependence of R on t for the various types for which $k=+1$ was subsequently given by de Sitter.[41] On requiring that M be constant and that p_m vanish (3.6) becomes

$$d(uR^3)+uR^2dR=0, \quad (p'=u/3),$$

whence

$$u=\beta/R^4, \quad E=Mc^2+\beta/R. \tag{9.1}$$

Since here both E and p are positive monotonically decreasing functions of R the classification of possibilities in §7 applies *in toto;* the universes of this class differ but little from those we have already discussed, and we therefore confine ourselves to the explicit solution of two cases of particular interest.

The first of these is the unique solution $k=+1$, $\lambda=\lambda_e$, $R_0>R_e$ which gives a universe which has been expanding forever, and is the immediate generalization of the corresponding Lemaître case of Friedmann's universes. Here R_e is the positive root of the quadratic

$$6R_e^2-3\kappa Mc^2R_e-4\kappa\beta=0, \tag{9.2}$$

and λ_e is determined in terms of it by

$$\lambda_e=1/R_e^2+\kappa\beta/3R_e^4. \tag{9.3}$$

The dependence of R on t is defined by

$$c(t-t_0)=3^{\frac{1}{2}}R_e\int_{R_0}^{R}\frac{xdx}{X^{\frac{1}{2}}(x-R_e)} \tag{9.4}$$

where

$$X=(1+\kappa\beta/3R_e^2)(x^2+2R_ex)+\kappa\beta; \tag{9.5}$$

the integral can be evaluated in terms of logarithms.[42] For any reasonable estimate of the pressure this solution is indistinguishable from Lemaître's (for which $p=0$), and we therefore carry the work no further; for a more detailed account the reader is referred to de Sitter's original papers.

A case of some theoretical interest, although without application to the actual world, is that in which there is only radiation, i.e., $M=0$. Eq. (3.7) then becomes

$$c(t-t_0)=3^{\frac{1}{2}}\int_{R_0}^{R}\frac{xdx}{(\lambda x^4-3kx^2+\kappa\beta)^{\frac{1}{2}}} \tag{9.6}$$

and is immediately integrable. The only universe of this type which has been expanding forever is that special case obtained from (9.2)–(9.5) on setting $M=0$:

$$R_e^2=2\kappa\beta/3, \quad \lambda_e=3/2R_e^2,$$
$$2^{\frac{1}{2}}c(t-t_0)=R_e\log\frac{R^2-R_e^2}{R^2-R_0^2}. \tag{9.7}$$

As pointed out by de Sitter, this solution parallels remarkably closely the one considered by Lemaître.

An investigation of the effect of the random motions of nebulae on the line element has been carried out by Lemaître[43] in order to explain an apparent discrepancy found by de Sitter.[44] On assuming motion in geodesics for which the integral $R^2du/ds=1/\gamma\neq 0$ [cf. (E. 5)] the pressure p_m due to matter is computed and incorporated into the fundamental equation for the determination of $R(t)$. But on replacing this R^2 by an appropriately chosen linear function of R^2, which differs but little from the identity, he shows that this complete equation reduces to precisely the same one as that one resulting from (9.1) above. Lemaître concludes that the average value of $1/\gamma$ will give a good approximation to the true situation—this is indeed the procedure on which the general development of §2 was based. The course of the various possibilities arising under this assumption has been sketched by Heckmann.[45]

The assumption on which this section is based, that matter is rigorously conserved, must, however, be modified in view of present tendencies in astrophysics, for if we are to account for the long life of the stars we seem forced to assume that in them matter is continually being transformed into radiation. The effect of this loss of

[40] See bibliography, 1927.1.

[41] See bibliography, 1930.3; 1931.7.

[42] See bibliography, de Sitter, 1930.3.

[43] See bibliography, 1930.2.

[44] See bibliography, 1930.3.

[45] See bibliography, 1932.2.

matter on the structure of the universe has been investigated by Tolman[46] and by de Sitter.[47] The former has attacked the problem by considering the effect of the various coefficients of t in the Taylor expansion of $R(t)$ (or rather of $\log R$) on the logarithmic derivative dM/Mdt of the total mass M. But, whereas the coefficient of t is determined uniquely by the red shift, the coefficients of the second and third powers enter into the present problem in such a way as to frustrate any attempt at a unique prediction until further observational data are available. de Sitter, on the other hand, assumes the definite law

$$(1/M)(dM/dt) = -\gamma(1/R)(dR/dt) \quad \text{or} \quad M = M_0(R_0/R)^\gamma \tag{9.8}$$

for the rate of decrease of matter, and determines the constant of proportionality γ to be of the order of 2×10^{-7}. The conservation equation (3.6) then allows him to conclude that

$$u = \frac{\gamma M_0 c^2}{1-\gamma}\frac{R_0^\gamma}{R^{3+\gamma}} + \frac{\beta}{R^4}, \quad E = \frac{M_0 c^2}{1-\gamma}\frac{R_0^\gamma}{R^\gamma} + \frac{\beta}{R}. \tag{9.9}$$

The pressure is a monotonically decreasing function of time, and this universe is accordingly included among the general types discussed in §7; however, this conclusion will only hold at all times in the past provided $\beta \geqq 0$. de Sitter has shown that the fact that the total amount of radiation in a fixed coordinate volume is decreasing, in spite of the annihilation of matter, can be attributed to the accompanying decrease in pressure under the expansion; he also rightly insists that the provisional law (9.8) can only be considered valid over a comparatively short range. The assumption of the "short time scale" for the life of the stars would of course still lessen the importance of this correction to the law of conservation of mass assumed in the above.

IV. Conclusion

10. The condensation problem

Our survey of all universes suitable for cosmology has led to the conclusion that under the reasonable assumption $p \geqq 0$ the only possibilities for a universe in equilibrium or for one which has not arisen in finite time from the singular state $R=0$ are included under that class in which space is finite and the cosmological constant is greater than zero. We are, therefore, tempted to conclude that the ideal background of the actual universe is of this type—that it has arisen from the unstable Einstein state by the influence of some perturbation and is now expanding at an ever increasing rate which will lead it eventually into the de Sitter state in which the observer and his more immediate physical system (nebula or intimate group of nebulae) are the only things existing within their own observable universe. The question which naturally arises concerns how this expansion originated—what processes in a universe in unstable equilibrium can be considered responsible for the impulse which set it off on its course? Here we are indulging in speculations which far transcend those contained in the foregoing, but the problem in one form or another is of such importance to modern astronomy that we can well afford to recount the progress which has been made in this direction.

The question naturally divides itself into two parts: How can we account for the initial expansion in terms of the quantities with which we have dealt so far, and to what physical agencies—perhaps without the province of our previous interest—can we attribute these intermediary effects? To the first we can give a definite answer, but with regard to the second we can only indicate the partial attempts which have been made to answer it. What, then, are the changes in such quantities as pressure or density which can cause a universe in equilibrium to begin to expand? In its equilibrium state such an Einstein universe is specified by a pressure p_e and a density ρ_e which are related to the cosmological constant $\lambda > 0$ and the radius R_e by Eqs. (5.2); we think of these quantities as having these constant values up to the time $t=0$, at which time they begin to vary. For simplicity, we suppose that they are continuous at $t=0$ and that for a sufficiently short time thereafter they may be expanded in powers of t—although the validity of the results is not dependent on this simplification. First, the Eqs. (3.2) which govern their behavior and the continuity condition require

[46] See bibliography, 1930.6–7.

[47] See bibliography, 1930.3.

that at time $t=0$ both R' and R'' vanish, and we may therefor write the excess ΔR of R over R in the form

$$\Delta R = R(t) - R_e = At^{n+2} + \cdots \quad (10.1)$$

(plus terms of higher order), where $n>0$. On substituting this expression for R into Eqs. (3.2), (3.4), (3.5), we find that the variation in ρ, p and the quantities derived from them accompanying such a variation in R are

$$\begin{aligned} \Delta\rho &= -[6A/\kappa c^2 R_e^3]t^{n+2} + \cdots, \\ \Delta p &= -[2(n+1)(n+2)A/\kappa c^2 R_e]t^n + \cdots, \\ \Delta\rho_m &= [6(n+1)(n+2)A/\kappa c^4 R_e]t^n + \cdots, \quad (10.2) \\ \Delta M &= [6(n+1)(n+2)R_e^3 A/\kappa c^4]t^n + \cdots, \\ \Delta E &= -[3p_e R_e^2 A]t^{n+2} + \cdots. \end{aligned}$$

From this result we can conclude that if any cause produces a diminution in the pressure or an increase in the proper mass or density of matter, the universe will begin to expand and in its subsequent course will be of type A_2, and if on the other hand these changes are of the opposite sense it will begin to contract.[48] By an elegant method, the result of which includes the foregoing as a special case, Lemaître has examined the effect of an instantaneous decrease of pressure to zero and has found that the present state would be attained in a period of time of the order of 10^{10}–10^{11} years.

This last result requires a few words concerning the time scale to which we are led by relativistic cosmology. All of the universes of types other than A_2 which have been investigated in detail lead to times of the order of the foregoing since the universe was in the singular state. Such a time scale seems far too short for modern astrophysics, according to which the ages of the stars and the time required to reach the observed degree of equipartition of energy may be of the order of 10^{12} years or more. Hence, unless we wish to revise these estimates downward—perhaps on the basis of an acceleration of the processes involved at a time when the radius was less than at present—or to admit that the radius has at some comparatively recent time gone through the singular or minimum state, we are forced to accept the unique possibility A_2 which allows an infinite time. But if we further require that $dp/dR=0$, as indicated in the above, this infinity is only a logarithmic one,[49] and the time which has elapsed since the perturbation set in is apt to be exceedingly short, as evidenced by Lemaître's result quoted above. Nevertheless we can obtain as ample a time scale as we wish by assuming sufficiently weak perturbations—and we are thereby brought face to face with the second question: What physical processes are responsible for the initial disturbance?

It was suggested by Eddington that the development of a condensation in the regular distribution of matter might be responsible for the initial impulse which started the expansion. To test this hypothesis McCrea and McVittie[50] undertook an investigation of the effect of one or more spherically symmetric condensations of mass m on the structure of a universe in equilibrium; however, they found that to terms of first order in m the volume of space was unaltered by the presence of such singularities—agreeing with the conclusions of Lemaître, who attacked the problem from a somewhat different standpoint.[51] Lemaître then showed that although the mere existence of such condensations could not account for the expansion, a process which he calls "stagnation," and which he attributes to the presence of condensations, causes a decrease in pressure, whence in accordance with (10.2) the universe will begin to expand. This stagnation is the result of kinetic energy, which would otherwise be free to wander through the uniform world, being captured, thereby causing a decrease in the pressure due to matter. Although his specific analysis has been criticized severely by McCrea and McVittie, their only pertinent objection—that this diminution in pressure is in fact a consequence of the formation of condensations—is partially answered on the ground that the decrease in pressure to which Lemaître refers is a secondary effect which could not occur if it were not for the existence of random motions.

[48] See bibliography, Lemaître, 1931.5. See also Eddington, 1930.1.

[49] To which Eddington, 1930.1 (see bibliography), objects on the ground that such infinities are usually found illusory in physics.

[50] See bibliography, 1931.6; 1932.3.

[51] See bibliography, 1931.5.

More recently McVittie[50] has carried his investigations further, and concludes from them that Jeans' time scale for the formation of nebulae can be reduced to one-quarter, bringing it nearer to 10^{10} years.

Throughout this report we have emphasized those universes of type A_2 which allow, at least in principle, and indefinite time scale in the past as well as in the future. However a recent investigation by Lemaître[52] of the problem of the formation of condensations in a universe of the type M_1, which began in the singular state $R=0$ at a finite time in the past, is based on the opinion that the relatively short time scale to which such a universe leads is not necessarily irreconcilable with modern astrophysics. This investigation leads to the interesting conclusion that whereas matter at a sufficient distance will recede from the particular condensation in question in much the same manner as in a homogeneous universe of type M_1, the nearer matter will fall back into the condensation as in a periodic universe of type O, thus contributing to the formation of a nebula.

11. Summary

We conclude with a brief statement of the results so far attained in the field of relativistic cosmology, with emphasis on the assumptions involved and on the relation of the observations to the general theory. In the first place, we accept the data, due primarily to Hubble and Shapley, on the uniform distribution of matter in the large within the visible universe, and extrapolate them to the universe as a whole. In addition to this uniformity of distribution, we accept the evidence showing that the relative peculiar motions of neighboring nebulae are extremely small compared with the velocity of light and base upon it an assumption, due in the first instance to Weyl, which allows us to introduce a cosmic time enabling us to speak in a significant manner of simultaneity in the world at large. The assumption concerning the uniform spatial density of matter (and radiation) now allows us to conclude, on the basis of the general theory of relativity and its precise formulation of Mach's principle, that the ideal background of the actual universe is one in which the spaces $t=\text{const.}$ are of constant Riemannian curvature and in which the lines of parameter t are the geodesics representing the mean motion of matter, as formulated more precisely by (2.2).

We have thus arrived at a form for the structure of the universe as a whole *without making use of the empirical data concerning the red shift observed in light from extra-galactic objects*, except insofar as we drew upon them to justify Weyl's postulate—which can, however, be made to stand upon its own by *a priori* reasoning acceptable to many. That such a universe will in general lead to a relation between distance and red shift, which is linear over a sufficiently restricted range, was a consequence of the previous work, although from it alone we could gain no estimate of the constant of proportionality.

Our survey of the possible types of universes led to the conclusion that the only ones in which the total pressure due to matter and radiation is never negative and in which the universe is never in the singular state $R=0$ (or its equivalent in which $E(R)=\infty$) are those for which space is finite and the cosmological constant greater than zero. Under the additional physically plausible assumption that the pressure does not increase with the radius of the spherical (or elliptic) space, we found among these a unique type E (the Einstein universe) which is in equilibrium, and a unique type A_2 which may be considered as arising in this equilibrium state and allowing an infinite time scale both in the past and in the future; of these we considered in some detail the case of Friedmann's energy-conserving universes discussed by Lemaître and the corresponding matter-conserving universe discussed by de Sitter. On interpreting the observed relationship between red shift and distance as that predicted by the theory we found that in general the meager time scale of the order of 10^{10} years offered by all other possibilities was against them, and that even those of type A_2 were subject to a corresponding—but in our opinion not necessarily insuperable[53]—difficulty in that their infinite past time was due to a logarithmic dependence between time and the variation of the radius

[52] As presented by Professor Lemaître in a seminar in Princeton University, November, 1932. To appear shortly in Phys. Rev.

[53] See bibliography, de Sitter, 1932.4, who seems also to be of this opinion.

from its equilibrium value. (If, as indicated by Oort,[19] Hubble's value (4.10) of the expansion constant h is too large, the time scales allowed by all types of universes will be correspondingly increased.) We also indicated in brief the progress made by Eddington, Lemaître and others toward an explanation of the disruption of equilibrium.

It is clear from this summary that the existence of the so-called velocity-distance relation formed no essential part of the deduction, which was based entirely on the evidence for the uniform distribution and state of motion of matter in the large, and on an acceptance of the general theory of relativity in the form proposed by Einstein in 1917—in which the cosmological constant λ is included. If we consider the observed red shift as arising from the nature of space-time we find in it additional evidence for the theory—and in particular for an asymptotic universe of the type A_2. It is therefore clear that any alternative explanation of the red shift (such as that proposed by Zwicky,[54] who would attribute it to a gravitational analogue of the Compton effect, or the kinematical model proposed by Milne[55]) which does not deny the general theory of relativity, is forced on the basis of other evidence to choose between one or another of the various types covered in this report—among which is to be counted the flat Minkowski space SE which most nearly approaches the Newtonian cosmology. In giving this survey of cosmologies we are convinced that the underlying theory forms an integral part of the theory of relativity, and that although the choice of a particular model may for the present be influenced by the predilections of the individual, we can hope that the future will reveal additional evidence to test its validity and to lead us to a satisfying solution.

V. Notes

A. Mathematical formulae

We place here for reference the explicit definitions of various expressions employed in the text. Greek indices μ, ν, etc., range from 0 to 3 and indices α, β, etc., from 1 to 3; repeated indices imply summation over their range. $g=|g_{\mu\nu}|<0$, determinant of the coefficients $g_{\mu\nu}$. $g^{\mu\nu}$=cofactor of $g_{\mu\nu}$ in g divided by g itself. $A^\mu=g^{\mu\nu}A_\nu$, $A_\mu=g_{\mu\nu}A^\nu$—any index may be raised with the aid of $g^{\mu\nu}$ or lowered with $g_{\mu\nu}$.

$$\left\{ {\sigma \atop \mu\nu} \right\} = \tfrac{1}{2} g^{\sigma\tau}\left(\frac{\partial g_{\mu\tau}}{\partial x^\nu}+\frac{\partial g_{\tau\nu}}{\partial x^\mu}-\frac{\partial g_{\mu\nu}}{\partial x^\tau}\right). \tag{A.1}$$

$$R^\sigma_{\mu\nu\tau}=\frac{\partial}{\partial x^\nu}\left\{ {\sigma \atop \mu\tau} \right\}-\frac{\partial}{\partial x^\tau}\left\{ {\sigma \atop \mu\nu} \right\}+\left\{ {\rho \atop \mu\tau} \right\}\left\{ {\sigma \atop \rho\nu} \right\}-\left\{ {\rho \atop \mu\nu} \right\}\left\{ {\sigma \atop \rho\tau} \right\}. \tag{A.2}$$

$$R_{\sigma\mu\nu\tau}=g_{\sigma\rho}R^\rho_{\mu\nu\tau}=\tfrac{1}{2}\left(\frac{\partial^2 g_{\sigma\tau}}{\partial x^\mu \partial x^\nu}+\frac{\partial^2 g_{\mu\nu}}{\partial x^\sigma \partial x^\tau}-\frac{\partial^2 g_{\sigma\nu}}{\partial x^\mu \partial x^\tau}-\frac{\partial^2 g_{\mu\tau}}{\partial x^\sigma \partial x^\nu}\right) + g_{\kappa\lambda}\left(\left\{ {\kappa \atop \mu\nu} \right\}\left\{ {\lambda \atop \sigma\tau} \right\}-\left\{ {\kappa \atop \sigma\nu} \right\}\left\{ {\lambda \atop \mu\tau} \right\}\right). \tag{A.3}$$

$$R_{\mu\nu}=R^\sigma_{\mu\nu\sigma}=\tfrac{1}{2}\frac{\partial^2 \log g}{\partial x^\mu \partial x^\nu} -\frac{\partial}{\partial x^\sigma}\left\{ {\sigma \atop \mu\nu} \right\}+\left\{ {\rho \atop \mu\sigma} \right\}\left\{ {\sigma \atop \rho\nu} \right\}-\tfrac{1}{2}\left\{ {\sigma \atop \mu\nu} \right\}\frac{\partial \log g}{\partial x^\sigma}. \tag{A.4}$$

$$R=R_\mu{}^\mu=g^{\mu\nu}R_{\mu\nu}. \tag{A.5}$$

For the line element defined by (2.2), (2.3), we obtain the following nonvanishing components:

$$\left\{ {\alpha \atop \beta\gamma} \right\}=\left\{ {\alpha \atop \beta\gamma} \right\}^*, \quad \left\{ {0 \atop \alpha\beta} \right\}=\frac{RR'h_{\alpha\beta}}{c^2}, \quad \left\{ {\alpha \atop 0\beta} \right\}=\frac{R'\delta_\beta{}^\alpha}{R} \tag{A.6}$$

where the asterisk on the Christoffel symbol indicates that is is to be computed from the coefficients $h_{\alpha\beta}$ of the line element (2.3).

$$R_{\alpha\beta\gamma\delta}=-R^2(k+R'^2/c^2)(h_{\alpha\gamma}h_{\beta\delta}-h_{\alpha\delta}h_{\beta\gamma}), \quad R_{0\alpha0\beta}=RR''h_{\alpha\beta}. \tag{A.7}$$

$$R_{\alpha\beta}=-[RR''+2(kc^2+R'^2)]h_{\alpha\beta}/c^2, \quad R_{00}=3R''/R. \tag{A.8}$$

$$R=6(RR''+R'^2+kc^2)/R^2c^2. \tag{A.9}$$

B. Riemannian curvature. Spherical space

The Riemannian curvature of a three-dimensional space V_3 is a directional quantity which is defined in terms of the more familiar total or Gaussian curvature of certain two-dimensional sub-spaces V_2 as follows. We first recall the definition of this latter; for this purpose consider the family of plane curves which are obtained by cutting an ordinary two-dimensional surface in three-dimensional Euclidean space by the family of planes containing the normal to a given point P on the surface. A curve of this family will have a curvature $1/\rho$ at P, where ρ is the radius of the osculating circle to the plane curve in question at P; let $1/\rho_1$, $1/\rho_2$, denote its maximal and minimal values. The total or Gaussian curvature of the surface at the point P is then the product $k(\mathrm{P})=1/\rho_1\rho_2$ of these two extremal values and is an intrinsic invariant, i.e., it is unchanged on making an arbitrary deformation of the

[54] F. Zwicky, *On the Red Shift of Spectral Lines through Interstellar Space*, Proc. Nat. Acad. Sci. **15**, 773–779 (1929).

[55] E. A. Milne, *World Structure and the Expansion of the Universe*, Nature **130**, 9–10 (1932).

surface without tearing or stretching, and can be computed from the coefficients of the line element and their first and second derivatives at P.

Returning to the three-space V_3, consider the surface defined by all geodesics (curves in V_3 for which the integral of ds between any two points on them is extremal) passing through a given point P, whose directions at P lie in the plane of directions

$$\lambda^\alpha = \xi_1\lambda_1{}^\alpha + \xi_2\lambda_2{}^\alpha \quad (\xi_1,\ \xi_2,\ \text{arbitrary}) \tag{B.1}$$

determined by the two vectors $\lambda_1{}^\alpha$, $\lambda_2{}^\alpha$ at P. The Riemannian curvature at P of the space V_3 in the two-spread of directions λ^α is defined to be the Gaussian curvature $k = k(\mathrm{P}, \lambda_1, \lambda_2)$ of the surface thus generated by the geodesics; it is given explicitly in terms of the coefficients $h_{\alpha\beta}$ of the line element of V_3 by[56]

$$k = \frac{R^*_{\alpha\beta\gamma\delta}\lambda_1{}^\alpha\lambda_2{}^\beta\lambda_1{}^\gamma\lambda_2{}^\delta}{(h_{\alpha\gamma}h_{\beta\delta} - h_{\alpha\delta}h_{\beta\gamma})\lambda_1{}^\alpha\lambda_2{}^\beta\lambda_1{}^\gamma\lambda_2{}^\delta} \tag{B.2}$$

where $R^*_{\alpha\beta\gamma\delta}$ is the Riemann-Christoffel tensor (A.3) formed from the $h_{\alpha\beta}$. These results are immediately applicable to spaces of a higher number of dimensions.

If the Riemannian curvature k is independent of direction the above allows us to conclude that

$$R^*_{\alpha\beta\gamma\delta} = k(h_{\alpha\gamma}h_{\beta\delta} - h_{\alpha\delta}h_{\beta\gamma}) \tag{B.3}$$

and it can be shown that k must also be independent of the point P (Schur's theorem); in this case we speak of the V_3 as having constant Riemannian curvature k. The coordinates x^α may be chosen in such a way that (2.3) assumes the form

$$du^2 = (dx_1{}^2 + dx_2{}^2 + dx_3{}^2)/(1 + kr^2/4)^2 \tag{B.4}$$

where r^2 is the sum of the squares of the coordinates. For $k = 1/\rho^2 > 0$ (B.4) describes *spherical space* of radius ρ, where the variables range from $-\infty$ to $+\infty$; the total volume of this space is readily found to be $2\pi^2\rho^3$ (where, as always in differential geometry, the volume of the elementary parallelopiped defined by the surfaces $x^\alpha = \text{const.}$, $x^\alpha + dx^\alpha = \text{const.}$ is $h^{\frac{1}{2}}dx_1dx_2dx_3$, where h is the determinant of the coefficients of the line element du^2). The analogy with an ordinary sphere in Euclidean three-space is strikingly brought out by considering spherical space as the hyper-surface

$$u_1{}^2 + u_2{}^2 + u_3{}^2 + u_4{}^2 = \rho^2 \tag{B.5}$$

in the Euclidean four-space whose line element is

$$du^2 = du_1{}^2 + du_2{}^2 + du_3{}^2 + du_4{}^2. \tag{B.6}$$

On introducing the parametric representation

$$u_\alpha = x^\alpha/(1 + r^2/4\rho^2), \qquad u_4 = \rho(1 - r^2/4\rho^2)/(1 + r^2/4\rho^2) \tag{B.7}$$

in terms of the spatial coordinates x^α we obtain from (B.6) the line element (B.4) for spherical space of curvature $k = 1/\rho^2$. Another form for the line element of a space of constant curvature $1/\rho^2$ is

$$du^2 = \rho^2(d\chi^2 + \sin^2\chi d\theta^2 + \sin^2\chi \sin^2\theta d\varphi^2) \tag{B.8}$$

where the coordinates vary over the range $0 \leqq \chi \leqq \pi$, $0 \leqq \theta \leqq \pi$, $0 \leqq \varphi < 2\pi$; this *elliptic space*, which results from spherical space on identifying antipodal points, has a total volume of $\pi^2\rho^3$.

C. Groups of motions. Periodic and stationary universes

The theory of groups of automorphisms or motions of a space into itself constitutes the natural mathematical tool for the investigation of spaces characterized by *a priori* symmetry conditions. As such, it is of particular value in the theory of relativity, for in virtue of Mach's principle it enables us to interpret directly in terms of the line element of space-time the effect of symmetry properties in the material and energetic distribution which determines the world geometry. Because of the lack of an adequate discussion of this discipline and its applications in the physical literature we have ventured to sketch briefly its main features and to indicate the application of it to the present problem.

Suppose that two points P, $\bar{\mathrm{P}}$ of space-time are physically equivalent in the sense that all intrinsic properties of the world as viewed by an observer O at P are indistinguishable from those as viewed by $\bar{\mathrm{O}}$ at $\bar{\mathrm{P}}$. Then by Mach's principle this is tantamount to the assertion that the geometrical structure of space-time is the same from the standpoint of O as from the standpoint of $\bar{\mathrm{O}}$; the equivalence of P and $\bar{\mathrm{P}}$ is then expressed by the assertion that $\bar{\mathrm{O}}$ can choose a coordinate system $\bar{x}^\mu$ in terms of which the new metric is defined by coefficients $\bar{g}_{\mu\nu}(\bar{x})$ which are exactly the same functions of the $\bar{x}^\mu$ as the $g_{\mu\nu}(x)$ are of the original coordinates $x^\mu : \bar{g}_{\mu\nu}(\bar{x}) = g_{\mu\nu}(\bar{x})$. On employing the law of transformation of the tensor $g_{\mu\nu}$ we conclude that the equations

$$g_{\sigma\tau}(\bar{x})(\partial\bar{x}^\sigma/\partial x^\mu)(\partial\bar{x}^\tau/\partial x^\nu) = g_{\mu\nu}(x) \tag{C.1}$$

are to admit a solution $\bar{x}^\mu = \bar{x}^\mu(x)$ which sends P into the point $\bar{\mathrm{P}}$ equivalent to it.

If we now assume that $\bar{\mathrm{P}}$ can be reached from P by a path consisting of points equivalent to P then the Eqs. (C.1) for an automorphism of the space on itself will admit a solution

$$\bar{x}^\mu = \psi^\mu(x, \sigma) \tag{C.2}$$

containing a continuous parameter σ, for some value—say $\sigma = 0$—of which (C.2) becomes the identical transformation $\bar{x}^\mu = x^\mu$. But these equations then obviously define a one-parameter transformation group in the sense of Lie,

[56] Cf. Eisenhart, *Riemannian Geometry*, pp. 79–88, for this and the following results.

[57] For a more complete development of the theory see chapter VI of Eisenhart's *Riemannian Geometry* and for application to Schwarzschild's problem of determining the field of a spherically symmetric body see J. Eiesland, *The Group of Motions of an Einstein Space*, Trans. Amer. Math. Soc. **27**, 213–245 (1925). The application of interest here is an amplification of Robertson, 1929.2 (see bibliography)

for if P_1 is equivalent to P and P to $\bar{P}$, then P_1 is equivalent to $\bar{P}$. Denoting the infinitesimal transformation belonging to this group by $\delta\sigma$:

$$\bar{x}^\mu = x^\mu + \xi^\mu \cdot \delta\sigma, \quad \text{where} \quad \xi^\mu(x) = \left[\frac{\partial\psi^\mu(x, \sigma)}{\partial\sigma}\right]_{\sigma=0}, \tag{C.3}$$

the Eqs. (C.1) for the automorphism become the equations of Killing

$$\frac{\partial g_{\mu\nu}}{\partial x^\tau}\xi^\tau + g_{\tau\nu}\frac{\partial \xi^\tau}{\partial x^\mu} + g_{\mu\tau}\frac{\partial \xi^\tau}{\partial x^\nu} = 0. \tag{C.4}$$

So far as symmetry properties which are expressible by a continuous transformation group are concerned we may replace the finite Eqs. (C.1) by the linear homogeneous Eqs. (C.4) involving only the generators ξ^μ of the group in addition to the $g_{\mu\nu}$. Dropping for the moment the restriction $n=4$ by allowing the indices to run through the range 0, 1, 2, $\cdots$, $n-1$, the fundamental result is that an n-space having the symmetry characterized by the r-parameter group with generators $\xi_u{}^\mu$ ($u=1, 2, \cdots, r$) is such that the coefficients $g_{\mu\nu}$ of its line element satisfy the Eqs. (C.4) for each of the $\xi_u{}^\mu$. In particular, it can be shown[58] with the aid of the conditions of integrability of (C.4) that an n-space can have at most an $[n(n-1)/2]$-parameter group of motions, and that such a space is of constant Riemannian curvature k.

To return to space-time and the original problem of determining the restrictions placed on (2.1) by the uniformity conditions discussed in the text, the equations of Killing become

$$\begin{gathered} g_{00}\frac{\partial \xi^0}{\partial t} = 0, \qquad g_{\alpha\beta}\frac{\partial \xi^\beta}{\partial t} + g_{00}\frac{\partial \xi^0}{\partial x^\alpha} = 0, \\ \frac{\partial g_{\alpha\beta}}{\partial t}\xi^0 + \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\xi^\gamma + g_{\gamma\beta}\frac{\partial \xi^\gamma}{\partial x^\alpha} + g_{\alpha\gamma}\frac{\partial \xi^\gamma}{\partial x^\beta} = 0. \end{gathered} \tag{C.5}$$

Now the group which specifies the spatial isotropy and homogeneity of space-time leaves t unaltered ($\bar{t}=t$, whence $\xi^0=0$) and involves six parameters, corresponding to the equivalence of all ∞^3 points in the three-space $t=$const., of all ∞^2 spatial directions about each point and of all ∞^1 two-spreads of directions through each direction at each point; Eqs. (C.5) then require that each of the six generators ξ^α be independent of t and satisfy the fundamental Eqs. (C.4) on restricting μ, ν, τ to the range 1, 2, 3. But then the $g_{\alpha\beta}(t, x)$ define a three-space which admits a six-parameter group of motions, and in accordance with the remarks above this three-space is consequently of (spatially) constant Riemannian curvature $k(t)$. Now it can furthermore be shown that this three-space must be of the form

$$g_{\alpha\beta}(t, x)dx^\alpha dx^\beta = -R^2(t)h_{\alpha\beta}(x)dx^\alpha dx^\beta \tag{C.6}$$

introduced in the text, as otherwise certain of the generators of the group would necessarily contain t, contrary to the above. A rigorous and more elegant derivation of this result, involving, however, more technical machinery than that sketched in the foregoing, is due to Fubini.[59]

The additional requirement that space-time be stationary is expressed by the requirement that there exist an additional one-parameter group of motions for which ξ^0 does not vanish and is at most a function of t. It can be shown[60] that (C.5) and their conditions of integrability allows us to conclude that

$$R(t) = R_0 e^{ct/a}, \qquad k/a = 0, \tag{C.7}$$

from which the two cases (2.4), (2.5), follow immediately.

As a further example of the methods here employed we derive, from a somewhat more elementary standpoint, the restriction imposed on the line element (2.2) by the condition that the structure of space-time be a periodic function of cosmic time. This investigation, which includes the above paragraph as the special case in which the period $\tau=0$, is suggested by work of Tolman[61] which is based on one of the possibilities obtained below. The condition is expressed mathematically by the requirement that there exist a transformation of the form

$$\bar{t} = t + \tau, \qquad \bar{x}^\alpha = \bar{x}^\alpha(t, x), \tag{C.8}$$

which preserves the form of the line element; on replacing $\bar{t}$ by $t+\tau$ the terms in dt^2 drop out, leaving us with the equation

$$R^2(t+\tau)h_{\alpha\beta}(\bar{x})d\bar{x}^\alpha d\bar{x}^\beta = R^2(t)h_{\alpha\beta}(x)dx^\alpha dx^\beta. \tag{C.9}$$

But it is at once evident from the definitions (B.3), (A.3) that the Riemannian curvature of the manifold defined by the line element on the left is $k/R^2(t+\tau)$, while the curvature of the one defined by that on the right is $k/R^2(t)$; hence we must have

$$k[R^2(t+\tau) - R^2(t)] = 0. \tag{C.10}$$

Consequently, if $k\neq 0$ we must have

$$R(t) = \text{periodic function of } t \text{ with period } \tau, \tag{C.11}$$

and the transformation $\bar{x}\to x$ may be taken as the identity —the case considered by Tolman. But we may alternatively have $k=0$, in which case the space (2.3) is Euclidean, and by a well-known theorem due to Liouville the only transformations satisfying (C.9) are combinations of translations and rotations with the dilatation

$$R(t+\tau)\bar{x}^\alpha = R(t)x^\alpha \tag{C.12}$$

where t is considered as a parameter. But t is actually a space-time coordinate, and in order to avert the appear-

[58] Cf. Eisenhart, *Riemannian Geometry*, p. 238.

[59] G. Fubini, *Sugli spazii a quattro dimensioni che ammettono un gruppo continuo di movimenti*, Ann. Mat. pura appl. [3] 9, 33–90 (1904)—in particular, p. 64.

[60] See bibliography, Robertson, 1929.2, p. 825. See also reference to Fubini given in footnote 59 above, p. 83.

[61] See bibliography, 1931.10.

ance in (C.9) of terms of the form $dx\,dt$ we must have $R(t+\tau)=CR(t)$, or

$$R(t)=e^{ct/a}P(t), \tag{C.13}$$

where $P(t)$ is a periodic function of t with period τ. These two periodic possibilities (C.11) and (C.13) obviously degenerate into the Einstein and de Sitter solutions on setting $\tau=0$.

D. Representation of space-time in five-dimensional flat space. The Minkowski diagram

In order to obtain a graphical representation of the general space-time (2.2), corresponding to the Minkowski diagram of the special theory of relativity, we first show that such a universe may be considered as a four-dimensional surface in a flat five-dimensional space,—i.e., in a space in which the coefficients of the line element may be taken as constants. To this end we restrict ourselves for the moment to the case $k=+1$ in which the auxiliary metric (2.3) defines spherical space, which seems to be of most physical interest, and make use of the formulae (B.5)–(B.7) for its representation as a surface in four-space. On defining the five variables z_0, z_α, z_4, as functions of t, x^α by

$$z_0=\int (c^2+R'^2)^{\frac{1}{2}}dt, \qquad z_\alpha=Ru_\alpha, \qquad z_4=Ru_4, \tag{D.1}$$

where the u's are defined by (B.7) for $\rho=1$, it is readily shown that the line element (2.3) may be written

$$ds^2=dz_0^2-(dz_1^2+dz_2^2+dz_3^2+dz_4^2); \tag{D.2}$$

space-time may therefore be considered as the hyper-surface of revolution

$$z_1^2+z_2^2+z_3^2+z_4^2=R^2 \tag{D.3}$$

in the five-dimensional z-space, where the $R(t)$ on the right is to be thought of as a function of the distance z_0 along the axis in virtue of the first of the Eqs. (D.1).[62] In particular the Einstein universe $R=$const. is represented by a cylinder.

In order to obtain the two-dimensional representation corresponding to the Minkowski diagram, we drop two of the spatial dimensions by setting $x_2=x_3=0$; the kinematical background of this abbreviated space-time is then the ordinary surface of revolution (Fig. 7)

$$z_1^2+z_4^2=R^2[t(z_0)] \tag{D.4}$$

whose parametric representation in terms of the coordinates t, $x_1=x$ is

$$z_0=\int (c^2+R'^2)^{\frac{1}{2}}dt, \quad z_1=R\frac{x}{1+x^2/4}, \quad z_4=R\frac{1-x^2/4}{1+x^2/4}. \tag{D.5}$$

The world lines of observers "at rest" at the distance $uR(t)$ apart are the intersections of this surface by meridian planes through the z_0- or time axis and including an angle u between them.

[62] See bibliography, Robertson, 1929.2, p. 826.

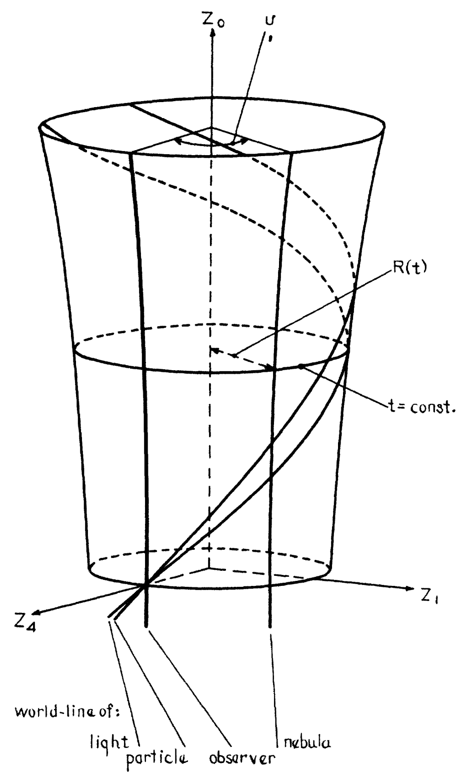

FIG. 7. Minkowski diagram for general cosmologies $k=+1$.

The representation of spaces for which $k=-1$ is similar to the one considered above—with the characteristic difference that dz_4^2 now appears in (D.2) with a positive sign. But we pass on to a brief statement of the corresponding results for the remaining case $k=0$, whose representation is surprisingly complicated compared with the above. We may here set

$$\begin{aligned} z_0&=(R/2b)(b^2+r^2)+f(t),\\ z_\alpha&=Rx^\alpha,\\ z_4&=(R/2b)(b^2-r^2)-f(t)[=bR-z_0], \end{aligned} \tag{D.6}$$

where

$$2bf(t)=c^2\int\frac{dt}{R'} \tag{D.7}$$

and b is an arbitrary constant which we take to be positive. Space-time is then the hyper-surface

$$z_0^2-(z_1^2+z_2^2+z_3^2+z_4^2)=2bRf, \tag{D.8}$$

where the function of t on the right is to be considered as a function of $z_0+z_4=bR(t)$. In contrast with the previous case, space-time is represented by only a portion of this surface, for we must have z_0+z_4 greater than some non-negative lower bound.

For the de Sitter universe (6.1) $R=e^{ct/a}$, whence

$$2bf(t)=-a^2/R(t). \tag{D.9}$$

This stationary universe is therefore represented by the pseudo-sphere

$$-z_0^2+z_1^2+z_2^2+z_3^2+z_4^2=a^2, \tag{D.10}$$

or parametrically in terms of the space-time parameters t, x^α by

$$\begin{aligned} z_0 &= a \sinh (ct/a)+(r^2/2a)e^{ct/a}, \\ z_\alpha &= x^\alpha e^{ct/a}, \\ z_4 &= a \cosh (ct/a)-(r^2/2a)e^{ct/a} \end{aligned} \tag{D.11}$$

on choosing $b=a$. The Minkowski diagram for the de Sitter universe is obtained from this result on suppressing two of the spatial coordinates by setting $z_2=z_3=0$; it is that part of the hyperboloid (D.10) of one sheet which lies above the plane $z_0+z_4=0$. This diagram is illustrated in Fig. 2 and described in the accompanying text of §6; the assertions there made concerning the relations existing between various world lines can be obtained with the aid of this representation.

E. Geodesics in space-time. Tolman's and Whittaker's definition of distance

The Eqs. (1.4), (1.5) are, in terms of the parameter u,

$$\frac{d^2x^\alpha}{ds^2}+\left\{\begin{matrix}\alpha\\ \beta\gamma\end{matrix}\right\}^*\frac{dx^\beta}{du}\frac{dx^\gamma}{du}=\left(\frac{d^2s}{du^2}\Big/\frac{ds}{du}-\frac{2}{R}\frac{dR}{du}\right)\frac{dx^\alpha}{du}, \tag{E.1}$$

$$\frac{d^2t}{du^2}+\frac{R}{c^2}\frac{dR}{dt}=\frac{d^2s}{du^2}\frac{dt}{du}\Big/\frac{ds}{du} \tag{E.2}$$

where

$$h_{\alpha\beta}\frac{dx^\alpha}{du}\frac{dx^\beta}{du}=1, \qquad \left(\frac{ds}{du}\right)^2=c^2\left(\frac{dt}{du}\right)^2-R^2. \tag{E.3}$$

On differentiating the second of Eqs. (E.3) with respect to u and eliminating dt/du, d^2t/du^2 by means of (E.2) and (E.3) we find

$$d^2s/du^2=(2/R)(ds/du)(dR/du), \tag{E.4}$$

whence (E.1) assumes the form (4.1) and in addition

$$ds/du=\gamma R^2. \tag{E.5}$$

Eq. (4.3) is then obtained from (E.3) on eliminating ds/du by means of (E.5).

The fact expressed by (4.1) that the spatial projection of a geodesic—including that of a null-line which represents the path of a beam of light— is itself a geodesic of the auxiliary space (2.3) enables us to investigate with ease a proposal due to Tolman and to Whittaker[63] concerning a definition of distance in space-time. Following the more general point of view of the latter, we define the distance d of an event P from a world line L as proportional to the square root of the normal area at the point of reception on L of light emitted in an infinitesimal solid angle at P, the constant of proportionality being so chosen that this definition agrees with the ordinary one for small distances. Since the projection onto the space (2.3) of the path of light is a geodesic we need only compute the normal area of a cone of geodesics at a point at distance u (as measured by the metric (2.3)) from the vertex of the cone; the product of the square root of this area into the function $R(t)$ computed at the time t of reception is then proportional to the distance d defined by Whittaker.

Considering in particular the case $k=+1$ in which (2.3) defines spherical space of unit radius, u is the angular distance between the points P′, L′ whose coordinates are the spatial coordinates of P and the event on L which represents the reception of light from P. Now by spherical geometry the normal area at L′ of the infinitesimal cone with vertex at P′ is proportional to $\sin^2 u$—as can be readily seen by cutting the spherical space with a fixed meridian plane, for the distance between the meridians on the resulting two-dimensional sphere, which are generators of the cone, is proportional to $\sin u$. Hence $d\backsim R(t) \sin u$, and in order that this agree with the ordinary definition of small distances the constant of proportionality must be unity, as expressed in Eq. (4.12).

Obviously for the case $k=0$ we have

$$d=R(t)u, \qquad (k=0), \tag{E.6}$$

agreeing with our (4.5), and a computation similar to that above enables us to conclude that for $k=-1$

$$d=R(t)\sinh u, \qquad (k=-1). \tag{E.7}$$

F. Tolman's relativistic thermodynamics

In following up the problem of the entropy of the universe as a whole, Tolman has recognized the desirability of supplementing the accepted mechanical principles of relativity, on which the development treated in the body of this report is based, by an appropriate generalization of thermodynamics. He has accordingly proposed a covariant form of the second law of thermodynamics—the first law being an adaptation of the conservation equations (1.3)—and has in particular applied it both to universes in static equilibrium and to cases of the more general nonstationary cosmologies defined by the line element (2.2).[64] Because of the importance of this problem, having as it does a possible bearing on the origin of cosmic radiation, we give here a summary of the, at first sight, rather startling conclusions to which Tolman has been led;

[63] See bibliography, Tolman, 1930.8; Whittaker, 1931.11.

[64] Insofar as it relates to the general cosmologies covered in this report Tolman's work is contained in 1931.8–10, 1932.5; Tolman-Ward, 1932.6 (see bibliography). For a statement of the general principles on which it is based see R. C. Tolman, *On the Use of the Entropy Principle in General Relativity*, Phys. Rev. [2] **35**, 896–903 (1930); R. C. Tolman and H. P. Robertson, *On the Interpretation of Heat in Relativistic Thermodynamics*, to appear shortly in Phys. Rev. For a general resumé and bibliography of previous work (including applications to the static Einstein cosmology) see Tolman 1931.8 (see bibliography).

that we have chosen to present it in the unified form of a brief note is motivated by the fact that in this way the continuity of development of the purely mechanical aspects of cosmology is not interrupted, and by the further fact that a more extended and authoritative account will be given soon by Tolman himself in the 1932 Gibbs' lecture *Relativity and Thermodynamics*.

The general covariant form of the second law of thermodynamics for a thermodynamic fluid with velocity $dx^\mu/d\tau$ states that the proper density of entropy φ_0 must satisfy the condition

$$(\partial/\partial x^\mu)[\varphi_0(-g)^{\frac{1}{2}}(dx^\mu/d\tau)]d\Sigma \geqq dQ_0/T_0, \qquad \text{(F.1)}$$

in which the equality holds for reversible processes; here dQ_0 is the proper measure of the net heat flowing into the space-time region $d\Sigma$ whose spatial boundaries share the motion of the fluid at points on them, and T_0 is the proper temperature of the boundary. Because of the spatial isotropy of the space-times considered in relativistic cosmology $dQ_0=0$, and furthermore $dx^\mu/d\tau=(1, 0, 0, 0)$; for such space-times the second law may therefore be written in the form

$$\partial\Phi/\partial t \geqq 0, \quad \text{where} \quad \Phi=\varphi_0 V \qquad \text{(F.2)}$$

is the proper entropy of that portion of the thermodynamic fluid occupying a volume $V=R^3$ whose coordinate volume in the auxiliary space (2.3) is unity, in accordance with the nomenclature of §3.

The first problem involving nonstationary cosmologies to which Tolman has applied his principle is that of determining whether a universe expanding or contracting at a finite rate can do so reversibly, a possibility which is suggested by the existence of periodic cosmologies of type O on purely mechanical grounds, but which is apparently in conflict with classical thermodynamics. With this in view he has considered in detail, among others, a universe filled with black-body radiation,[65] whose line element is of the form (9.6) for $k=1$, and a universe containing an equilibrium mixture of a perfect monatomic gas and black-body radiation.[66] In each of these equilibrium cases Tolman has been able to establish the reversibility from the equation[67]

$$d\Phi=dE/T_0+(p/T_0)dV \qquad \text{(F.3)}$$

for the increase $d\Phi$ of entropy in a given portion of the fluid; in fact, the first law, which here assumes the form (3.6), guarantees the vanishing of $d\Phi$.

More recently[68] Tolman has turned his attention to the investigation of universes in which irreversible processes occur, and has arrived at the possibility of such irreversible universes in which the entropy is not limited by an upper bound nor the free energy by a lower bound; this reasoning he has applied in detail to a universe with irreversible annihilation of matter.[69]

[65] See bibliography, 1931.8, p. 1653; 1931.10, p. 1767.

[66] See bibliography, 1931.9, p. 805; 1931.10, p. 1769.

[67] See bibliography, 1931.9, Eq. (14), p. 803; see also 1932.5, Eq. (18), p. 327.

[68] See bibliography, 1932.5.

[69] See bibliography, 1932.5, p. 334.

VI. Bibliography

This bibliography contains only those articles whose results have been incorporated into the unified account given in the report. For convenience the full titles have been given, together with a brief mention of the leading points discussed.

1917

1. A. Einstein, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, S.-B. preuss. Akad. Wiss. 1917, 142–152. Introduction of cosmological constant. Original treatment of Einstein static universe.
2. W. de Sitter, *On the Relativity of Inertia. Remarks Concerning Einstein's Latest Hypothesis*, Proc. Akad. Wetensch. Amsterdam 19, 1217–1225. Original treatment of de Sitter world.
3. ——, *On Einstein's Theory of Gravitation, and its Astronomical Consequences III*, Monthly Not. Roy. Astron. Soc. 78, 3–28. Includes account of original treatment of de Sitter universe.

1922

1. A. Friedmann, *Über die Krümmung des Raumes*, Zeits. f. Physik 10, 377–386. Finds line element (2.2) for $k=+1$ and discusses cases in which energy is conserved.
2. K. Lanczos, *Bemerkung zur de Sitterschen Welt*, Phys. Zeits. 23, 539–543. Contains derivation of nonstationary form for $k=+1$.

1923

1. K. Lanczos, *Über die Rotverschiebung in der de Sitterschen Welt*, Zeits. f. Physik 17, 168–189. Elimination of "mass horizon" and discussion of red shift in nonstationary form $k=+1$.
2. H. Weyl, *Zur allgemeinen Relativitätstheorie*, Phys. Zeits. 24, 230–232. Geometrical discussion of de Sitter universe. Introduction of postulate concerning coherency of matter and deduction of linear velocity-distance relationship.

1924

1. A. Friedmann, *Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes*, Zeits. f. Physik 21, 326–332. Extends 1922.1 to cases in which $k=-1$.

1925

1. G. Lemaître, *Note on de Sitter's Universe*, J. Math. and Physics (M.I.T.) 4, 188–192. Derivation and discussion of stationary form.

1926

1. E. P. Hubble, *Extra-galactic Nebulae*, Astrophys. J. 64, 321–369. Survey of these objects establishing their uniform distribution in visible portion of universe.

1927

1. G. Lemaître, *Un Univers Homogène de Masse Constante et de Rayon Croissant, Rendant Compte de la Vitesse Radiale de Nébuleuses Extra-galactiques*, Ann. Soc. Sci. Bruxelles **47**A, 49–59. Writes down line element (2.2) for $k=1$ and obtains condition (9.1) for conservation of mass. Detailed analysis Friedmann world for $k=1$, $\lambda=\lambda_e$, $R_0>R_e$. (Translation in Monthly Not. Roy. Astron. Soc. **91**, 483–490 (1931).)

1928

1. H. P. Robertson, *On Relativistic Cosmology*, Phil. Mag. [7] **5**, 835–848. Discussion of stationary form of de Sitter universe in coordinates employed by Lemaître, 1925.1.

1929

1. E. P. Hubble, *A Relation between Distance and Radial Velocity among Extra-galactic Nebulae*, Proc. Nat. Acad. Sci. **15**, 168–173. The empirical relationship deduced from observational data.
2. H. P. Robertson, *On the Foundations of Relativistic Cosmology*, Proc. Nat. Acad. Sci. **15**, 822–829. Derivation of most general line element by symmetry properties. Embedment in five-space.
3. R. C. Tolman, *On the Possible Line Elements for the Universe*, Proc. Nat. Acad. Sci. **15**, 297–304. Concludes Einstein and de Sitter worlds only possible static spherical symmetric ones.

1930

1. A. S. Eddington, *On the Instability of Einstein's Spherical World*, Monthly Not. Roy. Astron. Soc. **90**, 668–678. Proof of instability and discussion of Lemaître, 1927.1.
2. G. Lemaître, *On the Random Motion of Material Particles in the Expanding Universe. Explanation of a Paradox*, Bull. Astron. Inst. Netherlands **5**, No. 200, 273–274. Universe of constant mass for $k=+1$ when random motions are taken into account.
3. W. de Sitter, *The Expanding Universe. Discussion of Lemaître's Solution of the Equations of the Inertial Field*, Bull. Astron. Inst. Netherlands **5**, No. 193, 211–218. Universes of constant mass for $k=1$, $\lambda>0$. Effect of annihilation of matter. Geodesics.
4. ——, *Further Remarks on the Astronomical Consequences of the Theory of the Expanding Universe*, Bull. Astron. Inst. Netherlands **5**, No. 200, 274–276. Modification of 1930.3 in view of Lemaître, 1930.2. Effect of expansion on gravitational field of material bodies.
5. ——, *On the Distances and Radial Velocities of Extra-galactic Nebulae, and the Explanation of the Latter by the Relativity Theory of Inertia*, Proc. Nat. Acad. Sci. **16**, 474–488. Universe of constant mass for $k=1$, $\lambda>0$. Discussion of case $\lambda=\lambda_e$ of type A_2.
6. R. C. Tolman, *The Effect of the Annihilation of Matter on the Wave-Length of Light from the Nebulae*, Proc. Nat. Acad. Sci. **16**, 320–337. Derivation of (2.2) for $k=1$. Effect of annihilation of matter based on model $R=\exp(kt)$.
7. ——, *More Complete Discussion of the Time Dependence of the Non-static Line Element for the Universe*, Proc. Nat. Acad. Sci. **16**, 409–420. Effect of first three terms in expansion of R on logarithmic decrement of mass.
8. ——, *On the Estimation of Distances in a Curved Universe with a Non-static Line Element*, Proc. Nat. Acad. Sci. **16**, 511–520. Doppler effect. Effect of distance on luminosity; results in modification of definition of distance.
9. ——, *Discussion of Various Treatments which Have Been Given to the Non-static Line Element for the Universe*, Proc. Nat. Acad. Sci. **16**, 582–594. Comparative résumé of previous work.
10. H. Weyl, *Redshift and Relativistic Cosmology*, Phil. Mag. [7] 9, 936–943. Restatement of 1923.2 in coordinates employed by Lemaître, 1925.1, and Robertson, 1928.1.

1931

1. A. Einstein, *Zum kosmologischen Problem der allgemeinen Relativitätstheorie*, S.-B. preuss. Akad. Wiss. 1931, 235–237. Discussion of case $k=1$, $\lambda=0$, $p=0$. Expresses preference for $\lambda=0$.
2. O. Heckmann, *Über die Metrik des sich ausdehnenden Universums*, Nachr. Ges. Wiss. Göttingen 1931, 127–130. Suggests space may have zero or negative curvature (but see Friedmann, 1924.1; Robertson, 1929.2). Influence of random motion of matter on line element.
3. E. P. Hubble and M. L. Humason, *The Velocity-Distance Relation among Extra-galactic Nebulae*, Astrophys. J. **74**, 43–80. Empirical. Finds shift attributable to velocity of 558 km/sec. for every 10^6 parsecs.
4. M. von Laue, *Die Lichtfortpflanzung in Räumen mit zeitlich veränderlicher Krümmung nach der allgemeinen Relativitätstheorie*, S.-B. preuss. Akad. Wiss. 1931, 123–131. Wave, as opposed to geometrical, optics in such spaces.
5. G. Lemaître, *The Expanding Universe*, Monthly Not. Roy. Astron. Soc. **91**, 490–501. Finds condensations do not affect radius of universe in equilibrium (contrary to previous results of McCrea and McVittie—see 1931.6), but that a process of energy transfer (stagnation) does.
6. W. H. McCrea and G. C. McVittie, *The Expanding Universe*, Monthly Not. Roy. Astron. Soc. **92**, 7–12. Corrections to previous papers. Criticism of Lemaître, 1931.5.
7. W. de Sitter, *Some Further Computations Regarding Non-static Universes*, Bull. Astron. Inst. Netherlands **6**, No. 223, 141–145. Universes of constant mass for $k=1$, λ arbitrary.
8. R. C. Tolman, *On the Problem of the Entropy of the Universe as a Whole*, Phys. Rev. [2] **37**, 1639–1660. Generalized second law of thermodynamics and detailed application to universe filled with isotropic radiation.
9. ——, *Non-static Model of Universe with Reversible Annihilation of Matter*, Phys. Rev. [2] **38**, 797–814.

Detailed application of relativistic thermodynamics to universe consisting of monatomic gas in equilibrium with black-body radiation.

10. ——, *On the Theoretical Requirements for a Periodic Behavior of the Universe*, Phys. Rev. [2] **38**, 1758–1771. Non-existence of strictly periodic reversible universes (between nonzero limits). Quasi-periodic solutions of Einstein (p, $\lambda=0$), universe filled with radiation alone, universe with radiation and matter in equilibrium.

11. E. T. Whittaker, *On the Definition of Distance in Curved Space, and the Displacement of the Spectral Lines of Distant Sources*, Proc. Roy. London **133**A, 93–105. Definition of distance in accordance with astronomical usage (cf. Tolman, 1930.8). Application to de Sitter universe in Cayley-Klein coordinates.

12. J. H. Oort, *Some Problems Concerning the Distribution of Luminosities and Peculiar Velocities of Extragalactic Nebulae*, Bull. Astron. Institute Netherlands **6**, No. 226, 155–160. Disagrees with certain conclusions of Hubble and Humason 1931.3, and obtains tentative value of h half as large.

1932

1. A. Einstein and W. de Sitter, *On the Relation between the Expansion and the Mean Density of the Universe*, Proc. Nat. Acad. Sci. **18**, 213–214. Discussion of case $k=0$, $\lambda=0$, $p=0$.

2. O. Heckmann, *Die Ausdehnung der Welt in ihrer Abhängigkeit von der Zeit*, Nachr. Ges. Wiss. Göttingen 1932, 97–106. Plots all types of universes in which matter and radiation do not interact.

3. G. C. McVittie, *Condensations in an Expanding Universe*, Monthly Not. Roy. Astron. Soc. **92**, 500–518. Effect of condensation on radius. Revision of Jeans' work on condensation of matter into nebulae.

4. W. de Sitter, *On the Expanding Universe*, Proc. Akad. Wetensch. Amsterdam **35**, 596–607. Analysis and classification of all of Friedmann's worlds $p=0$.

5. R. C. Tolman, *Possibilities in Relativistic Thermodynamics for Irreversible Processes Without Exhaustion of Free Energy*, Phys. Rev. [2] **39**, 320–336. Establishes possibility of universe in which reversible processes take place without energy reaching a maximum or free energy a minimum.

6. —— and M. Ward, *On the Behavior of Non-static Models of the Universe when the Cosmological Term is Omitted*, Phys. Rev. [2] **39**, 835–843. Proves that universe for $k=1$, $\lambda\leqq 0$, $p\geqq 0$ is of type O.

7. H. Shapley and A. Ames, *A Survey of the External Galaxies Brighter than the Thirteenth Magnitude*, Ann. Astron. Obs. Harvard **88**, No. 2, 43–75.

For a general non-technical view of the field covered more thoroughly in this report the following accounts are to be recommended:

A. S. Eddington, *The Expanding Universe*, Proc. Phys. Soc. London **44**, 1–16 (1932).

E. P. Hubble, *The Exploration of Space*, The 1931 Vanuxem Lectures, delivered at Princeton University (to be published by Princeton University Press).

G. Lemaître, *La Grandeur de l'Espace*, Rev. d. Quest. Scientifiques, March 1929; *L'Expansion de l'Espace*, *do.* November 1932.

H. P. Robertson, *The Expanding Universe*, Science **76**, 221–226 (1932).

J. H. Reynolds, *Physical and Observational Evidence for the Expanding Universe*, Nature **130**, 458–462 (1932).

W. de Sitter, *The Expanding Universe*, Scientia **49**, 1–10 (1931). German translation available in Naturwiss. **19**, 365–369 (1931).

R. C. Tolman, *Models of the Physical Universe*, Science **75**, 367–373 (1932).