The cosmological constant and cosmological change

The constant originally proposed by Einstein, and recent ideas on evolutionary processes affecting galaxies, are factors in the latest attempts to discover what type of universe we live in.

Beatrice M. Tinsley

Is the Universe of infinite extent, or is it a finite system? Will it expand forever, or will it reach some maximum size before turning and collapsing upon itself like an inverse Big Bang? Just a few years ago, models of the conventional Friedman types were showing consistent, albeit tentative, evidence for an open, ever-expanding Universe.[1,2,3] Since then, further data and theory have inevitably conspired to blur the appealing simplicity of that picture. In this article I will show how the Friedman models fare in the light of new developments—particularly the recognition of a whole new class of evolutionary corrections to the properties of distant galaxies (see figure 1), and a proposed reinstatement of Einstein's disinherited cosmological constant. We shall see that the basic questions, posed above, are still unanswered.

Model testing

A conservative approach to testing cosmological models is to restrict one's attention to a class of models with sufficiently few free parameters that they can be determined by "local" tests alone. Then the data on distant galaxies, in which cosmological effects such as deceleration and curvature are obscured by effects of galactic evolution, can be interpreted through a "known" model to give empirical information on the past properties of the galaxies themselves. This is the approach that I shall finally adopt here, in view of the complexity of problems besetting the so-called "global" tests.

Even with a strong preference for simplicity, we must not forget that the basic assumptions of the allowed class of models may be too restrictive or even incorrect. Cosmology has seen lately a swing of interest towards models with theories of gravity or redshift (or both) that differ from conventional General Relativistic interpretations. These ideas are beyond the scope of my article, but many of the considerations and constraints that I shall discuss are relevant in other theoretical contexts. For example, any cosmological model that is not strictly steady-state must confront the systematic effects of evolution on the global tests, and every model must be consistent with local data on density and on the ages of stars and galaxies.

Some theories that postulate variation of physical constants with time are restricted by observational evidence. For example, relative redshifts of lines in the absorption spectra of distant quasars have been used to set limits on the rate of change of certain dimensionless ratios—the mass ratio of the proton and electron, the fine-structure constant, and the nuclear g-factor of the proton. The most recent analysis[4] leads to the conclusion that none of these ratios can have varied much (that is, by a fraction of order unity) during a lookback time that is a large fraction of the age of the Universe. Expressed in a form independent of particular cosmological models, this statement can be rephrased: these quantities do not change significantly during a large fraction of the "Hubble time." This time-scale is the inverse of Hubble's constant, H_0, and is the time that would have elapsed since the Big Bang if the expansion rate were constant. With H_0 in the usual units of kilometers per second per megaparsec, the Hubble time in years is approximately $10^{10}\,(100/H_0)$.

Tentative empirical evidence[5] for a possible variation in the gravitational constant G, with a time-scale of the order of $H_0{}^{-1}$ has yet to be confirmed. If true, variation of G would invalidate the models based on Einstein's General Theory of Relativity—as well as much of the astrophysics used to interpret the data. For now, the most fruitful philosophy appears to be to use the simplest general-relativistic models—the Friedman models.[6]

Friedman models . . .

Accordingly, we begin by adopting General Relativity with no cosmological constant, and by making the usual assumptions that the Universe on large scales is homogeneous and isotropic, and that its dynamically important contents now are noninteracting "particles," or in other words pressure-free "dust."[7] The time dependence of the scale factor of the Universe, $R(t)$, is then given by the set of Friedman models with zero cosmological constant; figure 2 summarizes their properties most relevant for empirical tests. This well known set includes the "open" Universe, expanding forever, and the "closed" Universe, which eventually collapses.

Two independent parameters define a model (in this case where we assume no cosmological constant). It is convenient to use H_0 as a scale factor for times and distances, and to use a dimensionless density parameter Ω_0 to define the shape of the function $R(t)$. The empirical definition of H_0 is the ratio of velocity to distance for galaxies with redshift $z \ll 1$; an equivalent definition in theoretical terms is

$$H_0 \equiv \dot{R}_0/R_0 \qquad (1)$$

where the zero subscripts denote present values. Units of km $\sec^{-1}$ Mpc^{-1} will be implied for H_0, except that whenever the

---

The author is an associate professor of astronomy at Yale University.

dimensionless product H_0t_0 is written, inverse units are implied. It is convenient to note that

$$(H_0/100 \text{ km sec}^{-1} \text{ Mpc}^{-1}) \times (t_0/10^{10} \text{ yr}) = 0.98 \times (\text{dimensionless } H_0t_0) \quad (2)$$

The definition of the density parameter Ω_0, with numerical values, is

$$\Omega_0 \equiv \rho_0/\rho_c = 8\pi G\rho_0/3H_0^2 = 5.33 \times 10^{28} (\rho_0/\text{gm cm}^{-3})(H_0/100)^{-2} \quad (3)$$

where the so-called "critical" density is

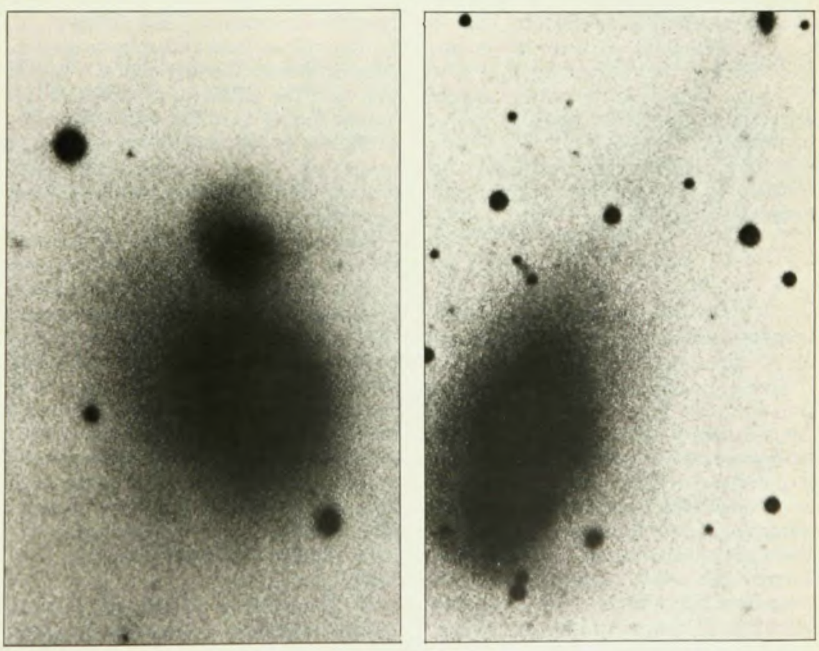
$$\rho_c \equiv 3H_0^2/8\pi G = 1.88 \times 10^{-29} (H_0/100)^2 \quad (4)$$

The models allow two possible ultimate futures for the Universe, bringing us to the key question: Will the Universe expand forever, or will it eventually stop and collapse? This question can be asked in several equivalent ways (tabulated in figure 2).

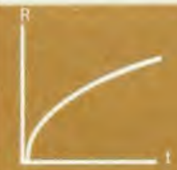
▶ Is the density parameter $\Omega_0 \leq 1$, or is $\Omega_0 > 1$? When the density is estimated from counts or kinematics of galaxies, the distance scale enters in such a way that one obtains a ratio ρ_0/H_0^2, that is, an estimate of Ω_0. The contribution of galaxies and their associated matter to Ω_0 (for example, extended halos or intracluster matter) is denoted Ω^*. A number of lines of evidence reviewed earlier[1,2] and more recent work on galaxy counts by Richard Gott and Edwin Turner[8] lead to the estimate $\Omega^* = 0.06 \pm 0.02$. Any significant mass density *not* included in Ω^* must be invisible and evenly distributed, requiring a very contrived set of hypotheses as to its nature. The value $\Omega^* = 0.04$ is shown as a lower limit to Ω_0 in figure 3, which contains various constraints on the models in the Ω_0–H_0 plane. The other constraints follow from alternative ways of asking the question posed above.

▶ Is the mean density $\rho_0 \leq \rho_c$, or is $\rho_0 > \rho_c$? If deuterium was made in the Big Bang, the only viable site at present, then an interstellar abundance X_D of around 2×10^{-5} and the standard, most straightforward assumptions about Big Bang conditions together require $\rho_0 < 4 \times 10^{-31}$ gm cm^{-3} (independently of H_0).[9] Deuterium is destroyed during galactic evolution; so the density must be smaller than this limit by 20% or more. It is therefore shown as an upper bound in figure 3.

▶ Is the dimensionless age parameter $H_0t_0 \geq 2/3$, or is $H_0t_0 < 2/3$? Experts on the value of H_0 disagree strongly as to whether a reasonable lower bound is 30 or 70. In the latter case, the condition for an



Merging galaxies. The apparent luminosities and sizes of distant galaxies have been regarded as the most promising probes for testing cosmological models, on the assumption that the intrinsic sizes and luminosities are either constant or evolving in a steady and predictable manner. Recently, however, it has been noted that large galaxies tend to swallow their companions, causing changes that greatly complicate the interpretation of cosmological tests. These two photographs, from H. C. Arp's *Atlas of Peculiar Galaxies* (California Institute of Technology, 1966), illustrate possible instances of (on the left) two galaxies beginning to merge, and (on the right) a composite galaxy evidencing a recent merger. Figure 1

FRIEDMAN MODELS WITH $\Lambda = 0$			
	Open	Critical	Closed
R(t)			
Future	Expand forever	Expand forever	Collapse
Curvature	$k<0$ Hyperbolic	$k=0$ Flat	$k>0$ Spherical
Deceleration	$0<q_0<½$	$q_0=½$	$q_0>½$
Density parameter $\Omega_0 = \rho_0/\rho_c$	$0<\Omega_0<1$	$\Omega_0=1$	$\Omega_0>1$
Age t_0	$1>H_0t_0>⅔$	$H_0t_0=⅔$	$H_0t_0<⅔$

Friedman models of the Universe. These are the three familiar cases that arise when no "cosmological-constant" term is included—the open Universe that expands forever, the closed Universe that collapses upon itself in a finite time, and the critical case between these two. $R(t)$ is a scale factor of the Universe (a function of time t), and k, the curvature parameter, can take values -1, 0 and $+1$ as illustrated. Also q_0 is the deceleration parameter, Ω_0 a dimensionless density parameter, and H_0t_0 is a dimensionless age parameter. Figure 2

ever-expanding Universe is met if the age $t_0 > 9.3 \times 10^9$ years, as is indeed suggested by analyses of globular clusters. But if $H_0 = 30$, the condition is not met unless the Universe is very old indeed, $t_0 > 22 \times 10^9$ years. Plausible limits from stellar ages, putting t_0 at 10–20 times 10^9 years, are indicated in figure 3.

▶ Is the deceleration parameter $q_0 \leq ½$, or is $q_0 > ½$? The definition of q_0 is (in all cosmological models)

$$q_0 \equiv -\frac{\ddot{R}_0 R_0}{\dot{R}_0^2} \qquad (5)$$

and in this class of models we have the simple relation

$$q_0 = \Omega_0/2 \qquad (6)$$

The deceleration can, in principle, be measured by determination of departures from a linear "velocity–distance" relation for galaxies, with their apparent magnitudes, angular sizes, and so on used as "distance" parameters. These global tests are afflicted by large and uncertain corrections for selection effects and the evolution of intrinsic properties of galaxies,[2,3] so I shall not consider them at this stage where we are restricted to models for which local data provide equivalent and more tractable tests.

▶ Is the curvature of space negative or zero, or is it positive? The curvature parameter k, which can take the values -1, 0, or $+1$, is given in these models by

$$kc^2/R_0^2H_0^2 = \Omega_0 - 1 \qquad (7)$$

Now the curvature can also, in principle, be determined from global tests, specifically by means of galaxy counts. But in this case, the effects of evolution are so overwhelming that the test is best regarded as a probe of the distant past properties of galaxies, requiring small "corrections" for uncertainties in the cosmological model.

Figure 3 summarizes the model constraints provided by local tests alone. Notice that the value of Ω^*, as a lower bound and probably a close approximation to Ω_0, is itself highly suggestive that the Universe is open and ever-expanding. The limits on t_0 then act chiefly as constraints on H_0, which are more restrictive than the limiting values currently proposed for the distance scale itself. If one accepts the additional physical assumptions of "standard" Big-Bang nucleosynthesis, then the abundance of deuterium clinches the case for an open Universe.

At the expense of making drastic, though conventional, assumptions about the allowed class of cosmological models, we have apparently been able to derive startlingly fundamental conclusions: the Universe is infinite in spatial extent, and will expand for an infinite future.

There are several motives for looking beyond this straightforward picture. To some people, the prospect of a monotonically expanding Universe is philosophically bleak; to others, the thought that a definitive model for the Universe itself could be reached is incredible (if not absurd) and to many, the empirical basis of the foregoing conclusions is altogether insecure. None of the "local" pieces of evidence for an open Universe is by itself fully convincing; for example, a smooth medium of stellar-mass black holes could fill the Universe ($\Omega_0 \geq 1$) and be undetectable either in dynamical estimates of Ω^* or by optical effects. Only the mutual consistency of the foregoing tests, leading without contrivance to a small but finite allowed area in the Ω_0–H_0 plane (figure 3) lends the results credence.

. . . with a cosmological constant

Yet another motive for complicating this scene is the possibility that the models do not use an adequate theory. I shall consider here only the least radical alternative: General Relativity with the cosmological constant, Λ.

The eventful early history of Λ as a parameter in General Relativity has been reviewed by J. D. North,[10] Vahé Petrosian,[11] and by James Gunn and myself.[12] Let us look at a few episodes in this history.

Einstein first introduced the Λ term (as a positive constant in his theory) to keep the Universe static in the face of gravity. When the recession of the galaxies was discovered, he admitted that it was "gravely detrimental to the formal beauty of the theory" [13] and dropped it with relief. Meanwhile, other cosmologists had become very attached to this constant. Georges Lemaître emphasized that the age of the solar system was very much greater than the Hubble time as estimated in his day (H_0^{-1} was 2×10^9 yr by early estimates), so he advocated the class of models known by his name, in which a near balance between gravitational attraction and the repulsive effect of Λ keeps the Universe at an almost constant scale (R) for an arbitrarily long time, depending on how close Λ is to the critical value introduced by Einstein. This rationale for keeping Λ disappeared with the downward revision of H_0. Arthur Eddington preferred models, with the Einstein critical value of Λ, that expand after an indefinite past time near the "Einstein radius." He deplored models with a Big Bang (that is, those that start a finite time ago with zero scale) for a reason that is quaint in the light of today's arguments *for* a Big Bang in order to thermalize the background radiation: Eddington wrote in 1932, "The theory recently suggested by Einstein and de Sitter . . . leaves me cold." [14] He regarded Λ as a fundamental constant of nature, and held that "if ever the theory of relativity falls into disrepute, the cosmical constant will be the last stronghold to collapse." [14]

Most cosmologists in later decades have

echoed Einstein's views about the formal and aesthetic advantages of assuming that $\Lambda \equiv 0$ unless empirical evidence demands the term.

In 1967 Petrosian, Edwin Salpeter and P. Szekeres[15] suggested that Lemaître models could account for an apparent excess of quasars with z approximately 2.0, but the statistical evidence for the excess soon vanished. A negative (attractive) Λ has been postulated to account for the high velocity dispersion of galaxies in clusters, but this suggestion would require a very large positive value of q_0 (as we shall see from the equations given below), which is almost out of the question.

A positive value of Λ was again suggested in 1975,[12] as the most plausible explanation of an apparent *acceleration* of the Universal expansion ($q_0 < 0$). Our immediate empirical evidence for a Λ term has, meanwhile, again been swallowed up, this time by a new evolutionary correction to magnitudes of galaxies,[16] which I shall discuss later. But there is no compelling reason to follow Einstein and reject his own creation, Λ, merely because it is not *demanded*. Let us therefore consider the properties of Friedman models with Λ.

Model properties

Three parameters are now needed to specify a model: figure 4 suppresses the scale parameter H_0 and indicates the nature of the functions $R(t)$ in the q_0–Ω_0 plane. Some useful relations define the curves that divide the q_0–Ω_0 plane by model type.

The curvature is given by a generalization of equation 7

$$\frac{kc^2}{R_0{}^2H_0{}^2} = \Omega_0 - 1 + \frac{\Lambda}{3H_0{}^2} \qquad (8)$$

which shows that the Universe can be closed ($k = +1$) because of a positive Λ, even if $\Omega_0 \ll 1$. Figure 4 shows how the one-to-one relation between sign of curvature and future behavior also breaks down if $\Lambda \neq 0$.

A critical value of Λ (an example of which was mentioned above as the value for the Einstein static Universe) is given by the relation

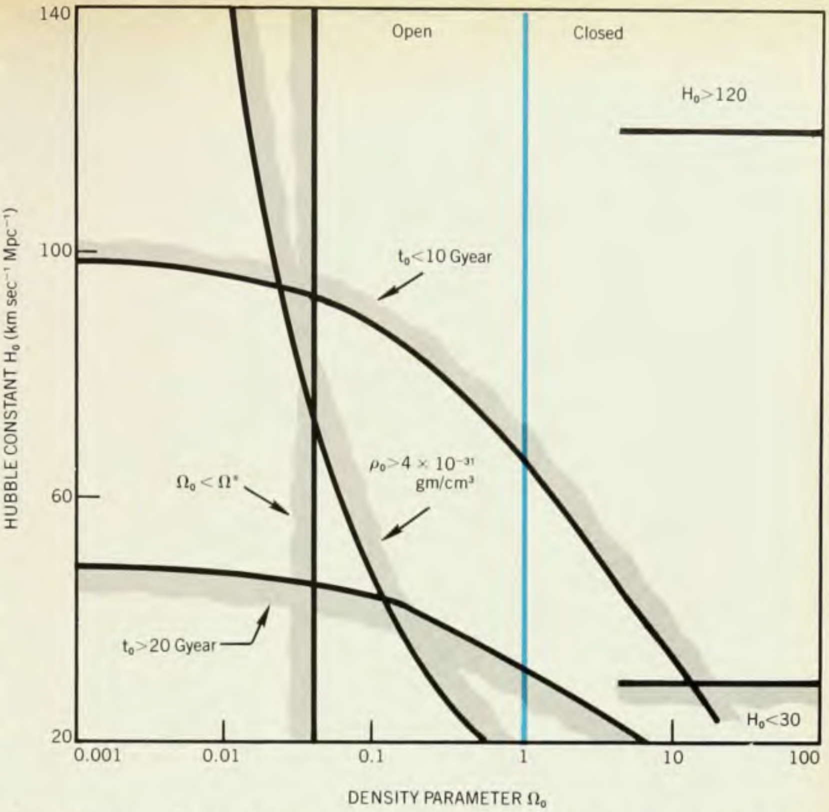
$$\Lambda_c = \left(\frac{c^3}{4\pi G \rho_0 R_0{}^3}\right)^2 \qquad (9)$$

This is equivalent to the following equation for the value of Ω_0 on the curve $\Lambda = \Lambda_c$, at a given value of q_0, showing that Λ_c exists only for $q_0 < -1$ and $q_0 > \frac{1}{3}$

$$\Omega_c = \frac{1}{3}(q_0+1)\left\{(q_0+1) \pm \left[(q_0+1)\left(q_0-\frac{1}{3}\right)\right]^{1/2}\right\} \qquad (10)$$

Models with $\Lambda < \Lambda_c$ and $q_0 < -1$ have no Big Bang (that is, there is no root to the equation $R = 0$).

Apart from the small area with $0 < \Lambda < \Lambda_c$ in the upper right of the diagram, all models with $\Lambda > 0$ expand forever at accelerating rates. This is because $\Lambda > 0$ acts like a repulsive force that increases with distance.



Constraints on the Friedman models of figure 2, illustrated on the H_0–Ω_0 plane. Models at and to the left of $\Omega_0 = 1$ will expand forever; those at the right will eventually collapse. Three pairs of constraints are shown; two for H_0, two for t_0 and two for density Ω_0. These constraints are provided by local tests alone. Note that the only allowed region of (H_0, Ω_0) is the small white triangle near the center of the figure; in other words, the rather conventional assumptions made here lead to the conclusion that the Universe is open. Figure 3

The generalization of equation 6,

$$q_0 = \frac{\Omega_0}{2} - \frac{\Lambda}{3H_0{}^2} \qquad (11)$$

shows that a negative Λ enhances effects of gravity in retarding the expansion, and that a positive Λ acts in the opposite sense. It is clear that a negative value of q_0 (that is, universal acceleration) implies $\Lambda > 0$ (although the converse is not necessarily true). Equation 11 also suggests that we associate Λ with a density defined by

$$|\rho_\Lambda| \sim \left|\frac{\Lambda}{4\pi G}\right| = \rho_0\left|1 - \frac{2q_0}{\Omega_0}\right| \qquad (12)$$

Because ρ_Λ is constant, while the density of matter decreases with time (as R^{-3}), this relation provides another way of understanding why the influence of Λ increases with time. Moreover, Λ will influence the dynamics of any region of the universe with a mean density not greater than ρ_Λ. Density enhancements, such as clusters of galaxies, would thus be affected if $|q_0|$ exceeded Ω_0 by an amount of order (mean cluster density)/ρ_0; this does not appear to be a significant possibility now.

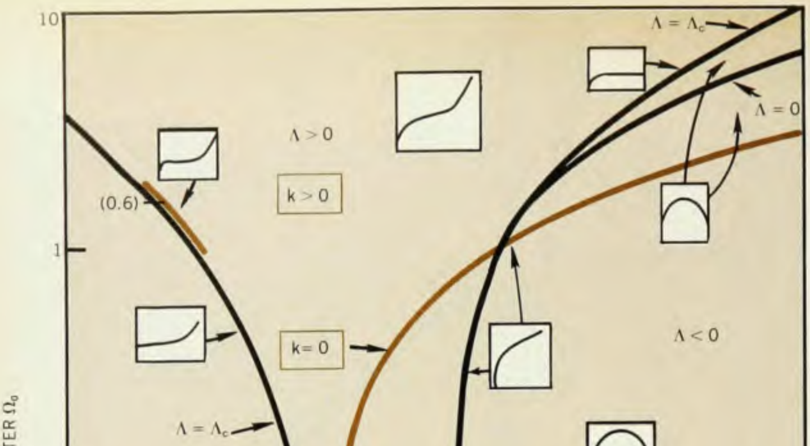
The Lemaître models can be found in figure 4 just above the left branch of the $\Lambda = \Lambda_c$ curve, and models of Eddington's type are on that curve. The minimum value of $R(t)$ in the latter models corresponds to a maximum redshift, z_m, given by

$$(1 + z_m)^3 = 1 - (2q_0/\Omega_0) \qquad (13)$$

Several values of z_m are marked in figure 4. Contracting–expanding models below the critical line have $z_m < 100$ as long as $\Omega_0 > 0.02$ and $q_0 > -3$. These minimum redshift values are independent constraints on such models: the quasar at $z = 3.5$ rules out Eddington models if $\Omega_0 > 0.025$, and, more drastically, the need for an epoch with $z > 100$ to thermalize the background radiation rules out all Eddington models (because they would need $\Omega_0 < 2 \times 10^{-6}$), as well as any models below the critical line that do not have very negative values of q_0.

Constraints on Λ

When we use this kind of reasoning to eliminate models with $\Lambda < \Lambda_c$ (if $q_0 < -1$) and set an upper limit to Ω_0, we already have an interesting upper limit to the value of Λ. For example, if $\Omega_0 < 0.1$, then



Types of Friedman model with a cosmological constant Λ, plotted on the Ω_0–q_0 plane. The small insets show the behavior of $R(t)$ in various regions. The plane is divided into regions of positive and negative curvature k and other regions of positive and negative Λ. The numbers in parentheses on the left branch of the Λ_0 line denote values of maximum redshift z_m. Figure 4

the condition $\Lambda > \Lambda_c$ requires $q_0 > -1.3$ (equation 10), so that (by equation 11)

$$\Lambda/3H_0^2 < 1.35 \quad (14)$$

Thus the distance-scale on which Λ may be important is given by

$$c\Lambda^{-1/2} > 1500\,(100/H_0)\ \mathrm{Mpc} \quad (15)$$

Of course, this distance corresponds to redshifts of order unity because Λ/H_0^2 is of order unity. (This method does not rule out a negative Λ with effects on smaller scales.) The constraint $\Lambda > \Lambda_c$ (if $q_0 < -1$) is indicated on a q_0–Ω_0 plane in figure 5.

The extreme Lemaître models, with Λ slightly greater than Λ_c and $q_0 < -1$, can also be eliminated fairly firmly, because their optical properties are strikingly unlike those of other Friedman models.[11] The absence (these days) of a pronounced peak in the quasar rédshift distribution, and the absence of very bright objects expected if the "antipole" (where the radial coordinate equals π) occurs at redshifts less than 3, rule out such models in the light-colored area of figure 5. The numerical limit to Λ is not very different from that in equation 14, but it is of some interest to be able to eliminate models displaying pronounced effects of a Lemaître "coasting" phase.

The density limits used for models with no cosmological constant still apply here because, as noted above, other constraints tell against a value of $|\Lambda|$ great enough to affect the main dynamical arguments relating to density. Figure 5 shows the limit $\Omega_0 > \Omega^* > 0.04$, and the upper limit to ρ_0 from deuterium production, for three values of H_0. If $H_0 > 75$, "standard" deuterium production is inconsistent with the adopted lower limit to Ω_0. With H_0 between 30 and 75 the foregoing constraints together allow both those models that collapse in the future ($k < 0$, $\Lambda < 0$) and those that expand forever (any sign of curvature, Λ greater than, or equal to, zero).

Arguments based on the age of the Universe are potentially powerful in this context. Figure 5 contains the loci $H_0t_0 = 0.5, 0.6, 1.0$ for illustration. In the present state of confusion over the values of H_0 and t_0, the dimensionless product may lie anywhere between say 0.3 (if $H_0 = 30$, $t_0 = 10 \times 10^9$ yr) and 2.4 (with $H_0 = 120$, $t_0 = 20 \times 10^9$ yr), thus spanning all model types in the q_0–Ω_0 plane that were not ruled out by the set of constraints near Λ_c at the left. It is nevertheless intriguing to consider "best guesses" of the parameter values. If we take $t_0 = (16 \pm 2) \times 10^9$ yr (to be consistent with recent work on globular clusters and elliptical galaxies), and $H_0 = 70 \pm 20$ (to overlap most recent determinations), we have $H_0t_0 = 1.1 \pm 0.4$. Now values of H_0t_0 greater than unity imply that the present expansion rate exceeds its average past value; such acceleration is impossible if $\Lambda \leq 0$, so we see that the current estimates of H_0t_0 come close to demanding that $\Lambda > 0$. The old argument for Λ used by Lemaître is reappearing, although in a much less extreme form! Obviously this argument is very insecure, and the estimates of H_0t_0 are by no means in conflict with the value of approximately 0.9 predicted by a model with $\Lambda = 0$ and $\Omega_0 \ll 1$; but the test should be kept in mind when estimates of H_0 and t_0 are refined.

The Hubble diagram and q_0

As a final constaint on Λ, let us consider the deceleration parameter q_0. This parameter can, in principle at least, be derived from global tests.[2] The test we will consider here is the Hubble diagram for giant elliptical galaxies—that is, the relation between their redshift and apparent magnitude.

The theoretical magnitude–redshift relation at $z \ll 1$ can be written

$$m = C + 5\log z - 1.09z(A + E + 0.65q_0) + 0(z^2) \quad (16)$$

Here the constant C includes H_0 (the distance scale), the constant A depends on the type of apparent magnitude used (such as monochromatic, or broad-band with K correction), and both A and the coefficient 0.65 include the aperture correction in analytical form, introduced by Gunn and J. Beverley Oke.[17] The parameter E allows for evolution of the luminosity L of that part of the galaxies that lies within the observer's aperture; in the first-order approximation, we have

$$E = -\frac{1}{H_0t_0}\left(\frac{d\log L}{d\log t}\right)_0 \quad (17)$$

For orientation, schematic first-order m–z relations are shown in figure 6. The black lines are for models with $q_0 = -1, 0,$ and $+1$, ignoring evolution; in each case the magnitudes are relative to those predicted for $q_0 = 0, E = 0$. At redshifts around 0.4, models with q_0 differing by 1 have magnitudes differing by about the intrinsic dispersion in absolute magnitudes of first-ranked cluster galaxies.

Equation 16 demonstrates the well known point that first-order deviations from linearity in the Hubble diagram measure the deceleration, with a correction for evolution; in this order, Λ enters only via H_0t_0 in the evolutionary correction. Ignoring evolution, one would infer an apparent value of q_0

$$q_{0a} = q_0 + E/0.65 \qquad (18)$$

so the "correction" to q_0, in first order, is

$$q_0 = q_{0a} - q_0 = 1.54\,E \qquad (19)$$

Thus if H_0t_0 is of order unity, evolution at a rate of a few per cent per 10^9 yr alters the apparent value of q_0 by about unity.

Stellar evolution in elliptical galaxies is currently estimated to make their luminosity grow fainter at a rate $(\partial \log L/\partial \log t)_* \approx -1$ ($\pm$ a few tenths, with more negative values more likely).[3] The correction for this effect is therefore given by

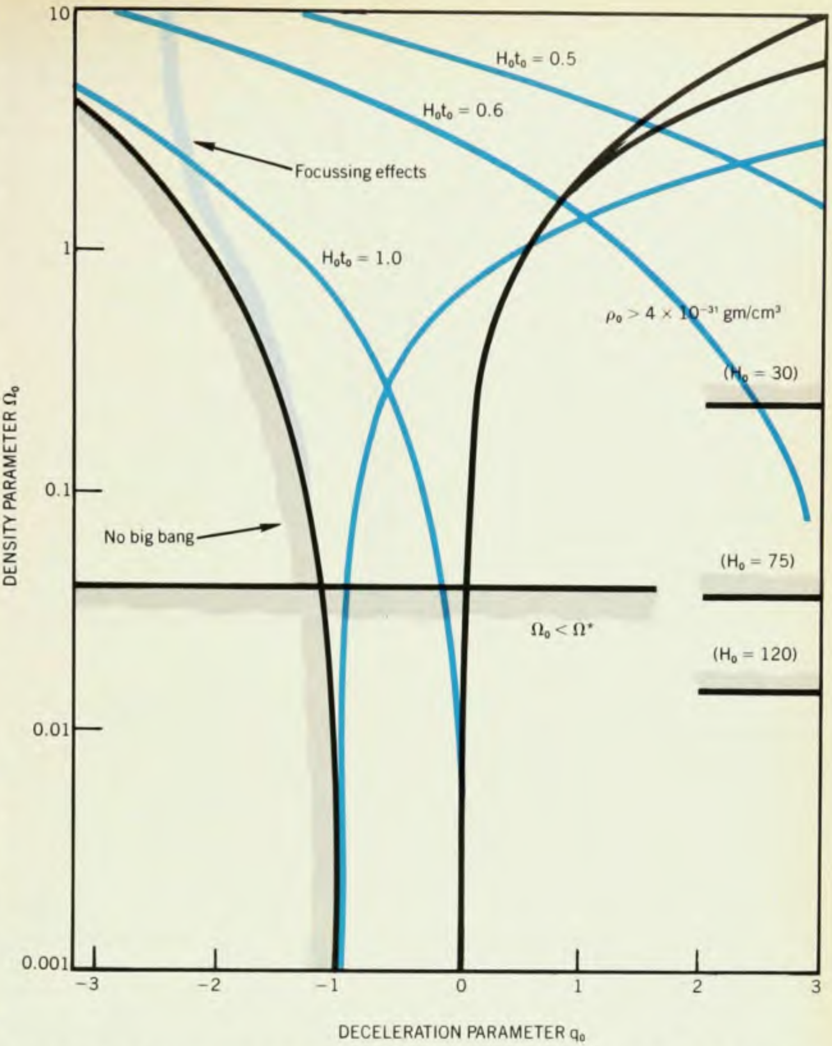
$$E_* \geq (H_0t_0)^{-1},$$
$$\Delta q_{0*} \gtrsim 1.5\,(H_0t_0)^{-1} \qquad (20)$$

The importance of stellar evolution to the m–z region is illustrated in figure 6, where the dotted lines include equation 20 in models with $q_0 = 0$ and -1 (each with a value of H_0t_0 appropriate to $\Omega_0 = 0.06$). Comparison with the black lines confirms that the shape of the m–z relation corresponds to an apparent value of q_0 that is too big by at least unity.

Herein lay the basis of Gunn's and my[12] reasoning for a positive Λ. Gunn and Oke[17] had recently obtained $q_{0a} = -0.15 \pm 0.57$ (1σ), which gives a true value at least 2σ below zero after correction for stellar evolution. As explained above, if $q_0 < 0$, Λ must be positive; moreover, the Universe must expand forever, because the repulsive effect of Λ becomes more important as expansion reduces the mean density of matter.

Counter-evolution

The suggestion of a negative value of q_0 created some problems, not only in reconciling the various density and age constraints,[12] but also in calling once more for the revival of Λ. An alternative interpretation of the Hubble diagram was not long in forthcoming. Jeremiah Ostriker and Scott Tremaine[16] suggested that the central cluster galaxies used in that test can grow by accretion of smaller galaxies, plausibly at a rate fast enough to counterbalance the dimming due to stellar evolution (see figure 1). Thus the net evolutionary correction could be near zero, and an apparently zero value of q_{0a} could arise in a Universe with $\Lambda = 0$ (and $0 < q_0 = \Omega_0/2$). The postulated accretion is due to dynamical friction, whereby smaller cluster galaxies entering the out-



Constraints on Friedman models with finite Λ, in the Ω_0–q_0 plane. The $k = 0$, $\Lambda = 0$ and $\Lambda = \Lambda_c$ lines are repeated from figure 4, with constraints on Ω_0 and ρ_0 added; three limits to ρ_0 (from deuterium) are shown corresponding to the three values of H_0 indicated. The light-colored region labelled "focussing effects" can be ruled out, by considerations of quasar red-shift distribution. The constraints permit open and closed, and expanding and collapsing, models. Figure 5

skirts of the central giant are slowed in their orbits so that they spiral in, and are eventually swallowed entirely by the captor.

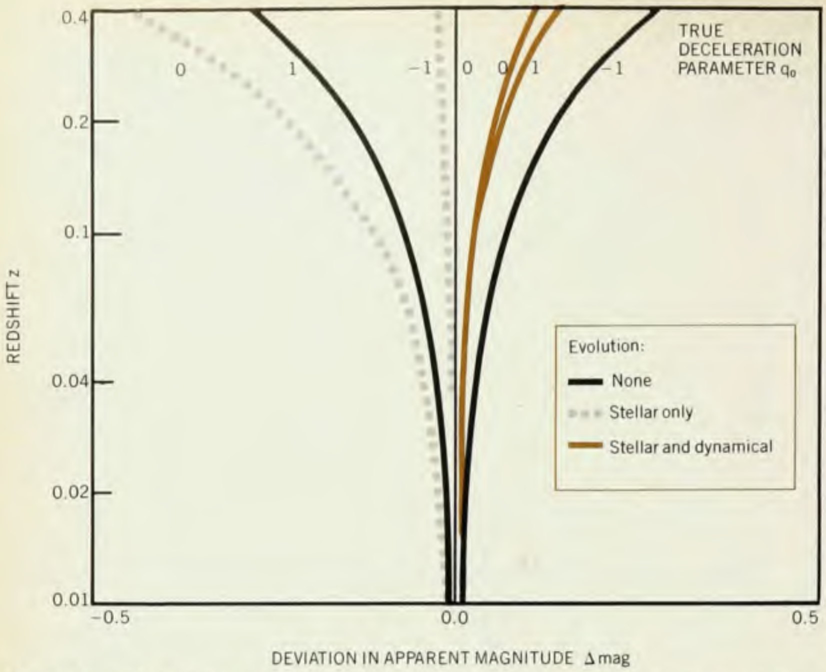
The implications for cosmology may be drastic.[16,18] It can be shown that, because of the factor H_0t_0 in the term E (equation 17), a component of evolution with $dL/dt > 0$ has a much less pleasant effect than to cancel the correction for stellar evolution: it makes the Hubble diagram very insensitive to q_0. Examples are given in figure 6. In this case, dynamical evolution is represented by a term

$$E_d = -\frac{0.5}{H_0t_0 - 0.33} \qquad (21)$$

which corresponds to the asymptotic case of dynamical friction with the assumption that accretion began at a cluster collapse time $0.33\,(H_0t_0)^{-1}$. The colored lines in figure 6 include also stellar evolution as before, and are for models with $\Lambda = 0$ and $q_0 = 0.03$ and 1. The first-order m–z relations for these models are now almost identical, with the greater q_0 even appearing slightly fainter because its smaller value of H_0t_0 enhances the importance of evolution.

Dynamical friction may get us out of the need for a cosmological constant, but at the expense of leaving q_0 indeterminate, at least by means of this hitherto most promising test.

We need detailed calculations of accretion rates, requiring much knowledge about the distributions and orbits of cluster galaxies, and about the density distributions of stars within the central and victim galaxies. It is not obvious that an effect as large as that indicated by equation 21 must arise. Tests have been suggested based on the fact that small elliptical galaxies are normally bluer than large ones; thus significant growth by accretion should make the central galaxy



Magnitude–redshift relations, shown schematically relative to the line for $q_0 = 0$ without evolution. Values of q_0 (rounded to an integer in each case) for no evolution, stellar evolution only, and stellar and dynamical evolution combined, are shown with the appropriate curves. Figure 6

measurably too blue for its magnitude. Even if the test is positive, the implications for the Hubble diagram are not clear. It is critical to know what fraction of the victim galaxy ends up inside the aperture radius, and to what extent stars are lost from the aperture by the swelling that accompanies accretion.[18]

The upshot is that we cannot, at present, determine q_0 from the Hubble diagram. Values of q_0 less than -2 or greater than 2 appear unlikely, but even with the constraints discussed above, the types of models allowed in the q_0–Ω_0 plane are still very diverse. Unless we can show that brightening by accretion is negligible, evolutionary effects in the Hubble diagram will remain very uncertain. It may be most fruitful to turn the interpretation of the diagram around, estimating q_0 by other tests (preferably local) then using the m–z diagram to study galactic evolution.

Present status, future plans

If we accept the Friedman models as a valid set for interpreting data on the large-scale structure of the Universe, their status may be summarized as follows:

▶ The Hubble diagram and other presently feasible global tests for q_0 are so extremely sensitive to the evolution of intrinsic galaxy properties that they will provide at best weak constraints on the model.

▶ If we assume that there is no cosmological constant, then density estimates strongly suggest that the Universe is open and will expand forever.

▶ If the cosmological constant is allowed, present constraints are consistent with a wide range of model types: the Universe could be spatially open or closed, and it could expand forever or eventually collapse. The presence of thermal background radiation, and the absence of extreme optical effects associated with the Lemaître models, eliminate models with $q_0 < -1$ and Λ less than, equal to, or fractionally greater than its critical value.

Could we ever convince ourselves that Λ does exist? The aesthetic arguments against Λ have lost some sway following Ya. B. Zeldovich's[19] suggestion that a Λ term may arise from quantum fluctuations of the vacuum; then Λ becomes part of the stress–energy tensor rather than part of General Relativity.

Empirical arguments for Λ must take the form of invalidating one of the predictions of the models with $\Lambda = 0$.

One test would be to show that $q_0 \neq \Omega_0/2$ (equation 11), but this approach is not promising because the evolutionary corrections must be precisely known.

Another possibility would be to show that the product $H_0 t_0$ does not have the value required by Ω_0 if $\Lambda = 0$. This test has the great advantage of requiring only local data. The astrophysical problems in finding the distance scale for H_0 and the ages of stars or elements for t_0 are formidable, but they are perhaps not quite as intractable as the problems involved in estimating differences at the 1% level between nearby elliptical galaxies and those with redshifts of several tenths!

Finally, let us return to the question of whether the Universe will expand forever, with reference to the model types illustrated in figure 4. It can be seen that if $\Omega_0 \lesssim 0.2$ (as seems almost certain from estimates of Ω^*, and the highly restricted possibilities for matter not included in Ω^*), then the Universe will expand forever if $\Lambda \leq 0$, that is, if $q_0 < +0.1$. The most tractable method of addressing this question seems to be via the time-scale: the condition $H_0 t_0 > 0.85$ is sufficient, though not necessary, to show that the Universe expands forever if $\Omega_0 < 0.2$. There is no firm answer yet, but local tests appear to be extremely promising.

* * *

This article is an adaptation of a paper originally presented at the Eighth Texas Symposium, in Boston, Mass., December 1976. The proceedings of that symposium are to be published in Annals of the New York Academy of Sciences.

The work reported in this article was supported in part by the National Science Foundation (Grant AST76-16329) and the Alfred P. Sloan Foundation.

References

1. J. R. Gott, J. E. Gunn, D. N. Schramm, B. M. Tinsley, Astrophys. J. **194,** 543 (1974), and *Scientific American,* March 1976, page 62.
2. J. E. Gunn, Ann. N.Y. Acad. Sci. **262,** 21 (1975).
3. B. M. Tinsley, Ann. N.Y. Acad. Sci. **262,** 436 (1975).
4. A. M. Wolfe, R. L. Brown, M. S. Roberts, Phys. Rev. Lett. **37,** 179 (1976).
5. T. C. Van Flandern, Ann. N.Y. Acad. Sci. **262,** 494 (1975).
6. A. Friedman, Zeits. f. Physik **10,** 377 (1922).
7. W. Rindler, PHYSICS TODAY, November 1967, page 23.
8. J. R. Gott, E. L. Turner, Astrophys. J. **209,** 1 (1976).
9. D. N. Schramm, R. V. Wagoner, PHYSICS TODAY, December 1974, page 40.
10. J. D. North, *Measure of the Universe,* Clarendon, Oxford (1965).
11. V. Petrosian, in *Confrontation of Cosmological Theories with Observational Data,* M. S. Longair, ed., Reidel, Dordrecht (1974).
12. J. E. Gunn, B. M. Tinsley, Nature **257,** 454 (1975).
13. A. Einstein, in *Principle of Relativity,* H. A. Lorentz *et al.,* Methuen, London (1922).
14. A. S. Eddington, *The Expanding Universe,* Cambridge U.P., Cambridge (1933).
15. V. Petrosian, E. E. Salpeter, P. Szekeres, Astrophys. J. **147,** 1222 (1967).
16. J. P. Ostriker, S. D. Tremaine, Astrophys. J. Lett. **202,** L113 (1975).
17. J. E. Gunn, J. B. Oke, Astrophys. J. **195,** 225 (1975).
18. J. E. Gunn, B. M. Tinsley, Astrophys. J. **210,** 1 (1976).
19. Ya. B. Zel'dovich, Sov. Phys. Uspekhi **95,** 209 (1968). □