1984ARA&A..22..157E

Ann. Rev. Astron. Astrophys. 1984. 22: 157–84
Copyright © 1984 by Annual Reviews Inc. All rights reserved

ALTERNATIVES TO THE BIG BANG

G. F. R. Ellis

Department of Applied Mathematics, University of Cape Town, Rondebosch 7700, Cape, South Africa,[1] and Department of Physics, University of Texas, Austin, Texas 78712

1. INTRODUCTION

At issue is one of the most significant questions in astronomy: What is the nature of the beginning of the Universe? While there is at present a widespread opinion that only one view (the "standard hot big bang," or SHBB) gives a viable picture of the physical origin of the Universe, there are alternatives that need consideration in any dispassionate review of models compatible with current evidence.

The concept of an expanding universe evolving from an initial singular state arose from the work (1922–30) of A. A. Friedmann, G. E. Lemaître, and A. S. Eddington, together with E. P. Hubble's (1929) determination of a linear velocity-distance relation for distant galaxies. With the work of H. P. Robertson and A. G. Walker (1933–35), the spatially homogeneous, isotropic Friedmann-Lemaître-Robertson-Walker (or FLRW) universes were established as the standard universe models. The steady-state universe (1952) of H. Bondi, T. Gold and F. Hoyle posited continuous creation in an expanding universe without a beginning. However, the discovery (1965) of the cosmic microwave background radiation (CMBR) firmly established the SHBB as the accepted theory, accounting not only for that radiation and the velocity-distance relation, but also for the observed abundances of elements.

In recent times, (*a*) alternative theories have been proposed for the redshifts observed for galaxies and QSOs; (*b*) new physical theories (such as the grand unified theories, or GUTs) have resulted in a reappraisal of the

[1] Permanent address.

0066–4146/84/0915–0157$02.00

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

effective equation of state of matter in the early Universe; (*c*) new theories of gravity have been proposed as alternatives to Einstein's general theory of relativity; and (*d*) consideration has been given to more complex geometries than the FLRW models that underlie the standard picture.

In each case, new possibilities arise in which the origin of the Universe is different from that in the standard picture. It is important to consider such alternatives, particularly as there are a number of problems that are not resolved by the SHBB; without such an examination, the argument for the standard view is clearly incomplete. (One cannot make a rational choice between alternative explanations on the basis of an examination of only one of them.)

2. BASIC RESULTS

2.1 *The Nature of Redshift*

The standard theory assumes that electromagnetic radiation ("light" for short) travels on null geodesics ($x^a(\lambda): k^a{}_{;b}k^b = 0,\ k^a k_a = 0$, where $k^a \equiv dx^a/d\lambda$) in the space-time of general relativity and is characterized by the relations $c = \nu\lambda$, $E = h\nu$, relating the wavelength λ, frequency ν, speed c, and energy E of a photon. Then the redshift $z \equiv \Delta\lambda/\lambda$ (where $\Delta\lambda$ is the change in wavelength measured from spectra for light emitted at wavelength λ) is a direct consequence of measurable time dilation between the source and the observer (25, 91, 111). There can be various sources for this time dilation.

In the cosmological context, there is assumed to be a well-defined average motion of matter at each point in the Universe, representing the "standard" motion of matter there (24, 25); galaxies moving in this standard way, characterized by a four-velocity u^a (where $u_a u^a = -1$), are called "fundamental particles," and an observer moving with four-velocity u^a is denoted as a "fundamental observer." The *cosmological redshift* z_c is the redshift observed when both the source and observer move with this velocity (at their respective space-time positions) and there are no local gravitational redshift effects. However, there will be *local Doppler redshifts* z_{Ds}, z_{Do} if the source or observer is not moving at the standard velocity, and there may be *local gravitational redshifts* z_{Gs}, z_{Go} generated at the source or the observer by local gravitational potential wells. (Notice that in each case the effect could be a blueshift rather than redshift; this corresponds to a negative value of z, and we shall assume this possibility without explicitly mentioning it.) These contributions to the total redshift z are not distinguishable from each other by any direct measurement of the received radiation; the measurable quantity is the total redshift z, given in terms of

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

the other quantities by

$$(1+z) = (1+z_{Ds})(1+z_{Gs})(1+z_c)(1+z_{Do})(1+z_{Go}). \qquad 1.$$

The cosmological redshift depends on the path of the light ray from the source to the observer. One can distinguish two contributions. The first arises from the increase in relative distance of fundamental observers due to integrated local relative motion of matter [characterized by the expansion tensor $\theta_{ab} = h^c_a h^d_b u_{(c;d)}$ (24, 25), where $h_{ab} \equiv g_{ab} + u_a u_b$ projects into the rest-space orthogonal to u^a]. The second arises from the noninertial motion of the fundamental observers, characterized by a nonzero acceleration vector $\dot{u}^a \equiv u^a{}_{;b} u^b$ (indicating that their world-lines are nongeodesic). The change of redshift when moving a distance dl, measured by a fundamental observer along a null geodesic in direction $e^a = h^a_b k^b/(u_c k^c)$, obeys the equation (24, 25)

$$dz = d\lambda/\lambda = (\theta_{ab} e^a e^b - \dot{u}_a e^a)\, dl. \qquad 2a.$$

Thus one can write

$$(1+z_c) = (1+z_{Ec}+z_{Gc}), \qquad 2b.$$

where z_{Ec} is the usual Hubble (expansion) term (obtained by integrating the first term in Equation 2a), and z_{Gc} is what might be termed a "cosmological gravitational redshift" (obtained by integrating the second term).

In the FLRW models, it is assumed that the other redshift contributions are very small compared with the basic cosmological expansion effect. Then

$$z = z_c = z_{Ec}, \qquad 1+z = R(t_o)/R(t_e), \qquad 3.$$

where $R(t)$ is the "radius function" of the Universe that represents the time behavior of the distance between any pair of fundamental particles as the Universe evolves (by the homogeneity and isotropy of these models, the distances between all pairs of fundamental particles scale in the same way); and t_e, t_o are the times of emission and observation of the light, respectively. The observation of systematic redshifts indicates that the Universe is expanding at the present time t_o; so $H_o \equiv (R^{-1}dR/dt)|_o > 0$.

2.2 *Basic Equations*

The standard results rest on three pillars: the conservation equations, the Raychaudhuri equation, and the energy conditions. The form of these equations depends on the nature of the matter in the Universe. It is assumed here that the usual description of matter on a cosmological scale as a "perfect fluid" is adequate at most times. Then (24, 25) the stress tensor of

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

each component of matter or energy present takes the form

$$T_{ab} = \mu u_a u_b + p h_{ab}, \tag{4a}$$

where μ is the (relativistic) energy density of the fluid and p its pressure. To complete the description of matter, we must give suitable equations of state for the matter variables; in the hot early Universe, the "standard" theory assumes (except during phases of pair annihilation) that these are

$$p = (1/3)\mu, \qquad \mu = aT^4, \tag{4b}$$

describing a "Fermi gas" with temperature T (107, 109).

THE CONSERVATION EQUATIONS The stress tensor T_{ab} of matter that does not interact with other matter or radiation obeys the equations $T^{db}{}_{;b} = 0$. If this matter is a perfect fluid (Equation 4a), the energy conservation equation $u_d T^{db}{}_{;b} = 0$ takes the form

$$d\mu/dt + (\mu + p)(3/R)\, dR/dt = 0 \tag{5}$$

where t is proper time measured along the fluid flow lines, showing the effect of expansion or compression of the matter on its energy density. (The radius R is related to the volume V of a fluid element by the relation $V \propto R^3$.) The momentum conservation equation $h^a_d T^{db}{}_{;b} = 0$ is

$$(\mu + p)\dot{u}^a = -h^{ab}\partial p/\partial x^b, \tag{6}$$

which shows how spatial pressure gradients generate acceleration (and thus observable gravitational redshifts).

THE RAYCHAUDHURI EQUATION Contracting the Ricci identity $u^a{}_{;bc} - u^a{}_{;cb} = R^a{}_{dbc}u^d$ for u^a and using Einstein's field equations

$$R_{ab} - (1/2)Rg_{ab} = \kappa T_{ab}, \tag{7}$$

it follows that the volume behavior of the fluid is controlled by the Raychaudhuri equation (82, 24, 25, 45)

$$3R^{-1}d^2R/dt^2 = 2(\omega^2 - \sigma^2) + \dot{u}^a{}_{;a} - (\kappa/2)(\mu + 3p), \tag{8}$$

where ω^2 is the squared magnitude of the fluid vorticity and σ^2 the squared magnitude of the shear (or rate of distortion). Because d^2R/dt^2 is the curvature of the curve $R(t)$, this equation shows how rotation tends to hold matter apart and distortion to make it collapse, with $(\mu + 3p)$ the "active gravitational mass density" that tends to cause collapse.

ENERGY CONDITIONS Various "energy conditions" have been proposed as physical restrictions on the matter (45); for present purposes, the major

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

such conditions are

$$\mu + p \geq 0, \qquad 9a.$$

$$\mu + 3p \geq 0, \qquad 9b.$$

with Equation 9a needed for physical stability of matter (as follows from Equations 5 and 6) and Equation 9b implying that the gravitational effect of matter is attractive (see Equation 8). It is usually assumed that "reasonable" matter obeys these conditions.

2.3 *The Standard Models*

The FLRW universe models (109, 45) are isotropic and spatially homogeneous, so their shear, vorticity, and acceleration are all zero. Thus the Raychaudhuri equation (8) reduces to

$$3R^{-1}d^2R/dt^2 = -(\kappa/2)(\mu + 3p). \qquad 10.$$

By energy condition Equation 9b, this implies that the curve $R(t)$ of the radius function against time always bends down. As the Universe is expanding at the present time t_o, a singular origin must have occurred less than a Hubble time previously:

$$\exists\tau_o : |t_o - \tau_o| < 1/H_o, \qquad R \underset{t \to \tau_o}{\rightarrow} 0. \qquad 11a.$$

Because of the definition of $R(t)$,

(*A*) at this time τ_o, the distance between all fundamental particles goes to zero like $R(t)$.

At early times, Equations 4b and 5 show that $T = T_o(R_o/R)$, where (from present observations of the CMBR) $T_o \approx 3$ K; thus,

$$t \to \tau_o \Rightarrow T \to \infty, \mu \to \infty, p \to \infty. \qquad 11b.$$

This is the basic origin of the hot big-bang prediction in the FLRW universes:

(*B*) The Universe originates, and all physics breaks down, at the time $t \to +\tau_o$ where the temperature, energy density of matter, and space-time curvature diverge.

(For a more careful discussion, see 18.) Note that it is not merely the matter in the Universe that originates here: The space-time itself (and indeed the laws of physics) do not exist before this time. Now, Equations 4b, 5, and 10

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

show that

$$w(t) = \int_{\tau_0}^{t} R(t)^{-1}\, dt$$

converges; this implies that

(*C*) for each fundamental observer O, there exist particle horizons limiting possible communication with other fundamental observers in each direction at each time t; those further than the particle horizon will be unobservable by any radiation, and no causal influence from them can affect events in O's history up to the time t (45, 75, 85, 102).

The major properties characterizing the SHBB are (*A*)–(*C*) above and Equation 11, which follow from Equations 4, 5, 9, and 10. The standard results on the CMBR and element formation (i.e. the CMBR is relic radiation from the hot early phase when $T > 10^3$ K, and the light elements were synthesized when $T \approx 10^9$ K in the early Universe) can then be deduced from these properties (107, 109). If current ideas on the GUT are correct, at earlier times baryon creation will take place (42, 73, 116).

These models explain many features of the observed Universe in a satisfactory way, in particular the CMBR and primordial element abundances. However, they leave various other issues obscure, e.g. why the Universe is so nearly uniform, and yet not uniform [physical processes taking place after the creation of the Universe cannot be responsible because of the particle horizons (*C*); (38, 57, 64)], and why the density of matter is so close to the critical value separating recollapsing universes from ever-expanding universes (20, 38). Thus, one has the problem of explaining why the Universe started off from very special initial conditions. Additionally, the very concept of the creation of the Universe at such a singular beginning is philosophically objectionable to some scientists; they wish to find an alternative picture of its origin (23, 41, 102).

2.4 *The Cosmological Constant*

When the cosmological constant Λ is taken into account, the field equations (Equation 7) are altered to

$$R_{ab} - (1/2)Rg_{ab} + \Lambda g_{ab} = \kappa T_{ab}, \qquad 12.$$

which can be rewritten in the form

$$R_{ab} - (1/2)Rg_{ab} = \kappa T'_{ab}, \qquad 13a.$$

where

$$T'_{ab} \equiv T_{ab} - (\Lambda/\kappa)g_{ab}; \qquad 13b.$$

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

that is, the field equations can be regarded as the same as before but with an altered stress-energy tensor. Writing the effective stress-energy tensor T'_{ab} in the form of Equation 4a, one finds that

$$\mu' = \mu + \Lambda/\kappa, \qquad p' = p - \Lambda/\kappa, \qquad (\mu+3p)' = (\mu+3p) - 2\Lambda/\pi, \tag{14a}$$

and so Equation 10 becomes

$$3R^{-1}d^2R/dt^2 = -\frac{\kappa}{2}(\mu+3p)' = -\frac{\kappa}{2}(\mu+3p)+\Lambda. \tag{14b}$$

If the Λ-term is large enough, the argument by which Equation 10 leads to Equation 11 no longer applies, for the effective active gravitational density $(\mu+3p)'$ can violate the energy condition Equation 9b. If it does so at all times, a singularity in the past can be avoided (cf. Equation 14b), and indeed it has long been known (86) that there are FLRW universes with a large positive cosmological constant that (*a*) start expanding from a state asymptotically like the Einstein static universe [Eddington (23) suggested that this was preferable to a "big bang" origin]; or (*b*) collapse from infinity to a minimum radius value $R_{\min} > 0$ and then reexpand to infinity.

There are two problems with these models. Firstly, the deceleration parameter $q_o \equiv -(1/R)_o(d^2R/dt^2)_o/(H_o^2)$ must be negative at all times in these universes (for if it ever becomes positive, the universes must come from a singular origin), but observations suggest that the present value of q_o is positive (114). Secondly, an alternative explanation for the CMBR must be found in these cases, for it is not possible for this radiation to arise from a hot early stage in such a universe. The reason is that if such a hot early stage were to explain the CMBR, the asymptotic phase or turnaround caused by Λ would have to occur during the early plasma phase before decoupling when $T > 3000$ K; thus inequality Equation 9b would have to be violated by the effective matter (see Equation 14) for values of R less than $(R_o/1000)$. But at that time, $(\mu+3p)$ would be at least 10^9 times larger than at present, while Λ would have the same value as at present (because it is constant); violation of Equation 9b at that time therefore implies a value of Λ so large that we would undoubtedly have detected it from q_o. It is difficult to find alternative explanations for the CMBR (54). In addition, these models require element abundances to be arbitrarily set by initial conditions in the Universe (for the argument above shows the turnaround would occur long before nucleosynthesis temperatures were achieved). Thus, these models leave so many questions open as to not be presently viable.

One should note here that it follows from Equations 5, 9, and 14b that a static FLRW universe model will necessarily be unstable. [Perturbing R from a static value R_c will lead either to collapse or expansion (23, 25).] This implies that even if we could find some other explanation for redshift

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

(abandoning Equations 1 and 2) we would not expect a FLRW universe to be static unless we abandon either Einstein's field equations or the energy conditions.

2.5 *Implications*

The discussion above makes clear the nature of the available alternatives if one is to avoid the conclusion that the Universe originates in a SHBB. One can question

1. the nature of the observed redshifts (Equations 1–3) by adopting either a different theory of light propagation or a different astrophysical interpretation;
2. the conservation laws and/or gravitational field equations (Equations 5–7);
3. the nature of matter in the Universe, e.g. by assuming some effective contribution to the matter stress tensor that violates the energy conditions (Equation 9).

In each case one can avoid the existence of an initial singularity but must carefully consider if the resulting theory gives a satisfactory account of the microwave background radiation and element abundances. An additional possibility is to consider different geometric assumptions. One can question

4. the assumption of exact spatial homogeneity and isotropy. Then at least one of the shear, vorticity, and acceleration will be nonzero (so Equation 10 does not follow from Equation 8). Singularities will still occur in the past (45, 102), but they can be so different from the SHBB in their geometry and physics as to represent quite different initial situations.

3. ALTERNATIVE VIEWS OF REDSHIFT

3.1 *Alternative Theories of Light*

TIRED LIGHT An alternative theory of light propagation is the "tired light" theory, in which light loses energy progressively while traveling across large distances of extragalactic space. Thus we abandon Equation 2; then the observed redshifts might occur in a static universe, where no SHBB occurs. When a detailed mechanism for such an effect is proposed, various problems occur (65); and attempts to check if it is true are negative (34).

ALTERNATIVE GEODESICS Two-metric theories may involve light traveling on geodesics of another metric than that specifying length and time measurements; thus we effectively abandon Equation 2a. In general, these theories will still require different field equations or energy violation to avoid a singularity (see below). Furthermore, any such dynamic two-metric theory is likely to run into causality problems (77).

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

A particular case is Segal's theory (92), where in effect two metrics are laid down a priori from geometrical rather than physical principles (and are singularity free). The theory therefore proposes a cosmology independent of any gravitational equations; it demands a quadratic (magnitude, redshift) distance relation rather than the linear one predicted by the FLRW universes. The interpretation of the observations is the subject of a dispute centering on the nature of the selection effects and the statistics used (70, 93, 98); but the weight of the evidence seems to be against the quadratic effect. Additionally, the distribution of absorption lines in QSO spectra appears to rule these models out (81).

3.2 *Local Redshift Theories*

LOCAL CAUSES The effective proposal is that for QSOs, the major source of redshift is either z_{Ds} or z_{Gs} (13, 90); thus, Equation 1 remains true but we abandon Equation 3. Then the redshift observed for these objects is no longer necessarily a result of the expansion of the Universe. However, the instability of the Einstein static universe still would lead us to expect an expanding Universe if we stay within the general relativity/FLRW framework; and as usual, it is accepted that galactic redshifts are evidence of cosmological expansion. Thus, this proposal results in local variation of the explanation, rather than an overall abandonment of the concepts of an expanding universe and the hot big bang.

NEW LOCAL REDSHIFT PHYSICS In the second case, because of observed associations between galaxies and QSOs with differing redshifts (2, 33), the observed QSO redshifts are attributed to local effects of unknown nature: thus we abandon Equation 1 in the case of QSOs. This interpretation depends critically on the statistics of the image associations (117). Until the nature of this "new physics" is specified, this is not a theory in competition with the SHBB. As in the previous case, one again ends up with the plausibility of an expanding universe with a singular origin anyhow.

4. ALTERNATIVE PHYSICS

4.1 *Alternative Matter*

ENERGY VIOLATION There are various proposals of alternative effective equations of state for matter, which result in violations of the energy condition (Equation 9b) through the existence of sufficiently large negative pressures. Thus, as well as transient solutions, one also obtains steady-state and periodic solutions (48); it is possible to avoid the initial singularity, e.g. by starting from a static initial situation that then becomes unstable or in a universe that goes from a collapse phase to a state of expansion (86). We discuss in turn several ways of generating negative pressures.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

If the matter in the Universe has a bulk viscosity with the coefficient of bulk viscosity proportional to density, negative pressures arise such that energy violations sufficient to avoid an initial singularity occur (66, 113). However, it is not expected that this coefficient would have such a dependence on density in the case of ultrarelativistic fluids.

A classical conformal massive scalar field can generate violations of the energy conditions (45), and a massless conformal field coupled to pressureless matter can cause energy violations sufficient to cause a "bounce" instead of a singularity in a FLRW universe model (6). No such field is known to exist, but this situation may be regarded as analogous to the pion field, which mediates strong interactions in nuclear matter. The field would only be significant when very high densities had been attained, so the energy violation and bounce would occur for nuclear densities. Insofar as this is an adequate model of the fields dominant at these early times (which is unknown), the possibility arises of avoiding the singularity itself; indeed, if $k = +1$ the universe model could "bounce" at each of a series of singularities, as envisaged in Tolman's oscillating universe (103).

Energy violation may be expected to arise when quantum effects are significant, for such violations may be regarded as the driving force for the Hawking radiation emitted by black holes (14, 44). For example, if one uses as the source of the gravitational field the expectation value of a quantum field theoretical energy-momentum tensor for a quantized scalar field possessing mass, coherence effects can give rise to negative pressures sufficiently large to violate the energy conditions of the singularity theorems. Then a Friedmann-like collapse can be arrested at elementary particle dimensions and changed to an expansion (74).

A particular case where energy violation is expected to arise is in the inflationary universe scenario (38, 53) of the GUTs, where a supercooled metastable state of unbroken symmetry (a "false vacuum") gives rise to an effective cosmological constant Λ dominating other fields present. This in turn gives rise to an effective energy density $\mu' = \Lambda/\kappa$ and effective pressure $p' = -\Lambda/\kappa$; so $(p+\mu)' = 0$ (the limiting case of Equation 9a), which allows energy increase while conserving energy-momentum. (By Equation 5 the matter expands while its density stays constant, as in the steady-state universe, because of the negative pressures.) Now $(\mu+3p)' = -2\Lambda/\kappa$, so if $\Lambda > 0$, then a violation of the energy condition (Equation 9b) *must* happen, causing an exponential expansion ("inflation"). This expansion solves the causal problems raised by horizons in the Universe because causally connected regions are now much larger than a Hubble radius (38, 53). While this energy violation cannot lead to avoidance of the initial singularity if thermal equilibrium is maintained (39), this condition is not necessarily fulfilled (79), and there is a possibility that the Universe could bounce and

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

thereby avoid the initial singularity because of the energy violation (79). In the standard scenario, the GUTs phase is preceded by a radiation-dominated phase, with the usual singularity and horizons at early times; but one could conceive of the Universe as being created in the exponentially expanding state. Then its early properties will be those of the steady-state universe, without a curvature singularity or horizon.

Overall, it is clear that energy violation, in principle, could allow for avoidance of the initial singularity and permit a study of the Universe in its pre-expansion phase (101). When quantum effects dominate the equation of state, it is quite likely that such energy violation can occur.

OTHER EQUATIONS OF STATE There are other proposals for alternative equations of state that do not violate the energy conditions but that disagree with the radiation form (Equation 4b); one still has an initial singularity, but it can be a cold or cool big bang, rather than a hot one (e.g. 40, 47, 84). The SHBB statements (*A–C*) and Equation 11 remain, except that the temperature T does not diverge; thus the basic view of the origin of the Universe remains, with only relatively small details being varied.

4.2 *Alternative Gravity*

CLASSICAL THEORIES There are a large variety of classical (i.e. non-quantum) theories of gravity proposed as alternatives to the general theory of relativity (36, 112). In each case, one can work out the corresponding application of the theory to the Robertson-Walker metrics and so obtain a corresponding version of the origin of the Universe.

Many of these alternative theories can be written as close variants of general relativity, with field equations of the form

$$R_{ab}-(1/2)Rg_{ab}+\Lambda_{ab}=\zeta T_{ab} \tag{15.}$$

for some suitable choice of Λ_{ab} and ζ [which will be given in terms of other auxiliary fields, e.g. the scale function ϕ in scale-covariant theories such as the Brans-Dicke theory (12, 106, 109), the creation field C in the original Hoyle-Narlikar theory (46, 109), or the second metric γ_{ab} in Rosen's bimetric theory (87, 88); these auxiliary fields will satisfy their own field equations]. Now just as in the transition from Equation 12 to Equation 13, one can rewrite this equation in the form

$$R_{ab}-(1/2)Rg_{ab}=\kappa T'_{ab}, \tag{16a.}$$

where

$$T'_{ab}\equiv(\zeta T_{ab}-\Lambda_{ab})/\kappa; \tag{16b.}$$

that is, we can regard the new theory as described by Einstein's equations

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

with a new form of matter. Although T_{ab} satisfies the energy condition (Equation 9b), T'_{ab} may not; if so, one may (as in the case of alternative matter) find cosmological solutions with a nonsingular origin. A further feature is now of importance. As is well known, the tensor $(R_{ab}-(1/2)Rg_{ab})$ has vanishing divergence; and indeed this fact is responsible for the standard form of the field equations (Equation 7). Equations 15 and 16 show that

$$\Lambda^{ab}{}_{;b} = (\zeta T^{ab})_{;b}, \qquad 17a.$$

$$T'^{ab}{}_{;b} = 0 \qquad 17b.$$

Now Λ_{ab} may have a nonzero divergence or $\zeta_{,b}T^{ab}$ may be nonzero. Thus, although T'_{ab} satisfies the conservation equations, in general T_{ab} does not, so the energy or momentum of the matter fields are not conserved in this theory; this is the origin of the possibility of matter creation in theories such as the Hoyle-Narlikar theory. However, one should note the comment by Lindblom & Hiscox (52): This procedure effectively constructs nonconserved quantities within a conservative theory. One can rewrite such theories in conservative form by putting the equations in the form of Equation 16 (cf. 61).

If an alternative theory of gravity leads to sufficiently large equivalent energy condition violations, a singularity can be avoided. This is possible, for example, in the Hoyle-Narlikar theory (46), in theories with torsion (83, 105), and in bimetric theories (80, 87). If the gravitational field equations are derived from a Lagrangian more complex than the general relativity Lagrangian, in general we may expect energy violations to be possible. Thus, if we introduce terms quadratic in the curvature invariants, a "bounce" may be possible (67); when nonlinear Lagrangians are considered that are arbitrary functions $f(R)$ of the scalar curvature R, the singularity can be avoided through a suitable choice of $f(R)$ (4, 50); and if the singularity is not avoided, one may nevertheless avoid the occurrence of horizons (4). Again, if the matter fields are nonminimally coupled to the curvature, singularity avoidance is, in general, possible (71).

CONFORMAL THEORIES A class of alternative theories that have attracted particular attention are the variable-G or scale-covariant theories, such as Dirac's theory (21), the Brans-Dicke theory, the later Hoyle-Narlikar theories, and the theory of Canuto and coworkers. In general, different conformal gauges can be chosen in these theories that relate different metric representations through appropriate choice of a conformal factor (15, 89). This allows alternative explanations for observed redshifts; effectively one can switch terms in Equations 1 and 2, for when conformal transformations are allowed, one can no longer clearly distinguish between expansion and

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

acceleration of the fundamental particles. [These depend on the conformal frame used (66).] In fact, conformal transformations can be used to transform the isotropic expansion of FLRW fundamental observers to zero; so one can remove the singularities in the matter flow in FLRW models by appropriate conformal transformation, i.e. by appropriate choice of a conformal frame. (In fact, it is well known that all the FLRW universe models are conformal both to flat space-time and to the Einstein static universe.) The effective matter tensor in the field equations (Equation 16) will therefore, in general, violate the energy conditions, for this is what makes it possible to transform the expansion to zero without violating the Raychaudhuri equation. This may explicitly be seen to be true in the revised Dicke theory for $(2\omega+3) < 0$ (83). It is therefore not surprising that in some cases the massless scalar field in the Brans-Dicke theory can prevent the occurrence of initial singularities in the Universe (37). The remarks above lead us to expect that this will be a general feature of gauge-covariant theories.

QUANTUM COSMOLOGY When investigating the behavior of the Universe at very early times, a quantum cosmological model may be used to take quantum gravitational effects into account; however, the construction of such models is considerably hampered by the lack of a complete and manageable quantum theory of gravity. Space limitations do not allow an adequate discussion of such models in this article; rather, the reader is referred to a recent excellent review article and the references therein (43).

Overall, it is easy to find alternative theories of gravitation where energy violation may occur. There is no compelling reason to support these theories against general relativity, except that it is clear that in the high-density and curvature limit a quantum theory of gravity is ultimately needed. Thus the quantum cosmology results (when agreement is reached on them) are clearly of considerable significance for any discussion of the origin of the Universe. Singularities might be avoided in the quantum case, but if not, it seems possible that they will be less severe than in the classical case; and at least horizons may not occur in these models.

5. ALTERNATIVE GEOMETRIES

There are two basically different variations of the standard FLRW geometries: models that are spatially homogeneous but anisotropic, and models that are spatially inhomogeneous.

5.1 *Anisotropies*

The family of spatially homogeneous but anisotropic evolving universe models have been extensively studied (see 49, 55, 56 for recent reviews).

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

While these universes may at some stage in their history be very similar to the FLRW universes, they can have totally different singularity structures. The kinds of singularity that can occur depend on whether the universes are orthogonal or tilted.

ORTHOGONAL UNIVERSES These are spatially homogeneous universes in which the matter flow vector is orthogonal (in the space-time sense) to the surfaces of homogeneity; thus, there is no motion of the matter relative to these homogeneous space sections. The fluid flow is geodesic and rotation free, and the surfaces of homogeneity are surfaces of simultaneity for all the fundamental observers. The models comprise Bianchi orthogonal models (the general case) and Kantowski-Sachs universes (a special "rotationally symmetric" case).

In all these cases, because $\omega = \dot{u}^a = 0$, the existence of a singularity follows directly from the energy conditions (Equation 9) and the Raychaudhuri equation (Equation 8); in essence, the Raychaudhuri equation shows that the normals to the surfaces of homogeneity must approach a point of intersection, but the fluid is moving along these normals and so runs into a singularity where the density of matter diverges.[2] The existence of the distortion term σ^2 in Equation 8 shows that the singularity is in general worse than in the FLRW case; indeed, usually the singularity is shear dominated, the effect of the matter term being negligible at early times. Limits on the amount of anisotropy come from observed limits on the anisotropy of the CMBR and from observed light-element abundances, compared with model predictions (55, 89, 100).

In these models, many different kinds of singularity structure can occur. In the simplest (Bianchi I) case, the solution resembles the vacuum Kasner solution; the singularity may be a "cigar" or "pancake" singularity (55, 100). If the cigar singularity occurs, horizons are broken in one spatial dimension (45, 64). Generically (in Bianchi VIII or IX) there will be oscillatory behavior, with a series of Kasner-like epochs taking place between "bounces" off a potential wall (7, 64). These possibilities have been discussed in a previous volume of this series (72). It seems agreed that the oscillatory behavior continues right back to the origin of the Universe, so that the initial behavior of the models can be characterized in a strict sense as "chaotic" (3); unfortunately, this does not suffice to break the horizons in all directions, as was once hoped (57).

While these singularities are geometrically quite different from the SHBB through their anisotropic expansion (and in many cases oscillatory behavior), which modifies (*A*), the essential features (*B*) remain; the most

[2] See (18) for a careful discussion of the Bianchi cases and (16) for the Kantowski-Sachs case.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

significant difference is the possibility of horizon breaking in some directions [so that (*C*) no longer holds in all directions].

TILTED UNIVERSES In this case, the surfaces of homogeneity are not orthogonal to the matter flow. Thus, although they are strictly spatially homogeneous in a rigorous sense, these universes appear inhomogeneous to the fundamental observers, who move relative to the homogeneous spatial sections (35). One can view this either as a result of their surfaces of instantaneity not coinciding with the surfaces of homogeneity (51) or because Doppler effects due to the peculiar velocity of the observer make a uniform distribution of matter appear anisotropic (30).

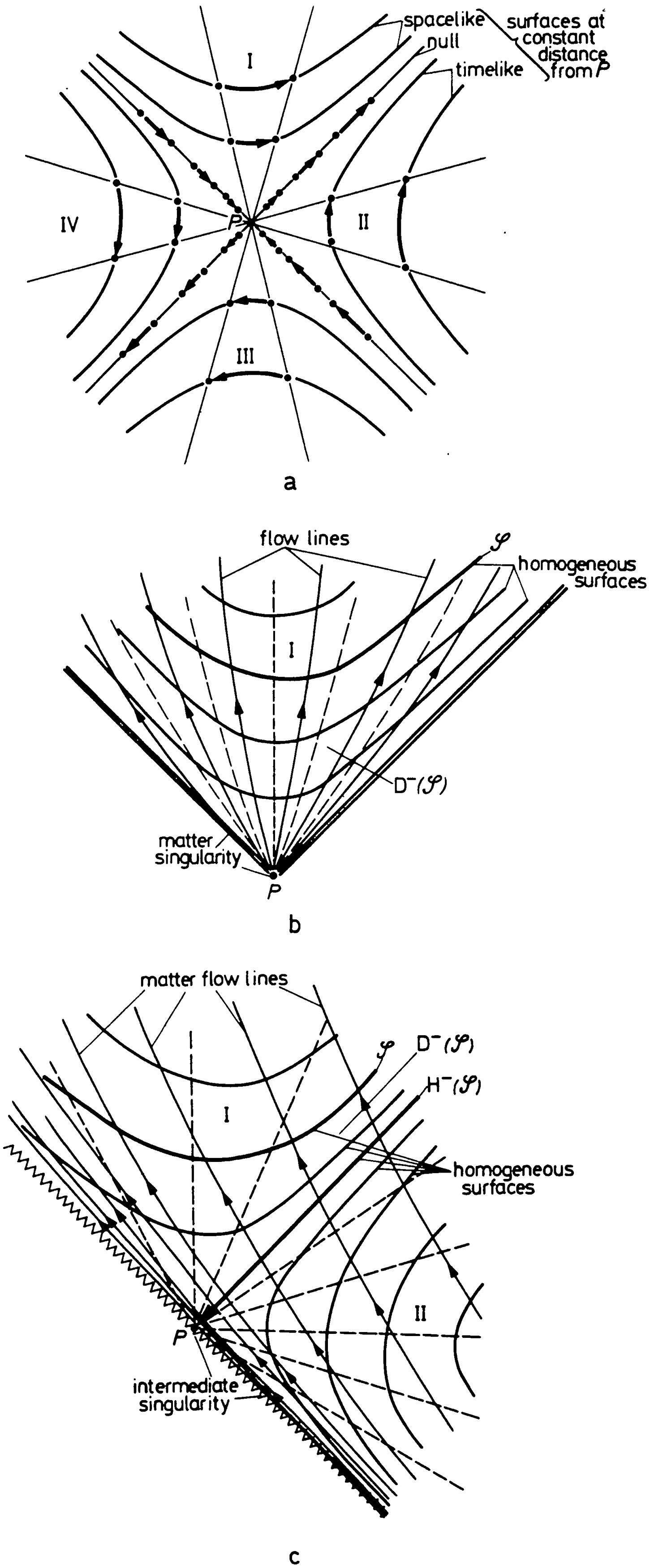
In tilted universes the matter can move with vorticity and will have nonzero acceleration unless the pressure is constant. Thus, the Raychaudhuri equation for the fluid no longer directly implies that a singularity occurs. However, the Raychaudhuri equation for the normals to the surfaces of homogeneity implies that these normals become singular within a finite proper distance; this result shows that the spatially homogeneous part of the Universe must end within a finite time. A careful analysis then demonstrates that the Universe must be singular at this time (45).

There are two rather different ways the spatially homogeneous region of the Universe can end in the past. (*a*) The first possibility is that an infinite density (spatially homogeneous) singularity can bound this region; then the situation is essentially the same as in the case of the orthogonal universes, except that the geometry can be more complex (because of the rotation and acceleration of the matter). The Universe originates at this infinite density singularity. (*b*) The second possibility (94) is that the density is finite at the boundary of the spatially homogeneous region of space-time (which represents late times in the Universe). The matter originates at early times in a spatially inhomogeneous, stationary region of space-time, separated from the spatially homogeneous region by a (Cauchy) horizon (18, 31).

It is easiest to understand this situation as follows (26, 27): consider the action of boosts of the Lorentz group through a hyperbolic angle β in two-dimensional Minkowski space-time (Figure 1*a*). If a single matter flow line through the origin *P* is spread over the space-time by repeating this action infinitely often, the (infinite) family of flow lines thus generated intersect at *P* and we have a model of Case (*a*): the spatially homogeneous region I comes to an end at a boundary where the matter density is infinite. This region comprises the entire Universe (Figure 1*b*). Suppose, however, a single flow line of matter that crosses from region I to region II is spread around by the same action of the Lorentz group. Then the surfaces of constant density change from being spacelike to being timelike as the fluid crosses the horizon $H^-(\mathscr{S})$ (Figure 1*c*), so while the late part I of the model is spatially homogeneous as before, the early part II is spatially inhomogeneous. This

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E



© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

provides a model of Case (*b*). The limiting value of the density at the boundary of these regions of space-time, shown by a zigzag line in Figure 1*c*, is finite: it is the same as the density on the horizon $H^-(\mathscr{S})$ (which is a surface of homogeneity). However, the Universe cannot be extended beyond this boundary, for a singularity occurs here. One can see this by noting that in Figure 1*c* the light cone of any point is at 45°. The past null geodesic to the left from any point in the space-time runs into the boundary in a finite distance, but in that distance it crosses an infinite number of world lines. Consequently, the density of matter measured in a frame of vectors parallel propagated down this null geodesic will diverge. This parallel-propagated frame is related to a group-invariant frame by a Lorentz transformation that diverges at the boundary; and the density is perfectly finite at each point if measured in this group-invariant frame. Thus, this singularity is a nonscalar singularity (45) [colloquially known as a "Whimper" (31)]. All curvature and matter invariants are regular as one approaches the boundary; the divergence occurs when the Ricci and matter tensors are measured in a parallel-propagated frame.

This model gives a good representation of what happens in the full four-dimensional Case (*b*) tilted Bianchi cosmologies. The spatially homogeneous region is bounded by a Whimper singularity on the one hand, and a Cauchy horizon $H^-(\mathscr{S})$ on the other; the fluid crosses the Cauchy horizon from an earlier spatially inhomogeneous region, which is also partially bounded by the Whimper singularity. This singularity is causally null; correspondingly (45, 75), there will be no particle horizon in the direction of the Whimper singularity (but there will be one in the opposite direction; cf. the causal diagrams in 18). It is a "mild" singularity in that the temperature and density of the matter is finite there. The fluid itself originates at a second singularity (not shown in Figure 1*c*) that is timelike. Here, in general, the density will diverge, but this will not always be the case: situations are known (18) where the density is finite at this initial singularity (but where the conformal geometry is singular). Thus, in such exact solutions of Einstein's equations—which have reasonable equations of state and obey the energy conditions—two singularities occur, but there is a finite maximum density for matter in the Universe; hence, there is also a bound to the temperature that will occur.

The plausibility of these universe models rests on their ability to give

Figure 1 Two-dimensional Minkowski space-time as a model for singularities. (*a*) The action of "boosts" about *P* is shown. (*b*) A matter line through *P*, moved over the space by the boosts about *P*, leads to a flow diverging from *P*. A matter singularity occurs. (*c*) A matter line not through *P*, moved over the space by the boosts about *P*, leads to a flow that "piles up" toward *P* but that does not converge there. A nonscalar singularity occurs.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

reasonable predictions for element production and compatibility with the microwave anisotropy. Initial investigation (31, 60) is encouraging; it seems that current observational evidence is quite compatible with such singularities. It should be mentioned that it is known that these models are of zero measure in the set of all spatially homogeneous models (95, 96), and this fact has been used as an argument against their occurrence. While this is an interesting line of argument, its status is unclear because of the uniqueness of the Universe. Additionally, it should be mentioned that if one pursues this line, the SHBB FLRW models are an even smaller subset of the spatially homogeneous models and thus, according to this argument, are even more unlikely to occur than the Whimpers.

OTHER EFFECTS As in the previous case, one can consider spatially homogeneous, anisotropic universes with other equations of state or governed by other gravitational theories. Because of the shear in these cosmologies, in many cases (e.g. for orthogonal fluids) one may expect the singularities to be worse than in corresponding FLRW universes. Thus, for example, while torsion can prevent singularities in FLRW universes, it cannot do so in Bianchi I universes (99).

Where quantum field effects seem to make a significant difference is in the evolution forward in time from the singularity; for in anisotropic universes there will be significant pair creation of conformally invariant particles, and this will dissipate anisotropy and so drive the Universe toward a state of isotropy (43, 115). Similarly, the effect of a false-vacuum inflationary expansion in a Bianchi universe will be to make it evolve exponentially toward the de Sitter (isotropic) solution (108). However, if the anisotropy in the Universe is too large at the end of the quantum gravity era, the inflationary phase will never take place (5a). Thus, quantum effects can contribute to a limited extent toward Misner's (63) vision of a situation where any anisotropy or nonuniformity will be dissipated by physical processes in the early Universe, so giving an "explanation" of presently observed isotropy.

5.2 *Inhomogeneities*

As the fluid can now move with acceleration and rotation, the Raychaudhuri equation does not directly imply the existence of a singularity in the past. However, the powerful Hawking-Penrose singularity theorems (45, 102) show that as long as suitable energy and causality conditions are satisfied, singularities will indeed occur in a general inhomogeneous universe model [essentially because the matter that thermalized the microwave background radiation has sufficient energy to cause a refocusing of all the null geodesics that generate our past light cone (45)].

In many cases the singularity will be essentially of the same character as

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

in the spatially homogeneous case, often showing the same kind of oscillatory behavior as the Bianchi IX cosmologies (7, 9; but see 5, 8, 57). The major difference is that one can now additionally get timelike singularities.

TIMELIKE INITIAL SINGULARITIES The existence of inhomogeneities at early phases in a universe that is like a FLRW universe at late times implies that "the initial big bang is not necessarily simultaneous. One can formulate initial conditions in such a way that expansion of some parts of the universe are delayed" (116). It has been proposed that such "lagging cores" could be responsible for violent astrophysical phenomena such as QSOs (69, 71a). The situation could be the time-reverse of the collapse to a black hole: this case is called a "white hole" (71a). In this case, the singularity will be surrounded for most of its life by an absolute particle horizon, into which no particles may fall but from which particles may be ejected (76). Difficulties arise concerning the existence of "white holes," particularly in view of the particle creation process that would tend to make them rapidly evaporate away (76, 78).

However, there appears to be nothing preventing the existence of more general timelike singularities that do not occur instantaneously, but continue over a time (22), and that are not associated with such horizons (104); the timelike part of the singularity may have the nature of either a negative- or positive-mass singularity (62, 104). A new dynamics now occurs, for the singularity both emits matter and radiation to the Universe and receives them from it; thus, it both acts on and can be acted upon by the Universe. This relation of the singularity to the Universe is quite unlike that in the spacelike singularity case where the singularity emits matter and radiation to the Universe, and indeed determines all its initial conditions, but cannot be affected by it. One now has the possibility of investigating the dynamics of interaction of the singularity with the Universe.

These timelike singularities can occur as inhomogeneous segments of a singularity that is spacelike elsewhere (62, 104). Indeed, it is clear that unless one chooses very special conditions in the Universe (by setting to zero the "decaying modes" of density perturbations), inhomogeneities will dominate at early stages (97). Penrose argues on entropy grounds that the early Universe should obey such rather special conditions (77, 78); his argument depends on probability assumptions about the set of all possible universes—assumptions that may be queried in view of the uniqueness of the Universe.

The most intriguing possibility is if one or more isolated timelike singularities occur, with most of the matter passing between them rather than originating at them (45). This situation is closely related to the hope that many people have that the present expanding phase of the Universe

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

should result from a previous collapse phase, with the formation of isolated singularities and subsequent reexpansion. Then the hot big bang phase is like an inhomogeneous SHBB but is preceded by a previous collapsing phase. The overall picture is like "cyclic" FLRW universes collapsing from infinity or from a previous expansion phase and then reexpanding (with most of the matter avoiding the singularities; these must form because the energy conditions are satisfied).

We do not have exact cosmological models where this situation occurs; but two examples indicate it may be possible, namely the Reissner-Nordstrom solution with test matter in it (45), and certain classes of Whimper singularities where two horizons may occur (31) (if these exist; no exact solutions of this kind are known at present). In both these cases the situation is probably unstable, but they show that it is not obviously prohibited by the field equations. If it does occur, the conditions of the present phase of expansion are determined by the two-way interaction between the Universe and the singularity formed from the previous collapse phase; no energy violations or quantum effects are needed to cause such a "bounce."

ESSENTIALLY INHOMOGENEOUS UNIVERSES These are universe models containing no epochs where the space-time is approximately spatially homogeneous; that is, they are never like a FLRW universe model. This assumption, of course, goes against the usual "cosmological principle," which states a priori that the Universe is spatially homogeneous (109); but that is an unverified and, indeed, possibly unverifiable (28) philosophical statement that could be incorrect.[3]

One possibility is asymptotically flat models—an isolated dense universe surrounded by empty space. An example of this kind is the recent Narlikar & Burbidge (68) proposal, in which all detectable matter lies in an expanding super-supercluster. While this kind of situation is a possibility, many features remain to be explained, e.g. the element abundances and the microwave background radiation. The specific Narlikar-Burbidge version appears to be ruled out by QSO absorption-line measurements (81).

A related concept is that of a hierarchical universe, where clustering occurs over all possible scales, so that each observer is within an effective spherical inhomogeneity but the average density of matter goes to zero in larger and larger volumes. In this model the Universe is spatially homogeneous (19). However, this is difficult even to describe mathematically (as the description used depends completely on the scale of averaging); it is not clear that the models proposed so far (10, 110) in fact represent the hierarchical situation fully.

[3] It is false in the currently fashionable inflationary universe scenario, where major inhomogeneities occur at the bubble walls.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

If we abandon statistical homogeneity, the only universe models that give isotropic observations[4] are inhomogeneous, spherically symmetric models, where we are near a center of symmetry (11). In order to explain the microwave background radiation in such models, an intriguing possibility is that there are *two* centers of symmetry in the Universe, with one in our neighborhood and the other at the location of a timelike singularity. This singularity would be surrounded by a massive fireball emitting blackbody radiation; thus the space sections of the Universe would look like Figure 2*b*, with the fireball—like a supermassive star—at the antipodes. A space-time diagram can be drawn as in Figure 2*a*. The timelike singularity at the center of the fireball would continually emit radiation and particles into the Universe as well as receive them from the Universe—it would be dynamically interacting with the Universe (as discussed above). A remarkable duality between such models and the FLRW universes is possible, with each stage of the FLRW universes that occurred in the past as a result of time evolution also occurring in these universes spatially as a result of spatial inhomogeneity (32).

Perhaps the most radical such proposal is the completely static two-centered universe possibility (17, 32), where the redshifts observed for distant galaxies are purely cosmological gravitational redshifts (cf. Equation 2b). High pressures are needed in these models to cause the acceleration that underlies the gravitational redshift; this is not impossible but is perhaps uncomfortable. Two problems arise: Firstly, because of the spherical symmetry, the (redshift, distance) law in such a universe must be quadratic at the center, so (as in Segal's theory discussed in Section 3.1) we can test this model by seeing if a quadratic relation is indeed observed. The usual data show a linear relation, contrary to expectation in such universes; however, these data are analyzed on the assumption the Universe is a FLRW universe, and selection effects play a crucial role in determining which is the correct law. (Essentially the same analysis will apply here as in the Segal theory.) The data still have not been examined using an analysis that takes surface brightness effects fully into account. Secondly, philosophical problems may arise because of the Earth being near the center in these models. However, these qualms can be offset by noting that physical conditions can favor the occurrence of life at this center in such universes and by considering the a priori improbability of the FLRW universe models (27a). On the positive side, it is intriguing to note that the background radiation, still our best evidence for isotropy of the Universe, would appear isotropic at *every* point in such universes (because its temperature will be determined by the ratio of gravitational potentials at the points of emission and reception of the radiation; and this is

[4] See, for example, (55) for a summary of the evidence that the Universe is isotropic about us.

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

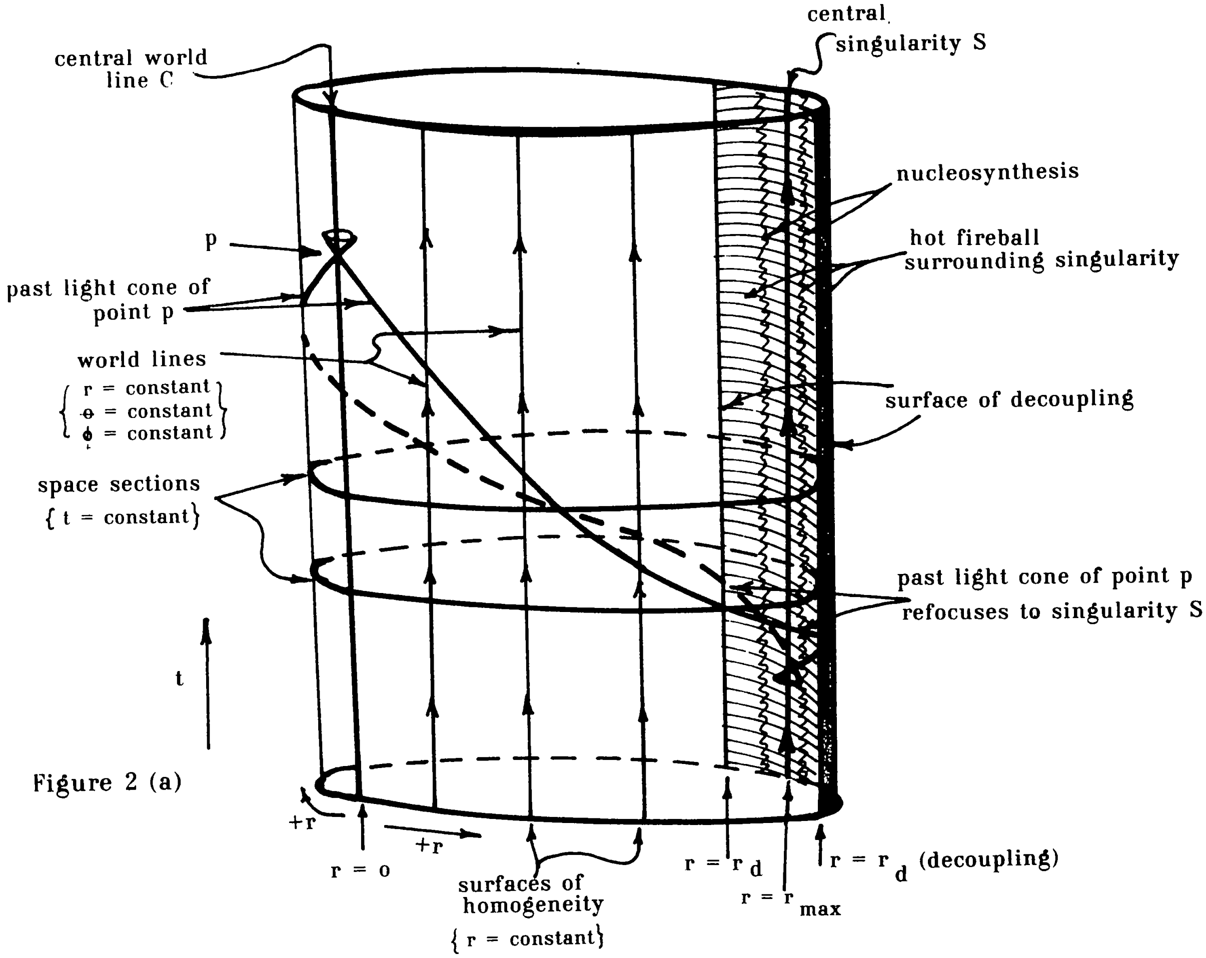


Figure 2 (a)

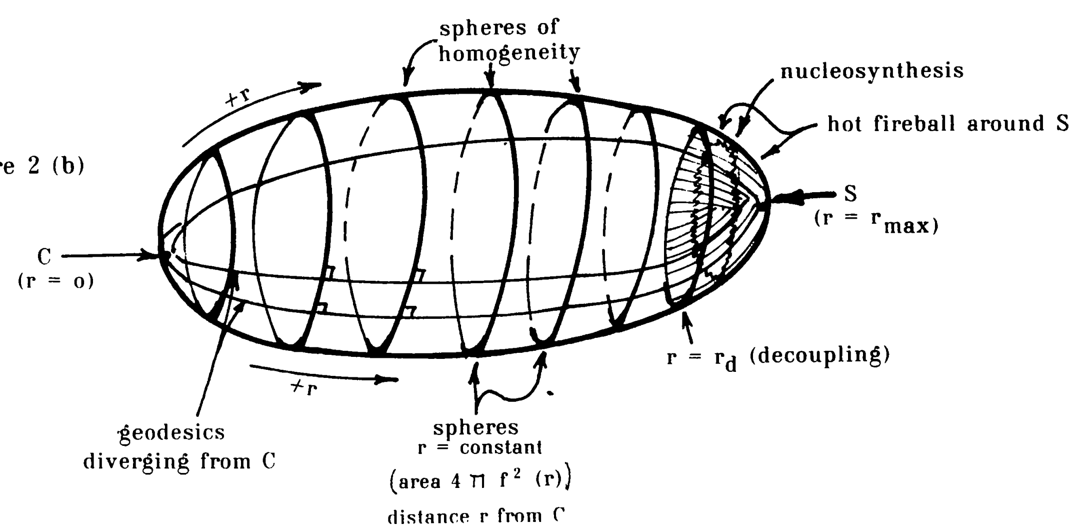


Figure 2 (b)

1984ARA&A..22..157E

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

independent of the path the light follows to the observer). Thus, observed isotropy of this radiation does not, in this case, imply that the Universe is isotropic about our position. On the other hand, isotropy of galactic redshifts would carry this implication.

The major interest of this model is the way in which it highlights the quite different role the singularity can play as a continuing influence in the evolution of the Universe. It is not a once-and-for-all event, but is instead a continuing source and sink of matter and information; in this case, it enables a steady state. The singularity has a different *meaning* than in the FLRW universe.

While it is difficult to make a convincing case for a static universe of this kind, it seems likely that expanding versions maintaining the essential features of Figure 2 could account for all present observations. If the expansion originated in a singularity, the model would then be essentially like the delayed core models mentioned above; but the expansion could, for example, conceivably be preceded by a stationary phase or by a contraction phase to some minimum radius with the singularity always localized in one region of the Universe.

FURTHER EFFECTS As in the previous cases, one can consider the nature of inhomogeneous cosmologies if energy violation, other gravitational theories, or quantum effects are taken into account.

Perhaps the most intriguing new possibility is that the energy conditions may be violated for essentially geometric reasons. The basic point here is that the Einstein field equations are tested, and believed to hold, on a particular scale (say, that of the solar system). In an inhomogeneous situation it is not clear that the geometrically averaged space-time metric representing the situation at other scales (say, that of clusters of galaxies) will obey the same field equations; indeed, one would expect correction terms (similar to the polarization terms in electromagnetic theory) to allow for the spatial averaging that is taking place (29). However, there is no guarantee that these terms will obey the energy conditions; and, indeed, in a turbulent situation they might be able to cause effective negative energy densities so large as to avoid a singularity (59). Thus the space-time metric averaged out to describe the Universe at the larger scale could be singularity free (but small-scale singularities could still occur, e.g. local black holes resulting from gravitational collapse of stars).

Figure 2 A two-centered universe, with our Galaxy near one of the centers. (*a*) Space-time diagram of the Universe (two spatial dimensions suppressed). It is a cylinder, with a timelike singularity at the other center. Our past light cone is shown. (*b*) Space section of the Universe (one spatial dimension suppressed). The singularity lies behind a fireball, which is like a supermassive star "over there."

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

6. CONCLUSION

If we assume the standard FLRW geometries, then different equations of state or field equations resulting in sufficiently large effective energy violation imply the possibility that there is no singularity at the beginning of the present phase of expansion of the Universe. This expansion could start from a stationary state (either a static phase or an exponentially expanding phase) or could be a bounce from a previous collapse. Any "bounce" that occurs must have taken place at very early times, if our present picture of the origin of the microwave background radiation and the elements is correct. It is plausible that quantum field effects could cause a bounce at extremely early times, but the foundations of quantum cosmology need clarification before such a statement is on firm footing; if this does not happen, the quantum field effects probably at least make the singularity less severe (e.g. by eliminating particle horizons).

A different geometry implies the possibility of quite different kinds of singularity, which entail different concepts of creation. Finite density singularities can occur at the beginning of the Universe in the case of tilted homogeneous cosmologies, lessening the severity of the initial singularity. (One must have some real singularity at any "beginning" of the Universe, for otherwise we can continue the Universe past this event to earlier times, thus showing that it was a "fictitious" singularity rather than a real one.) A major difference in the case of inhomogeneous universes is the possibility of timelike singularities, which continue to interact with the Universe over a period of time—perhaps for the entire history of the Universe—rather than the once-and-for-all interaction of the singularity in the standard theory, where creation takes place at one time and then the singularity is only an incident in the past.

There are two kinds of steady-state universe that seem to have the potential of giving a reasonable description of the astrophysical evidence: the original steady-state universe of Bondi, Gold, and Hoyle; and the two-centered static universe of Ellis, Maartens, and Nel. In both cases creation is a continuously proceeding process—in the first case diffused through space and in the second localized at a singular center of the Universe. However, both of these models run into difficulties when compared in detail with observational evidence.

The next major possibility is that the Universe started off in a steady-state situation and then changed to an evolutionary phase, as in the Eddington-Lemaître expansion from an Einstein static state or the Starobinski expansion from a de Sitter steady-state phase; there may also be suitable models expanding from an Ellis-Maartens-Nel static phase. A major

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

problem with any such stationary-state origin is how the Universe chooses when to break this phase and start the present evolutionary phase: "Since the initial metric has a finite life-time, it could not exist as $t \rightarrow -\infty$, and therefore some other metric should exist before it" (53).

If a bounce takes place, it either comes from a state that has existed for indefinitely long in the past, perhaps collapsing from an infinite radius; or else it is a rebounce following a previous state of expansion from a state of high density, as in the "phoenix" universes that perpetually oscillate, and in many ways give an attractive understanding of the history of the Universe (20, 58). Each rebirth can take place either through effective energy violation or quantum effects (when no singularity occurs) or through the occurrence of isolated singularities that are sidestepped by most of the matter and that interact with the Universe for a brief period.

A singularity will exist unless effective energy violation takes place; it will provide an initial boundary to the Universe as we know it. The singularity could be "weak," so that quantum mechanically one can follow the Universe through to the era preceding the initial singularity; that is, one may be able to provide a theory of creation. Any such theory of creation of the Universe must by necessity postulate some preexistent structure that provides the basis for the equations used to describe creation (e.g. a previously existing oscillating universe, Minkowski space, a de Sitter universe, superspace, or pregeometry). These issues cannot be pursued further here; we refer the reader to the Appendix of Linde's (53) paper and the references therein.

There are then a variety of possibilities available, both within standard general relativity (where a variety of geometries are considered) and in variations of that theory, for understanding the nature of the origin of the Universe. The SHBB is one of this family but is by no means the only conceivable member; further possibilities are not completely excluded experimentally or theoretically.

Literature Cited

1. AAPT. 1979. *Cosmology*, ed. L. C. Shepley, A. A. Strassenburg. SUNY Stony Brook, N.Y.: Am. Assoc. Phys. Teach.[1]
2. Arp, H. C. 1980. *Ann. NY Acad. Sci.* 336:94–112
3. Barrow, J. D. 1982. *Phys. Rep.* 85:1–49
4. Barrow, J. D., Ottewill, A. C. 1983. *J. Phys. A* 16:2757–76

[1] Various key papers on cosmology are reprinted in this volume; they are indicated in this list of references by the designation "AAPT".

5. Barrow, J. D., Tipler, F. J. 1979. *Phys. Rep.* 56:371–402
5a. Barrow, J. D., Turner, M. S. 1981. *Nature* 292:35–38
6. Bekenstein, J. D. 1975. *Phys. Rev. D* 11:2072–75
7. Belinsky, B. A., Khalatnikov, I. M., Lifshitz, E. M. 1970. *Adv. Phys.* 19:525–73
8. Belinsky, B. A., Khalatnikov, I. M., Lifshitz, E. M. 1980. *Phys. Lett. A* 77:214–16; 83:321
9. Belinskii, V. A., Khalatnikov, I. M.,

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

Lifshitz, E. M. 1982. *Adv. Phys.* 31:629–67
10. Bonnor, W. B. 1972. *MNRAS* 159:261–68
11. Bonnor, W. B. 1974. *MNRAS* 167:55–61
12. Brans, C., Dicke, R. 1961. *Phys. Rev.* 124:925–31 (AAPT)
13. Burbidge, E. M. 1967. *Ann. Rev. Astron. Astrophys.* 5:410–11, 433–37
14. Candelas, P., Sciama, D. W. 1977. *Phys. Rev. Lett.* 38:1372–75
15. Canuto, V. M., Narlikar, J. V. 1980. *Ap. J.* 236:6–23
16. Collins, C. B. 1977. *J. Math. Phys.* 18: 2116–24
17. Collins, C. B. 1983. *J. Math. Phys.* 24: 215–19
18. Collins, C. B., Ellis, G. F. R. 1979. *Phys. Rep.* 56:65–105
19. de Vaucouleurs, G. 1970. *Science* 157: 1203–13 (AAPT)
20. Dicke, R. W., Peebles, P. J. E. 1979. In *General Relativity: An Einstein Centenary Survey*, ed. S. W. Hawking, W. Israel, pp. 504–17. Cambridge Univ. Press
21. Dirac, P. A. M. 1938. *Proc. R. Soc. London Ser. A* 165:199–208
22. Eardley, D., Liang, E., Sachs, R. K. 1972. *J. Math. Phys.* 13:99–107
23. Eddington, A. S. 1930. *MNRAS* 90: 668–78
24. Ehlers, J. 1961. *Abh. Akad. Wiss. Lit. Mainz, Math. Naturwiss. Kl.* 11:791–837
25. Ellis, G. F. R. 1971. In *General Relativity and Cosmology*, ed. R. K. Sachs, pp. 104–82. New York: Academic
26. Ellis, G. F. R. 1978. *Comments Astrophys. Space Phys.* 8:1–7
27. Ellis, G. F. R. 1975. *Ann. NY Acad. Sci.* 262:231–40
27a. Ellis, G. F. R. 1979. *Gen. Relativ. Grav.* 11:281–89
28. Ellis, G. F. R. 1980. *Ann. NY Acad. Sci.* 336:130–60
29. Ellis, G. F. R. 1984. *GR10 Conf. Proc.*, ed. A. Pascolini, et al. In press
30. Ellis, G. F. R., Baldwin, J. 1984. *MNRAS* 206:377–81
31. Ellis, G. F. R., King, A. R. 1974. *Commun. Math. Phys.* 38:119–56 (AAPT)
32. Ellis, G. F. R., Maartens, R., Nel, S. D. 1978. *MNRAS* 184:439–65
33. Field, G. B., Arp, H., Bahcall, J. N. 1973. *The Redshift Controversy*. Reading, Mass: Benjamin
34. Geller, M., Peebles, P. J. E. 1972. *Ap. J.* 174:1–5
35. Gödel, K. 1952. *Proc. Int. Congr. Math.*, ed. L. M. Graves, et al., pp. 175–81. Providence, R.I.: Am. Math. Soc.
36. Goenner, H. 1984. *GR10 Conf. Proc.*, ed. A. Pascolini, et al. In press
37. Gurevich, L. E., Finkelstein, A. M., Ruban, V. A. 1973. *Astrophys. Space Sci.* 22:231–42
38. Guth, A. H. 1983. In *The Very Early Universe*, ed. G. W. Gibbons, S. W. Hawking, S. T. C. Siklos, pp. 171–204. Cambridge Univ. Press
39. Guth, A. H., Sher, M. 1983. *Nature* 302:505–6
40. Harrison, E. R. 1970. *Nature* 228:258–60
41. Harrison, E. R. 1981. *Cosmology*. Cambridge: Cambridge Univ. Press
42. Hartle, J. B. 1981. In *Quantum Gravity 2*, ed. C. J. Isham, R. Penrose, D. W. Sciama, pp. 313–28. Oxford Univ. Press
43. Hartle, J. B. 1983. In *The Very Early Universe*, ed. G. W. Gibbons, S. W. Hawking, S. T. C. Siklos, pp. 59–89. Cambridge Univ. Press
44. Hawking, S. W. 1975. In *Quantum Gravity: An Oxford Symposium*, ed. C. J. Isham, R. Penrose, D. W. Sciama, pp. 219–67. Oxford Univ. Press
45. Hawking, S. W., Ellis, G. F. R. 1973. *The Large-Scale Structure of Space-Time*. Cambridge Univ. Press
46. Hoyle, F., Narlikar, J. V. 1964. *Proc. R. Soc. London Ser. A* 278:465–78
47. Huang, K., Weinberg, S. W. 1970. *Phys. Rev. Lett.* 25:895–97
48. Jones, J. E. 1974. *Proc. R. Soc. London Ser. A* 340:263–86
49. Jantzen, R. T. 1983. *Adv. Ser. Astrophys.*, Vol. 1
50. Kerner, R. 1982. *Gen. Relativ. Grav.* 14:453–69
51. King, A. R., Ellis, G. F. R. 1973. *Commun. Math. Phys.* 31:209–42
52. Lindblom, L., Hiscox, W. A. 1982. *J. Phys. A* 15:1827–30
53. Linde, A. D. 1983. In *The Very Early Universe*, ed. G. W. Gibbons, S. W. Hawking, S. T. C. Siklos, pp. 205–49. Cambridge Univ. Press
54. Longair, M. S., Rees, M. J. 1973. In *Cargese Lectures in Physics*, ed. E. Schatzmann, 6:340–45. New York: Gordon & Breach
55. MacCallum, M. A. H. 1979. In *General Relativity: An Einstein Centenary Survey*, ed. S. W. Hawking, W. Israel, pp. 533–80. Cambridge Univ. Press
56. MacCallum, M. A. H. 1979. In *Physics of the Expanding Universe, Lecture Notes in Physics No. 109*, ed. M. Demianski, pp. 1–59. Berlin/New York: Springer

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

57. MacCallum, M. A. H. 1981. In *The Origin and Evolution of Galaxies*, ed. B. J. T. Jones, J. E. Jones, pp. 9–39. Dordrecht: Reidel
58. Markov, M. A. 1983. In *The Very Early Universe*, ed. G. W. Gibbons, S. W. Hawking, S. T. C. Siklos, pp. 353–71. Cambridge Univ. Press
59. Marochnik, L. S., Pelikov, N. V., Vereshkov, G. M. 1975. *Astrophys. Space Sci.* 34:249–95
60. Matravers, D. R., Vogel, D. L. 1983. Preprint, Dept. Appl. Math., Univ. Cape Town
61. McCrea, W. H. 1951. *Proc. R. Soc. London Ser. A* 206:562–75
62. Miller, B. D. 1979. *J. Math. Phys.* 20:1356–61
63. Misner, C. W. 1968. *Ap. J.* 151:431–57 (AAPT)
64. Misner, C. W. 1969. *Phys. Rev. Lett.* 22:1071–74 (AAPT)
65. Misner, C. W., Thorne, K. S., Wheeler, J. A. 1973. *Gravitation*. San Francisco: Freeman
66. Murphy, G. L. 1973. *Phys. Rev. D* 8: 4231–33
67. Nariai, H., Tomita, K. 1971. *Prog. Theor. Phys.* 46:776–86
68. Narlikar, J. V., Burbidge, G. R. 1981. *Astrophys. Space Sci.* 74:111–29
69. Ne'eman, Y. 1965. *Ap. J.* 141:1303–5
70. Nicholl, J. F., Segal, I. E. 1982. *Proc. Natl. Acad. Sci. USA* 79:3913–17
71. Novello, M., Heintzmann, H. 1983. Preprint CBPF-NF-010/83, Cent. Bras. Pesqui. Fis., Rio de Janeiro
71a. Novikov, I. D. 1964. *Sov. Astron. AJ* 8:857, 1075
72. Novikov, I. D., Zel'dovich, Ya. B. 1973. *Ann. Rev. Astron. Astrophys.* 11:387–412
73. Parker, L. 1982. *Fundam. Cosmic Phys.* 7:201–39
74. Parker, L., Fulling, S. A. 1973. *Phys. Rev. D* 7:2357–74 (AAPT)
75. Penrose, R. 1964. In *Relativity, Groups, and Topology*, ed. C. de Witt, B. de Witt, pp. 565–84. New York: Gordon & Breach
76. Penrose, R. 1979. *General Relativity: An Einstein Centenary Survey*, ed. S. W. Hawking, W. Israel, pp. 600–11. Cambridge Univ. Press
77. Penrose, R. 1980. In *Essays in General Relativity*, ed. F. Tipler, pp. 1–12. New York: Academic
78. Penrose, R. 1981. *Quantum Gravity 2*, ed. C. J. Isham, R. Penrose, D. W. Sciama, pp. 244–72. Oxford Univ. Press
79. Petrosian, V. 1982. *Nature* 298:506–7, 805–8
80. Petry, W. 1981. *Gen. Relativ. Grav.* 13: 1057–71
81. Phillipps, S., Ellis, R. S. 1983. *MNRAS* 204:493–506
82. Raychaudhuri, A. K. 1955. *Phys. Rev.* 98:1123–26
83. Raychaudhuri, A. K. 1980. *Relativistic Cosmology*. Cambridge Univ. Press
84. Rees, M. J. 1978. *Nature* 275:35–37, 343
85. Rindler, W. 1956. *MNRAS* 116:662–77
86. Robertson, H. P. 1933. *Rev. Mod. Phys.* 5:62–90
87. Rosen, N. 1980. *Gen Relativ. Grav.* 12: 493–510
88. Rosen, N. 1983. *Found. Phys.* 13:363–72
89. Rothman, T., Matzner, R. 1982. *Ap. J.* 257:450–55
90. Schmidt, M. 1969. *Ann. Rev. Astron. Astrophys.* 7:540–44
91. Schrödinger, E. 1956. *Expanding Universes*. Cambridge Univ. Press
92. Segal, I. E. 1976. *Mathematical Cosmology and Extragalactic Astronomy*. New York: Academic
93. Segal, I. E. 1982. *Ap. J.* 252:37–38
94. Shepley, L. C. 1969. *Phys. Lett. A* 28: 695–96
95. Siklos, S. T. C. 1977. *Commun. Math. Phys.* 58:255–72
96. Siklos, S. T. C. 1981. *Gen. Relativ. Grav.* 13:433–41
97. Silk, J. 1977. *Astron. Astrophys.* 59:53–58
98. Soneira, R. M. 1979. *Ap. J. Lett.* 230: L63–65
99. Stewart, J., Hajicek, P. 1973. *Nature Phys. Sci.* 244:96
100. Thorne, K. S. 1967. *Ap. J.* 148:51–68
101. Tipler, F. J. 1978. *Phys. Rev. D* 17: 2521–28
102. Tipler, F. J., Clarke, C. J. S., Ellis, G. F. R. 1980. In *General Relativity and Gravitation*, ed. A. Held, 2:97–206. New York: Plenum
103. Tolman, R. C. 1934. *Relativity, Thermodynamics and Cosmology*. Oxford Univ. Press
104. Tomita, K. 1978. *Prog. Theor. Phys.* 59:1150–69
105. Trautman, A. 1973. *Nature Phys. Sci.* 242:7–8
106. van den Bergh, N. 1983. *Gen. Relativ. Grav.* 15:441–48
107. Wagoner, R. V. 1980. In *Physical Cosmology*, ed. R. Balian, J. Audouze, D. N. Schramm, pp. 394–442. Amsterdam: North-Holland
108. Wald, R. M. 1983. Preprint MPI-PAE/PTh 38/83, Max Planck Inst. Phys., Munich

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System

1984ARA&A..22..157E

109. Weinberg, S. E. 1972. *Gravitation and Cosmology*. New York: Wiley
110. Wesson, P. S. 1975. *Astrophys. Space Sci.* 32:273–284, 305–30
111. Weyl, H. 1923. *Phys. Z.* 24:230–32
112. Will, C. M. 1979. In *General Relativity: An Einstein Centenary Survey*, ed. S. W. Hawking, W. Israel, pp. 24–89. Cambridge Univ. Press
113. Woszczynn, A., Betkowski, W. 1982. *Astrophys. Space Sci.* 82:489–93
114. Yahil, A., Sandage, A., Tammann, G. 1980. In *Physical Cosmology*, ed. R. Balian, J. Audouze, D. N. Schramm, pp. 127–59. Amsterdam: North-Holland
115. Zeldovich, Ya. B. 1970. *JETP Lett.* 12:307
116. Zeldovich, Ya. B. 1979. In *Physics of the Expanding Universe, Lecture Notes in Physics No. 109*, ed. M. Demianski, pp. 60–80. Berlin/New York: Springer
117. Zuiderwijk, E. J., de Ruiter, H. R. 1983. *MNRAS* 204:675–89

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System