NONSTANDARD COSMOLOGIES

by

Jayant V. Narlikar

Tata Institute of Fundamental Research
Bombay 400005, INDIA

I. INTRODUCTION

The adjective 'nonstandard' applied to cosmologies is intended to describe those ideas about the origin and structure of the universe that are different from the Friedmann-Robertson-Walker models based on Einstein's general theory of relativity.

To fix ideas, these latter models, called the 'standard models' are given by the Robertson-Walker line element

$$ds^2 = c^2\, dt^2 - S^2(t) \left[\frac{dr^2}{1-kr^2} + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2)\right] \qquad (I.1)$$

where (r,θ,ϕ) are the comoving coordinates of a typical 'fundamental observer' and t the cosmic time. The spaces t = constant are homogeneous and isotropic and are completely specified by the curvature parameter $k = 0$, $+1$ or -1. In terms of observable physics, the above statement is called the 'cosmological principle' implying that all fundamental observers at any given epoch find the universe in the same physical state and that to the typical observer viewing the universe there is no preferential direction.

The function $S(t)$ called the 'scale factor' or the 'expansion factor' is determined by Einstein's equations:

$$R_{ik} - \frac{1}{2} g_{ik} R = - \frac{8\pi G}{c^4} T_{ik} . \tag{I.2}$$

The 'physics' is contained in the T_{ik} on the right hand side. The standard Friedmann models describe the universe as originating in a spacetime singularity at $S = 0$ (the 'big bang'). The cosmic clock begins from this instant. In the early stages the universe was radiation-dominated while in later stages (including the present) it was dust-dominated. Whether it will continue to expand for ever ($k = 0, -1, \dot{S} > 0$) or it will eventually contract ($k = +1, \dot{S} > 0$ now, $\dot{S} < 0$ later) depends on how much gravitating matter it has per unit volume. The critical density (corresponding to $k = 0$) is

$$\rho_c = \frac{3H^2}{8\pi G} , \quad H = \frac{\dot{S}}{S} . \tag{I.3}$$

Thus $k = -1$ models have density $\rho < \rho_c$ while $k = +1$ models have $\rho > \rho_c$. The constant H is of course the Hubble constant.

It is convenient to define

$$\Omega = \frac{\rho}{\rho_c} \quad , \quad q = -H^{-2} \frac{\ddot{S}}{S} \tag{I.4}$$

as the 'density parameter' and the 'deceleration parameter' respectively. We will denote the values of H, Ω and q at the present epoch ($t = t_0$) by H_0 , Ω_0 and q_0 and write

$$H_0 = h_0 \times 100 \text{ km s}^{-1} (\text{Mpc})^{-1} . \tag{I.5}$$

current estimates place h_0 in the range 0.5 to 1.

In more recent times the standard models have had an important input - that of the inflationary era brought in by the phase transitions of matter. Since the main thrust of these lectures is on nonstandard cosmologies, we will not discuss these and other details of the standard models here but assume them to be known. [For details see refs 1, 4 and 5]

Nonstandard cosmologies are many and it would not be possible to do justice to all of them within the limited timespan of these lectures. I will concentrate on some of those which have played significant roles in the ongoing cosmological debate. Even those that I describe are presented without too many details, but with the aim of giving their motivation, march of ideas and confrontation with observations. This last topic of observational tests will be considered at the end for all models taken together.

II THE STEADY STATE THEORY

II.1 MOTIVATION FOR A NON-BIG BANG COSMOLOGY

In 1948, around the same time that George Gamow was initiating detailed studies of the physical properties of the universe close to the big bang epoch, three astronomers proposed an entirely new approach to cosmology. This model, now famous (or notorious!) as the steady state model, does not have a singular big bang type epoch; indeed, it does not have either a beginning or an end on the cosmic time axis. The cosmological scene was considerably enlivened for two decades after the inception of the steady state model by the observers' attempts to shoot this rival model down. What was the motivation that led Hermann Bondi, Thomas Gold, and Fred Hoyle to the steady state cosmology?

First of all, in 1948 the measured value of $T_o \equiv H_o^{-1}$ was only $\sim 1.8 \times 10^9$ years. Consequently the age of a standard Friedmann model could not exceed T_o - a value lower than the geological age of the Earth! Thus a prima facie case existed for doubting the conclusion that the universe began $\sim$1 to 1.8 billion years ago.

Secondly, if a model (like the Friedmann models) proposes that the universe began at $t = 0$, it should provide a physical discussion of the beginning. At least it should leave the question tractable for a future, more sophisticated physical theory. The spacetime singularity at the $t = 0$ epoch precludes any such discussion. For example, the question as to how the matter and radiation we see around us came into existence in the first place remains unanswered.

Finally, on a more fundamental level we could raise the following doubt. The universe by definition contains everything-even the physical laws that describe the behavior of the matter and so on contained in it. Have we any guarantee that the physical laws that we use here and now have always remained the same? We could have assumed this to be the case had the universe itself not changed considerably in the course of time. This, however, was not the case for the Friedmann universes. A typical standard model changes considerably in its physical content and properties from soon after $t = 0$ to the present day. So the assumption that the laws of physics have remained unchanged throughout the history of the standard models is more an article of faith than a verifiable fact.

Today, as we shall see later, the age problem is still with us, although not in such a severe form as the low value of T_o in 1948 suggested. The questions of singularity and matter creation still remain with the standard models: the work discussed therein does not tell us what happened at $t = 0$. Hoyle's approach to the steady state theory was designed to attack the problem of primary creation of matter. His colleagues Bondi and Gold, however, considered the last issue discussed above as of paramount importance.

II.2 THE PERFECT COSMOLOGICAL PRINCIPLE

Bondi and Gold argued that the cosmological principle goes some way towards ensuring that the locally discovered laws of physics have universal validity: but it does not go far enough.

This principle tells us that at any given cosmic time t, all fundamental observers see the same large-scale features of the universe. Thus we are justified in assuming no spatial variation in the basic physical laws at any given cosmic time. But there is no justification from the cosmological principle to assume that the laws remain unchanged with time.

To provide such a justification Bondi and Gold strengthened the cosmological principle in what they called the perfect cosmological principle (PCP). The PCP states that in addition to the symmetries implicit in the cosmological principle, the universe in the large is unchanging with time. Thus the geometrical and physical properties of the hypersurfaces t = constant do not change with t.

It is important to emphasize the qualification "in the large". On a small enough scale the observed part of the universe will change. For example, stars in a galaxy will grow older, a small cluster of galaxies may evolve with time in shape and composition, and so on. However, according to the PCP the statistical properties on a large scale do not change.

For example, Hubble's constant should remain the same whether it is measured now or at any other time past or present, since its accurate measurement involves sampling a largish region in our neighborhood.

This requirement tells us immediately that

$$H = \frac{\dot{S}}{S} = \text{constant} = H_o . \qquad \text{(II.1)}$$

Further, the curvature of a t = constant hypersurface is given by k/S^2. This could in principle be measured at different times and found to be changing unless $k = 0$. Thus the PCP leads us to the unique line element.

$$ds^2 = c^2\, dt^2 - e^{2H_o t}\, [dr^2 + r^2\, (d\theta^2 + \sin^2\theta\, d\phi^2)]. \qquad \text{(II.2)}$$

Notice that we have arrived at the line element of the steady state universe without having to solve any field equations, as we had to do to determine $S(t)$ and k in standard cosmology. Bondi and Gold cited this result as an example of the deductive power of the PCP. Two other examples of deductions from this principle are given below.

Expansion of the Universe

The line element (II.2) is completely characterized by H_o. It is possible to have $H_o = 0$, $H_o < 0$, or $H_o > 0$, all consistent with the PCP. If, however, we take account of the local thermodynamic conditions, we are able to deduce that $H_o > 0$. For our observations show that the universe in our local neighborhood is far from being in a state of thermodynamic equilibrium. Stars radiate; regions of high and low temperatures exist within the Galaxy and outside it. If $H_o = 0$ we would have a static, infinitely old Euclidean universe. Such a universe should have reached a thermodynamic equilibrium by now, as implied by the Olbers paradox. If $H_o < 0$ we would have a contracting universe in which radiation from distant objects would be blueshifted, Such radiation would lead to an infinite radiation background, even worse than that indicated by the calculations of Olbers. Thus our local observations preclude $H_o \leq 0$, leaving the case $H_o > 0$, which is consistent with the finite and low night sky background. Hence the universe must expand: a conclusion arrived at without looking at any nearby galaxies!

Creation of Matter

It is easily seen that a proper 3-volume V bounded by fixed (r, θ, ϕ) coordinates increases with time as

$$V \propto \exp 3H_o t,$$

that is

$$\frac{\dot{V}}{V} = 3H_o. \qquad \text{(II.3)}$$

By the steady state hypothesis the density of the universe must remain constant at $\rho = \rho_o$. Then the amount of matter within

V must increase its mass $M \equiv V\rho_0$ as

$$\dot{M} = 3H_0 V \rho_0 .$$

In other words,

$$Q = 3H_0\rho_0 \qquad \text{(II.4)}$$

denotes the rate of creation of matter per unit volume. If we use cgs units we get

$$Q = 2 \times 10^{-46} \frac{\rho_0}{\rho_c} h_0^3 \text{ g cm}^{-3} \text{ s}^{-1}, \qquad \text{(II.5)}$$

where ρ_c and h_0 have been defined in §I.

The small value of Q shows that there is a very slow but continuous creation of matter going on, in contrast to the explosive creation at $t = 0$ of the standard models.

II.3 THE CREATION FIELD

Attractive though the above deductive approach is, it has its limitations. For example, we do not have a quantitative relation connecting H_0 to say, the mean density ρ_0 as we have in Friedmann cosmologies. Nor do we have any physical theory for such an important phenomenon as the continuous creation of matter. Is the sacrosanct law of conservation of matter and energy being violated in the process of matter creation? Bondi and Gold appreciated the fact that questions like these could be answered through a dynamic theory rather than from their deductive approach. However, they felt that the PCP together with local observations fix the large-scale properties of the universe in a form that can be tested by observations (see § II.4). Therefore they attached a greater importance to testing the PCP by observations than to a dynamic theory that might determine H_0, ρ_0, and so on quantitatively.

Fred Hoyle, on the other hand, took the opposite view. He looked for a process - that is, a field theory - that could account for the phenomenon of primary creation of matter. After several attempts he finally adopted the formulation suggested by M.H.L. Pryce. This formulation, known as the C-field theory, was used extensively by Hoyle and the author in the early 1960s. The details of the C-field theory are given below.

The Action Principle

The C-field theory involves adding more terms to the standard Einstein-Hilbert action to represent the phenomenon of creation of matter. Using Occam's razor, the additional field to be introduced is a scalar field with zero mass and zero charge. We denote this field by C and its derivative with respect to the spacetime coordinate x^i by C_i. The action is then given by

$$\mathcal{A} = \frac{c^3}{16\pi G} \int R \sqrt{-g}\, d^4x - \sum_a m_a c \int ds_a$$

$$- \frac{1}{2c} f \int C_i C^i \sqrt{-g}\, d^4x + \sum_a \int C_i \, da^i. \qquad \text{(II.6)}$$

Instead of the electromagnetic terms (which might be present if we had charged particles), we have in (II.6) the C-field terms. To appreciate the difference between the two interactions, note that the last term of (II.6) is path-independent. If we consider the world line of particle a between the end points A_1 and A_2, we have

$$\int_{A_1}^{A_2} C_i \, da^i = C(A_2) - C(A_1).$$

Normally such path-independent terms do not contribute to any physics derivable from the action principle. So why include such a term? The answer to this question lies in the notion of "broken" world lines. A theory that discusses creation (or annihilation) of matter per se must have world lines with finite beginnings or ends (or both). The C-field interaction term picks out precisely these end points of particle world lines. If we vary the world line of a and consider the change in the action $\mathcal{A}$ in a volume containing the point A_1 where the world line begins (see Figure 1), we get at A_1 (which is now varied)

$$m_a \, c \, \frac{da^i}{ds_a} \, g_{ik} - C_k = 0. \qquad \text{(II.7)}$$

This relation tells us that overall energy and momentum are conserved at the creation point. The 4-momentum of the created particle is compensated by the 4-momentum of the C-field. Clearly, to achieve this balance the C-field must have negative energy. We will return to this point later.

We also note that since the interaction term is path-independent, the equation of motion of a is still that of a geodesic:

$$m_a \left[\frac{d^2 a^i}{ds_a^2} + \Gamma^i_{kl} \frac{da^k}{ds_a} \frac{da^l}{ds_a} \right] = 0 \qquad \text{(II.8)}$$

The constant f in the action (II.6) is a coupling constant. The variation of C gives the source equation in the form

$$C^k{}_{;k} = cf^{-1} n \qquad \text{(II.9)}$$

where n = number of net creation events per unit proper 4-volume. In calculating n we attach a + sign to the points like A_1 where a world line begins and - sign to the points like A_2 where a world line ends. Again we see in (II.9) the relationship between the C-field and the creation/annihilation events.

Finally, the variation of g_{ik} leads to the modified Einstein field equations

$$R^{ik} - \frac{1}{2} g^{ik} R = - \frac{8\pi G}{c^4} \left\{ \underset{(m)}{T^{ik}} + \underset{(c)}{T^{ik}} \right\} \qquad \text{(II.10)}$$

where $\underset{(m)}{T^{ik}}$ is the matter tensor as in standard cosmology while

$$\underset{(c)}{T^{ik}} = - f \quad C^i C^i - \frac{1}{2} g^{ik} C^l C_l \quad . \qquad \text{(II.11)}$$

Again we note that $T^{44}_{(c)} < 0$ for $f > 0$. Thus the C-field has negative energy density that produces a repulsive gravitational effect. It is this repulsive force that drives the expansion of the universe.

The above effect may resolve one difficulty usually associated with the quantum theory of negative energy fields. Because such fields have no lowest energy state, they normally do not form stable systems. A cascading into lower and lower energy states would inevitably occur if we perturb the field in a given state of negative energy. However, this conclusion is altered if we include the feedback of (II.11) on spacetime geometry. This feedback results in the expansion of space and in the lowering of the magnitude of field energy. Both these effects tend to stabilize the system.

Cosmological Equations

Using the Robertson-Walker line element and the assumption that a typical particle created by the C-field has mass m, we get the following equations out of (II.7) through (II.11):

$$\dot{C} = mc^2 \tag{II.12}$$

$$mf \left\{ \ddot{C} + 3 \frac{\dot{S}}{S} \dot{C} \right\} = \left\{ \dot{\rho} + 3 \frac{\dot{S}}{S} \rho \right\} c^2 \tag{II.13}$$

$$2\,\frac{\ddot{S}}{S} + \frac{\dot{S}^2 + kc^2}{S^2} = \frac{4\,Gf}{c^4}\,\dot{C}^2 \tag{II.14a}$$

$$3\,\frac{\dot{S}^2 + kc^2}{S^2} = 8\pi G\left\{\rho - \frac{f}{2c^4}\,\dot{C}^2\right\}. \tag{II.14b}$$

It is easy to verify that the steady state solution (II.2) follows from these equations for

$$k = 0,\quad S = e^{H\,t},\ \rho = \rho_o = \frac{3H^2_o}{4\pi G} = fm^2\ . \tag{II.15}$$

Notice that both H_o and ρ_o are given in terms of the elementary creation process; that is, in terms of the coupling constant f and the mass of the particle created. Thus the Hoyle approach gives the quantitative information lacking in the deductive approach of the PCP.

A first order perturbation of the above equations and of the solution (II.15) also tells us that the solution is stable. Indeed, a stability analysis brings out the key role played by (II.7). This tells us that the created particles have their world lines along the normals to the surfaces C = constant. Hoyle has argued that such a result gives a physical justification for the Weyl postulate: it tells us why the world lines of the fundamental observers are orthogonal to a family of space-like hypersurfaces. In the C-field cosmology these hypersurfaces are not just abstract notions but have a physical basis.

Because the velocity field given by (II.7) is the gradient of a scalar, the vorticity or spin vector ω^i must be zero. Thus spinning universes would be excluded. Hoyle and the author have argued that provided matter creation is always going on, the newly created matter will tend to introduce the regularity implicit in the Weyl postulate. Thus they conjectured that starting from arbitrary initial conditions the universe may be driven by the C-field creation process to the homogeneous and spinless state it is in today. Hence the coincidence of §III.1 that led to Mach's principle is explained as a dynamic outcome of matter creation.

Explosive Creation

Although the C-field was introduced primarily to account for the continuous creation of matter, the author showed in 1973 that it also describes explosive matter creation such as is required in the big bang cosmology. We illustrate below how this is achieved for the case $k = 0$.

In equations (II.12) through (II.14), we make use of the idea that all matter is created in an explosive process at $t = 0$. Then the right-hand side of (II.13) is like a delta function $\delta(t)$, leading to the solution

$$\dot{C} = \frac{A}{S^3} .$$

Notice that this solution is inconsistent with (II.12) except at one epoch, $t = 0$. This is hardly surprising, since we have assumed no creation of matter subsequent to $t = 0$. Thus the creation condition (II.9) is not satisfied at $t > 0$.

Substituting for $\dot{C}$ in (II.14a) we can integrate for S and obtain a solution

$$S(t) = \left[1 + \frac{(t + t_1)^2}{t_o^2}\right]^{1/3} \qquad \text{(II.16)}$$

where t_o and t_1 are constants related to the initial conditions at $t = 0$.

The scale factor given by (II.16) behaves like that for the standard Einstein-de Sitter model for $t \gg t_o$. In the C-field model not only is the spacetime singularity at $t = 0$ averted, but we also see the present matter as arising from a primordial explosion that conserves energy and momentum.

This conservation of energy and momentum must follow as a general deduction for any C-field model, since the governing equations are derived from an action principle. Hence criticism based on the unexplained origin of new matter, which could be validly applied to the explosive creation of the standard cosmology or to the continuous creation in the Bondi-Gold version of the steady state model, does not apply to the C-field cosmology.

In physical terms the creation is explained by a process of interchange of energy and momentum between the negative energy C-field and the matter. The divergence of (II.10) gives the mathematical formula for energy conservation:

$$\underset{(m)}{T}^{ik}{}_{;k} = fC^{i}C^{k}{}_{;k}. \qquad \text{(II.17)}$$

It is easy to verify that the idea would not work for a positive energy field.

II.4 OBSERVABLE PARAMETERS OF THE STEADY STATE THEORY

Leaving aside the dynamics of the model, we now come to some of the observable features of the steady state theory. Here we deal essentially with the line element (II.2) and the geometrical properties deducible from it. Indeed, as Bondi and Gold emphasized in their original paper, the steady state model makes precise predictions and is therefore vulnerable to observational disproof, in contrast to the big bang models, which can always be fed with arbitrary parameters. (This comment will become clearer when we discuss observational cosmology later).

Since these calculations of observable features are similar to those for standard models, We will be brief here and simply state the results.

The Redshift

The redshift of a galaxy G_1 at (r_1, θ_1, ϕ_1) emitting light at t_1 that is received by the observer O at $r = 0$ at the present epoch t_o is given by

$$z_1 = e^{H_o(t_o - t_1)} - 1 = r_1 \frac{H_o}{c} e^{Ht_o} . \qquad (II.18)$$

The Luminosity Distance

This is given for the above galaxy by

$$D_1 = \frac{c}{H_o} z_1 (1 + z_1) . \qquad (II.19)$$

Equation (II.19) is the Hubble law for steady state cosmology.

Angular Size

The angle $\Delta\theta(\ll 1)$ subtended at O by an astronomical source of projected linear size d and redshift z is given by

$$\Delta\theta = \frac{H_o}{c} d . \left\{ \frac{1+z}{z} \right\} . \qquad (II.20)$$

Thus the angular size tends to a finite minimum as $z \to \infty$.

Flux Density

The formula for bolometric flux becomes in this case

$$\mathcal{F}_{bol} = \frac{L_{bol}}{4\pi \frac{c}{H_o}^2 z^2 (1 + z)^2} . \qquad (II.21)$$

For flux density at frequency ν_o we get

$$\tilde{\mathcal{J}}(\nu_o) = \frac{L\, J(\nu_o\ 1 + z)}{4\pi \frac{c}{H_o}^{2}\ z^2\ (1 + z)} \qquad \text{(II.22)}$$

Number Count

The number of sources with redshift less than z is given by

$$N(z) = 4\pi n \left(\frac{c}{H_o}\right)^3 \left[\ln\ (1+z) - \frac{3z^2 + 2z}{2(1 + z)^2} \right] . \qquad \text{(II.23)}$$

The Age Distribution of Galaxies

New galaxies are always being formed in the steady state universe. Since the universe expands, the galaxies, once formed, move away from each other. Thus the older a population of galaxies, the more sparse its distribution will be. Since the volume bounded by galaxies increases with time as $\exp(3H_ot)$, we have the following simple result for the age-density relation of galaxies:

$$Q(\tau) = e^{-3Ho\tau} \qquad \text{(II.24)}$$

where $Q(\tau)d\tau$ is the proper number number density of galaxies with ages in the range $\tau, \tau + d\tau$. The average age is therefore

$(3H_o)^{-1}$. However, caution is necessary in the interpretation of (II.24), as we shall see in the following section.

II.5 PHYSICAL AND ASTROPHYSICAL CONSIDERATIONS

This section briefly outlines some of the ideas proposed from time to time in the context of the steady state theory to discuss such problems as the nature of created particles, the formation of galaxies, the origin of the microwave background radiation, and so on. Some of these concepts might still be relevant whether or not the steady state cosmological picture survives.

The Hot Universe

In 1958 Gold and Hoyle proposed the hypothesis that the created matter was in the form of neutrons. The creation of neutrons does not violate any standard conservation laws of particle physics except the constancy of the baryon number. Although this was considered an objection in 1958, today the baryon number is no longer regarded as invariant. Indeed today scenarios are being proposed in the context of the early universe to account for the observed baryon number in the universe. In the Gold-Hoyle picture the created neutron undergoes a β-decay:

$$n \rightarrow p + e^- + \bar{\nu} \ . \qquad \text{(II.25)}$$

The conservation of energy and momentum results in the electron taking up most of the kinetic energy and thereby acquiring a high kinetic temperature of $\sim 10^9$K. Gold and Hoyle argued that such a high temperature produced inhomogeneously, would lead to the working of heat engines between the hot and cold regions, which in turn would result in condensations of the size of $\gtrsim$ 50 Mpc, while pure gravitational forces are known not able to provide a satisfactory picture of galaxy formation. The temperature gradients set up in the hot universe of Gold and Hoyle help in this process.

The resulting system, however, is not a single galaxy, but a supercluster of galaxies containing $\sim 10^3$ to 10^4 members. Such large-scale inhomogeneities in the distribution of galaxies were first referred to mainly through the work of G.O. Abell and G.de Vaucouleurs.

It is worth noting that inhomogeneities on such a large scale as $\sim$ 50 Mpc caution us against applying the cosmological principle too rigorously. For example, the formula (II.24) for the age distribution of galaxies will hold over a region considerably larger than 50 Mpc in such a model. If we are in a particular supercluster, we expect to see a preponderance of galaxies of age similar to that of ours in our neighborhood out to say 20 or 30 Mpc. Thus it will not be surprising if our local sample yields an average age much larger than the universal average of $(3\,H_o)^{-1} \simeq 3 \times 10^9\,h_o^{-1}$ years.

Although newly created electrons have a kinetic temperature of $\sim 10^9$K, the temperature tends to drop because of expansion.

The average temperature is 3/5 of this value, that is, around 6×10^8 K. It was suggested by Hoyle in 1963 that such a hot intergalactic medium would generate the observed X-ray background. However, quantitative estimates by R.J. Gould soon showed that the expected X-ray background in the hot universe would be considerably higher than what is actually observed, thus making the hot universe untenable. Although the present background measurements do not rule out such a hot universe for $h_o = 0.5$, astrophysicists are inclined to look for other explanations for the origin of the X-ray background.

The Bubble Universe

In 1966 Hoyle and the author discussed the effect of raising the coupling constant f by $\sim 10^{20}$. As the formulae (II.15) show, we would then have a steady state universe of very large density ($\rho_o \sim 10^{-8}$ g cm^{-3}) and very short time scale ($H_o^{-1} \sim 1$ year!). If in such a dense universe creation is switched off in a local region, that is, if we locally have

$$C^i_{;i} = 0, \qquad \text{(II.26)}$$

then this local region will expand according to (II.16). Being less dense than the surroundings, such a region will simulate an air bubble in water.

According to this model, this bubble is all that we see with our surveys of galaxies, quasars, and so on. Hence our observations tell us more about this unsteady perturbation than about the ambient steady state universe. There are, however, observable effects that give indications of the high value of f.

For example, these authors showed that particle creation is enhanced near already existing massive objects and that the resulting energy spectrum of the particles would simulate that of high-energy cosmic rays. Although this result applies for any value of f, the actual energy density of cosmic rays requires the high value of f chosen here.

Another useful idea to come out of this picture was that galaxies, especially the elliptical ones, are examples of a small bubble whose expansion is controlled by a local massive object. The basic calculation is given below. Now that it is being realized that ellipticals cannot have arisen from condensation of a pregalactic cloud the above expansion idea may well contain a germ of the truth.

Consider first an expanding bubble as a cloud of gas moving radially outwards. The Einstein-de Sitter model is simulated by this cloud, in which each particle has just the right velocity to escape to infinity:

$$\dot{r}^2 = \frac{2GM(r)}{r} . \tag{II.27}$$

Here $M(r)$ is the mass interior to radius r.

Suppose now that when $r = r_o$ an object of mass μ appears at the origin. The appearance of this object will influence the subsequent motion of the cloud according to the equation

$$\dot{r}^2 = \frac{2G[M(r) + \mu]}{r} - \frac{2G\mu}{r_o} \quad . \tag{II.28}$$

Here we have assumed that the velocities are not affected by the introduction of μ; thus (II.28) is continuous with (II.27) at $r = r_o$. The mass μ now exerts its gravitational pull so that the cloud is unable to escape to infinity. In (II.28) r attains a maximum value given by

$$r_{max} = \left\{1 + \frac{M}{\mu}\right\} r_o \approx \frac{Mr_o}{\mu} \text{ for } M \gg \mu. \tag{II.29}$$

In our description it is assumed that there is no inward/outward crossing of cloud particles, so $M(r)$ is fixed for each cloud element.

What is r_o? This radius can be fixed in the following way. Although we expect μ to be small compared to M, we cannot use the above Newtonian calculation unless r_o is large enough so that any general relativistic corrections to our calculation are negligible. At r_o, these corrections are of the order $(2GM/r_oc^2)$. This quantity must be smaller than the Newtonian quantity $2G\mu/r_oc^2$. This requirement gives us a lower limit on r_o:

$$r_o \geq \frac{M}{\mu} \cdot \frac{2GM}{c^2} \quad . \tag{II.30}$$

Setting r_o equal to the right-hand side gives for r_{max}

$$r_{max} \cong \frac{2GM}{c^2} \left(\frac{M}{\mu}\right)^2 . \qquad \text{(II.31)}$$

If we now set $r_{max} \cong 3 \times 10^{22}$ cm, a typical galactic radius, (II.31) gives

$$\frac{M}{M_{\odot}} \approx 5 \times 10^5 \left(\frac{\mu}{M_{\odot}}\right)^{2/3} . \qquad \text{(II.32)}$$

Thus a central condensation of $\mu \sim 10^9\ M_{\odot}$ can control the shape of a galaxy of mass $M \sim 5 \times 10^{11}\ M_{\odot}$. Anisotropy of expansion can lead to ellipsoidal shapes, the important result being that galaxies formed this way should have no rotation. Apart from the lack of rotation, ellipticals are now believed to contain massive nuclei, the most dramatic discovery in recent years being that of a supermassive object at the center of the galaxy M87.

The Origin of Elements

One of the beneficial influences of the steady state cosmology on astrophysics was that it prompted work on stellar nucleosynthesis. Since the model does not have a high-temperature epoch, it cannot draw on the calculations given in Chapter 5 to explain how nuclei are made from protons and neutrons.

Since centers of stars provide sites for high temperature and density, astrophysicists looked for nucleosynthesis in such places. The pioneering work of E.M. Burbidge, G.R. Burbidge, W.A. Fowler, and F. Hoyle in 1957 demonstrated in a comprehensive manner how the whole observed range of nuclei can be produced in stellar processes as stars evolve. Thus it became established that the bulk of the nuclei are produced in stars rather than in the early hot universe, as Gamow had envisaged.

Light Nuclei and the Microwave Background

Between 1964 and 1965 the steady state model received two near-fatal blows. The realization that the observed helium abundance in several parts of the Galaxy is considerably higher than that generated in the stars led astronomers back to Gamow's ideas once again. The case for the hot big bang became even stronger with the discovery of the microwave background in 1965.

The steady state model has not quite recovered from these two blows. Indeed, if it is to survive as a viable alternative to the big bang it must produce an astrophysical interpretation for both the above observations, as well as for the observed abundances of other light nuclei besides helium, like deuterium, Li, Be, and so on.

Energetically, it is realized that increased stellar activity is required to account for the observed helium, and the resulting additional starlight has to be thermalized to produce the microwave background possible scenarios exist in which dust grains in the intergalactic space act as thermalizers. In working such scenarios into the steady state model a further constraint has to be placed on any calculations. This is the constraint demanded by the PCP, that is, that the universe in the past was no different from the way it is now.

The main difficulties of such attempts are as follows. Although increased stellar activity can generate sufficient helium, the production of deuterium in stars (or supermassive objects) has not proved so easy, since the deuterium produced is quickly destroyed. Also, the extreme homogeneity of the microwave background places severe limits on any theory that attempts to generate it from discrete sources.

Although attempts have been made by pro-steady state astrophysicists to construct a viable explanation, the goal has not yet been achieved. Nor has the problem been abandoned as insoluble; so the fate of the steady state model hangs in the balance!

Inflation and the C-field

There is considerable similarity between the C-field cosmology and the inflationary scenarios currently fashionable. The idea of negative stresses of C-field cosmology is echoed in the negative stresses of the vacuum of inflationary models. The bubble universe model in which locally the C-field switches from the creative to noncreative mode is very similar to the emergence of the Friedmann universe when phase transition is completed in a given region of the inflationary universe. Further, the result that inflation wipes out memories of the state of the universe existing before-hand was obtained earlier for the C-field cosmology by F. Hoyle and the author.

III MACHIAN COSMOLOGIES

III.1 INERTIA AND COSMOLOGY

Next we will consider some of the cosmological models inspired by the ideas of the philosopher - scientist Ernst Mach. Mach's principle itself arose out of the observation that the local inertial frame, earlier identified by Newton as the absolute space, is one relative to which the distance parts of the universe are nonrotating. Let us now examine the possible implications of this observation further.

When expressed in the framework of the absolute space, Newton's second law of motion take the familiar form

$$\underset{\sim}{P} = m\underset{\sim}{f} \ . \tag{III.1}$$

This law states that a body of mass m subjected to an external force $\underset{\sim}{P}$ experiences an acceleration $\underset{\sim}{f}$. Let us denote by S the coordinate system in which $\underset{\sim}{P}$ and $\underset{\sim}{f}$ are measured.

Newton was well aware that his second law has the simple form (III.1) only with respect to S and those frames that are in uniform motion relative to S. If we choose another frame S' that has an acceleration $\underset{\sim}{a}$ relative to S, the law of motion measured in S' becomes

$$\underset{\sim}{P}' \equiv \underset{\sim}{P} - m\underset{\sim}{a} = m\underset{\sim}{f}' \ . \tag{III.2}$$

Although (III.2) outwardly looks the same as (III.1), with $\underset{\sim}{f}'$ the acceleration of the body in S', something new has entered into the force term. This is the term $m\underset{\sim}{a}$, which has nothing to do with the external force but depends solely on the mass m of the body and the acceleration a of the reference frame relative to the absolute space. Realizing this aspect of the additional force in (III.2), Newton termed it "inertial force." As this name implies, the additional force is proportional to the inertial mass of the body. Newton discusses this force at length in his Principia, citing the example of a rotating water-filled bucket in which the water surface is curved due to such forces.

According to Mach, the Newtonian discussion was incomplete in the sense that the existence of the absolute space was postulated arbitarily and in an abstract manner. Why does S have a special status in that it does not require the inertial force? How can one identify S without recourse to the second law of motion, which is based on it?

To Mach the answers to these questions were contained in the observation of the distant parts of the universe. It is the universe that provides a background reference frame that can be identified with Newton's frame S. Instead of saying that it is an accident that Earth's rotation velocity relative to S agrees with that relative to the distant parts of the universe, Mach took it as proof that the distant parts of the universe somehow enter into the formulation of local laws of mechanics.

One way this could happen is by a direct connection between the property of inertia and the existence of the universal background. To see this point of view, imagine a single body in an otherwise empty universe. In the absence of any forces (III.1) becomes

$$m\ddot{\underset{\sim}{r}} = \underset{\sim}{0} \ . \tag{III.3}$$

What does this equation imply? Following Newton we would conclude that $\ddot{\underset{\sim}{r}} = \underset{\sim}{0}$, that is, the body moves with uniform velocity. But we now no longer have a background against which to measure velocities! Thus $\dot{\underset{\sim}{r}} = \underset{\sim}{0}$ has no operational significance. Rather, $\underset{\sim}{\dot{r}}$ should be completely indeterminate. And it is not difficult to see that such a conclusion is not inconsistent with (III.3) provided we argue that

$$m = 0. \tag{III.4}$$

In other words, the measure of inertia depends on the existence of the background in such a way that in the absence of the background the measure vanishes! This aspect introduces a new feature into mechanics not considered by Newton. The Newtonian view that inertia is the property of matter has to be augmented to the statement that inertia is the property of matter as well as of the background provided by the rest of the universe.

Such a Machian viewpoint not only modifies local mechanics, but it also introduces new elements into cosmology. For, except in the universe following the perfect cosmological principle, there is no basis now for assuming that particle masses would necessarily stay fixed in an evolving universe. This is the reason for considering cosmological models anew from the Machian viewpoint. The ideas presented here give some instances of how different physicists have given quantitative expression to Mach's principle.

II.2 THE BRANS-DICKE THEORY OF GRAVITY

In 1961 C. Brans and R.H. Dicke provided an interesting alternative to general relativity based on Mach's principle. To understand the reasons leading to their field equations, we first note that the concept of a variable inertial mass just arrived at itself leads to a problem of interpretation. For how do we compare masses at two different points in spacetime? Masses are measured in certain units, such as masses of elementary particles, which are themselves subject to change! We need an independent unit of mass against which an increase or decrease of a particle mass can be measured. Such a unit is provided by gravity, the so called Planck mass:

$$\left(\frac{\hbar c}{G}\right)^{1/2} \cong 2.16 \times 10^{-5} \text{ g}. \tag{III.5}$$

Thus the dimensionless quantity

$$\chi = m\left(\frac{G}{\hbar c}\right)^{1/2} \tag{III.6}$$

measured at different spacetime points can tell us whether masses are changing. Or alternatively, if we insist on using mass units that are the same everywhere, a change of χ would tell us that G is changing. This is the conclusion Brans and Dicke arrived at in their approach to Mach's principle . They looked for a framework in which the gravitational constant G arises from the structure of the universe, so that a changing G could be looked upon as the Machian consequence of a changing universe.

In 1953 D. W. Sciama gave general arguments leading to a relationship between G and the large-scale structure of the universe. We have already come across one example of such a relation in Friedmann cosmologies;

$$\rho_0 = \frac{3H_0^2}{4\pi G} q_0 .$$

If we write $R_0 = c/H_0$ as a characteristic length of the universe and $M_0 = 4\pi\rho_0 R_0^3/3$ as the characteristic mass of the universe, then the above relation becomes

$$\frac{1}{G} = \frac{M_0}{R_0 c^2} \sim q_0^{-1} \sim \frac{M_0}{R_0 c^2} \sim \sum \frac{m}{rc^2} . \qquad \text{(III.7)}$$

Given a dynamic coupling between inertia and gravity, a relation of the above type is expected to hold. Brans and Dicke took this relation as one that determines G^{-1} as a linear superposition of inertial contributions from masses all over the universe, a typical contribution m/rc^2 being from a mass m at a distance r from the point where G is measured. Since m/r is a solution of a scalar wave equation with a point source of strength m, Brans and Dicke postulated that G behaves as the reciprocal of a scalar field ϕ:

$$G \sim \phi^{-1}, \tag{III.8}$$

where ϕ is expected to satisfy a scalar wave equation whose source is all the matter in the universe.

The Action Principle

These intuitive concepts are contained in the Brans-Dicke action principle, which may be written in the form

$$\mathcal{A} = \frac{c^3}{16\pi} \int_{\mathcal{V}} (\phi R + \omega\phi^{-1} \phi^k \phi_k) \sqrt{-g} d^4x + \Lambda \tag{III.9}$$

Notice first that the coefficient of R is $c^3\phi/16\pi$ instead of $c^3/16\pi G$ as in the Einstein-Hilbert action. The reason for this lies in the anticipated behavior of G as given in (III.8). The second term, with $\phi_k \equiv \partial\phi/\partial x^k$, ensures that ϕ will satisfy a wave equation, while the third term includes, through a Lagrangian density L, all the matter and energy present in the spacetime region $\mathcal{V}$. The energy momentum tensor T^{ik} is derived from by variation of g_{ik}. ω is a coupling constant.

The variation of $\mathcal{A}$ for small changes of g^{ik} leads to the field equations

$$R_{ik} - \frac{1}{2} g_{ik} R = - \frac{8\pi}{c^4\phi} T_{ik} - \frac{\omega}{\phi^2} \left\{ \phi_i\phi_k - \frac{1}{2} g_{ik} \phi^l\phi_l \right\}$$

$$- \frac{1}{\phi} (\phi_{ik} - g_{ik} \Box \phi). \tag{III.10}$$

Similarly, the variation of ϕ leads to the following equation for ϕ :

$$2\phi \Box \phi - \phi_k \phi^k = \frac{R}{\omega} \phi^2. \tag{III.11}$$

This latter equation can be simplified by substituting for R from the contracted form of (III.10). We finally get

$$\Box \phi = \frac{8\pi}{(2\omega + 3)c^4} T, \tag{III.12}$$

where T is the trace of $T^i{}_k$. Thus (III.12) leads to the anticipated scalar wave equation for ϕ with sources in matter.

Because it contains a scalar field ϕ in addition to the metric tensor g_{ik}, the Brans-Dicke theory is often referred to as the scalar-tensor theory of gravitation.

Solar System Measurements of ω

It is clear from these field equations that as $\omega \to \infty$ the Brans-Dicke theory tends to general relativity. For $\omega \approx O(1)$ the theory makes significantly different predictions from general relativity in a number of Solar System tests. These tests are the same as those for general relativity.

The computation of perihelion precession of the planet Mercury gives the theoretical prediction of this theory as $(3\omega + 4)/(3\omega + 6)$ times the value given by general relativity. Dicke and his colleagues suggested during the 1970s that if the Sun is oblate, with a quadrupole moment parameter of $\sim 2.5 \times 10^{-5}$, then the resulting change in its gravitational field would lead to a perihelion precession of about 7% of the observed (unexplained) value of $\sim$ 43 arc second per century. Had this been the case the relativistic value of $\sim$ 43 arc second would have been too high, while a Brans-Dicke value of $\omega = 6$ would have correctly accounted for the residual of $\sim$ 40 arc second per century. However, external studies of the Sun's surface do not conform with oblateness even of this order. Hence this test does not give any evidence for ω as small as 6.

The bending angle of a light ray grazing massive spherical object in the Brans-Dicke theory is $(2\omega + 3)/(2\omega + 4)$ of the relativistic value. Since the accuracy of the radio and microwave measurements of the bending angle is $\sim$ 5% and the angle agrees with the relativistic value within this error, the parameter ω has to be as high as $\sim$ 10.

The lunar laser-ranging experiments, however, lead to the conclusion that $\omega > 29$. Here again the general relativistic value of the Earth-Moon distance is in excellent agreement with observations, and any departures from it, if they are to be tolerated by the observations, have to be small enough to demand a large value of ω. Radar ranging to probe landers on Mars places an even more severe limit on ω by requiring that $\omega \gg 500$.

It therefore follows that at the Solar System level the Brans-Dicke theory has to have a large value of ω in order to survive, thus making it practically indistinguishable from general relativity.

However, even for a large ω this theory can produce interesting departures from general relativity at the cosmological level. The following section outlines these differences.

III.3 COSMOLOGICAL SOLUTIONS IN THE BRANS-DICKE THEORY

We will consider only the homogeneous and isotropic cosmological models in the Brans-Dicke theory. Accordingly we start with the Robertson-Walker line element and the energy tensor for a perfect fluid, as in standard cosmology. The scalar-field ϕ is now a function of the cosmic time only. Thus the field equations become

$$\frac{2\ddot{S}}{S} + \frac{\dot{S}^2 + kc^2}{S^2} = -\frac{8\pi p}{\phi c^2} - \frac{2\dot{\phi}\dot{S}}{\phi S} - \frac{\omega\dot{\phi}^2}{2\phi^2} - \frac{\ddot{\phi}}{\phi} \qquad \text{(III.13)}$$

$$\frac{\dot{S}^2 + kc^2}{S^2} = \frac{8\pi\varepsilon}{3\phi c^2} - \frac{\dot{\phi}\dot{S}}{\phi S} + \frac{\omega\dot{\phi}^2}{6\phi^2} . \qquad \text{(III.14)}$$

Compare these equations with the corresponding ones of the Friedmann cosmologies. The conservation equation is the same:

$$\frac{d}{dS}(\varepsilon S^3) + 3pS^2 = 0 . \qquad \text{(III.15)}$$

In addition, we have the field equation for ϕ :

$$\frac{1}{S^3}\ \frac{d}{dt}\,(\dot{\phi}S^3) = \frac{8\pi}{(2\omega + 3)c^2}\,(\epsilon - 3p). \tag{III.16}$$

We anticipate that big bang solutions will emerge from these equations and set the big bang epoch at $t = 0$. Then the integral of (III.16) gives

$$\dot{\phi}S^3 = \frac{8\pi}{(2\omega + 3)c^2} \int_0^t (\epsilon - 3p)S^3\,dt + C, \tag{III.17}$$

where C is a constant. Two types of solutions are obtained, depending on whether $C = 0$ or $C \neq 0$.

$C = 0$

We will consider a simple example of this type, with $k = 0$, $p = 0$, $\epsilon = \rho c^2$. This solution is therefore analogous to the Einstein-de Sitter model of general relativity. Write

$$S = S_o \left(\frac{t}{t_o}\right)^A, \quad \phi = \phi_o \left(\frac{t}{t_o}\right)^B . \tag{III.18}$$

So that $p \propto t^{-3A}$ and the field equations give

$$A = \frac{2\omega + 2}{3\omega + 4} \ , \quad B = \frac{2}{3\omega + 4} \ , \tag{III.19}$$

and

$$\rho_o = \frac{(2\omega - 3)B\phi_o}{8\pi t_o^2} \ . \tag{III.20}$$

The temporal behavior of S and $G(\propto\phi^{-1})$ is illustrated in Figure 2. It can be verified that as $\omega \to \infty$ this solution tends to the Einstein-de Sitter model.

An analogue of the radiation model can be obtained in this theory. H. Nariai obtained solutions for $p = n\epsilon$ with n in the range $0 \leq n \leq 1/3$.

$C \neq 0$

In this case the ϕ-terms dominate the dynamics of the universe in the early stages. Thus for small enough t we have

$$\frac{8\pi}{(2\omega + 3)c^2} \int_o^t (\epsilon - 3p) \ S^3 \ dt < |C| , \tag{III.21}$$

both for the cases of dust and of radiation. For our power law solutions for the case $p = 0$, we have at small enough t

$$3A + B = 1. \quad t_o = \frac{S_o^3 \phi_o^B}{C} . \tag{III.22}$$

In the case of a radiation-dominated universe $p = 1/3\varepsilon$ and we can again try a solution of the form (III.18) to get as $t \to 0$

$$A^2 = - AB + \frac{\omega B^2}{6} . \tag{III.23}$$

Taking into account (III.22) we can solve (III.23) to get

$$A = \frac{\omega + 1 \pm \sqrt{(2\omega/3) + 1}}{3\omega + 4} . \quad B = \frac{1 \pm 3\sqrt{(2\omega/3) + 1}}{3\omega + 4} . \tag{III.24}$$

The upper sign holds when $C > 0$ and the lower sign when $C < 0$. For $C > 0$, $\phi \to 0$ when $S \to 0$, while for $C < 0$, $\phi \to \infty$ for $S \to 0$. These conclusions hold irrespective of the values of k or of the equation of state, since at small values of S the dynamics of the universe are controlled by the ϕ-term.

TABLE 1

MASS FRACTIONS OF ^{2}H AND ^{4}He IN
BRANS-DICKE COSMOLOGY FOR
MATTER-DOMINATED MODELS*

	ρ_o (g cm^{-3})		
ω	10^{31}	10^{-30}	10^{-29}
5	7.6×10^{-4}	2.6×10^{-5}	3.4×10^{-8}
	0.26	0.33	0.40
10	7.6×10^{-4}	2.1×10^{-5}	$\sim 10^{-9}$
	0.26	0.30	0.35
∞	6.6×10^{-4}	1.3×10^{-5}	$\sim 10^{-11}$
	0.25	0.27	0.29

* The deuterium fraction is given above the helium fraction.

Production of Light Nuclei

Dicke and G.S. Greenstein independently investigated the nucleosynthesis problem in the early Brans-Dicke universe. Greenstein followed the same physical approach as for standard cosmologies, for the case $C = 0$. The results obtained by him for $h_o = 1$ are given in Table 1.

For each of three values of the present density of matter ρ_0. Table 1 gives three sets of values for the deuterium and helium abundance, corresponding to $\omega = 5$, $\omega = 10$, and $\omega = \infty$. The last case is of course that of general relativity. The differences between the Brans-Dicke theory and general relativity are noticeable for $\omega = 5$ at high values of ρ_0 when more ^2H and ^4He are formed in the former theory. For $\omega > 30$, the present observed abundances set an upper limit of $\rho_0 < 5 \times 10^{-30}$ g cm^{-3} in the Brans-Dicke cosmology.

In the ϕ-dominated models the constant C can be adjusted to produce any desirable abundances, high or low. For cosmic abundances lower than the above value one has to choose suitably low value of $|C|$.

There is, however, another observational handle on C, which is described briefly below.

The Variation of G

Since $G \propto \phi^{-1}$, a time-dependent ϕ will mean a time-dependent gravitational constant. As seen from (III.18), we have for $C = 0$

$$\frac{\dot{G}}{G} = -\frac{2}{3\omega + 4} \cdot \frac{1}{t} = -\frac{H}{\omega + 1} . \qquad \text{(III.25)}$$

Thus $|\dot{G}|G|$ is of the order of Hubble's constant unless ω is large and its sign indicates that the gravitational constant should decrease with time (see Figure III.1)

However, for a large enough $|C|$, the ϕ-dominated solutions differ significantly from the matter-dominated ones even at the present epochs. In this case for C large and negative we can have G increasing with time even at relatively recent epochs.

We will review the evidence for or against G-variation later.

III.4 THE HOYLE-NARLIKAR COSMOLOGIES

We next consider another gravitation theory that may claim to have given the most direct quantitative expression to Mach's principle. This theory was first proposed in 1964 by Fred Hoyle and the author, and we will refer to it here as the HN theory and to the cosmological models based on it as HN cosmologies. Throughout this discussion we will set $c = 1$.

Like general relativity and the Brans-Dicke theory, the HN theory is formulated in the Riemannian spacetime. There is one important difference, however, between this theory and all other cosmological theories we have discussed so far. The difference lies in the fact that general relativity, the Brans-Dicke theory, and so on are pure field theories, whereas the HN theory is based on the concept of direct interparticle action. The difference between the two types of theories is best seen in a description of electromagnetism to which we will frequently refer in this section and the next for comparison. Until the advent of Maxwell's field theory, it was customary to describe electrical and magnetic interactions as instances of direct action at a distance between particles. The success of Maxwell's theory established the field concept in physics at the expense of the concept of action at a distance (see Figure 3).

Since Mach's principle (implying as it does a connection between the local and the distant) suggests action at a distance, even an early convert to it like Einstein later became skeptical as to its validity. By the early 1960s, however, it became clear that action at a distance can successfully describe electrodynamics and that it has interesting cosmological implications. Since Hoyle and the author had played an active role in these developments (see ref. [3]), they naturally adopted an action-at-a-distance approach to Mach's principle.

Accordingly, we use here the somewhat unfamiliar notation of action at a distance. Let us denote by a,b... the particles in the universe, m_a, e_a being the mass and charge of the a^{th} particle. As implied by Mach, the mass m_a is not entirely an intrinsic property of particle a: it also owes its origin to the background provided by the rest of the universe. To express this idea quantitatively, write

$$m_a(A) = \lambda_a \sum_{b=a} m^{(b)}(A). \qquad \text{(III.26)}$$

The above expression means the following. At a typical world point A on the world line of particle a, the mass acquired by a is the net sum of contributions from all other particles b ($\neq$ a) in the universe.

The contribution from b at A is given by the scalar function $m^{(b)}$ (A). The coupling constant λ_a is intrinsic to the particle a. Notice, however, that if a were the only particle in the universe $m_a = 0$ and we have the conclusion arrived at in (III.4).

A Digression into Electromagnetic Theory

What are these functions $m^{(b)}$ (X)? That they communicate the property of inertia from particles b to any particle placed at the spacetime point X is clear from the context. To arrive at a suitable form for them we take hints from action-at-a-distance electromagnetism, in which it is usual to introduce electromagnetic disturbances that arise specifically from sources, that is, from moving electrical charges. Accordingly, we introduce the 4-potential $A_i^{(b)}$ (X) as denoting the electromagnetic effect at X from the electric charge b. The $A_i^{(b)}$ (X) satisfies the wave equation

$$\Box A_i^{(b)} + R_i^k A_k^{(b)} = 4\pi J_i^{(b)} \qquad \text{(III.27)}$$

where $J_i^{(b)}$ is the 4-current density generated by the charge b. The solution of (III.27) may be written in the integral form

$$A_i^{(b)} (X) = 4\pi \int e_b G_{ik} (X,B) db^k \qquad \text{(III.28)}$$

where G_{ik} (X,B) is a Green's function of the wave operator $(g_i^k \Box + R_i^k)$.* The well-known Coulomb potential is a special case of (III.28).

The Green's function is not uniquely fixed from the form of the wave operator alone. Boundary conditions must also be specified. The customary boundary condition is that imposed by causality; that is, the influence from B to X must vanish if X lies outside the future light cone of B. The Green's function satisfying this condition is called the retarded Green's function. We will denote such a Green's function with a superscript R. Similarly a Green's function confined to the past light cone of B is called the advanced Green's function and is denoted with a superscript A (see Figure 4).

These Green's functions have played a key role in action-at-a-distance theories. It was originally believed that action at a distance must be instantaneous and hence inconsistent with the framework of special relativity. However, K. Schwarzschild, H.Tetrode, and A.D. Fokker demonstrated during the first three decades of this century that a relativistically consistent action at a distance theory can indeed be formulated. If we consider two spacetime points A and B with s^2_{AB} as the invariant square of the relativistic distance between them, then $\delta(s^2_{AB})$, where δ is the Dirac delta function, is a convenient function for transmitting physical influences between A and B. For this function acts only when A and B are connectible by a light ray (that is, when $s^2_{AB} = 0$). This delta function therefore necessarily occurs as the main component in any Green's function in the action-at-a-distance theory. The action principle, which is the basis of the electromagnetic theory in

Riemannian spacetime, is described below. We start with the action.

$$\mathcal{A} = - \sum_{a} \sum_{< b} 4\pi e_a e_b \int \int \bar{G}_{ik} \, da^i \, db^k \tag{III.29}$$

where $\bar{G}_{ik}$ is the symmetric Green's function given by

$$\bar{G}_{ik} (A,B) \equiv \frac{1}{2} \left[G^R_{ik} (A,B) + G^A_{ik} (A,B) \right] . \tag{III.30}$$

Thus $\bar{G}_{ik}$ (A,B) = $\bar{G}_{ik}$ (A,B) and each term in the action is completely symmetric between each pair of particles. The action (III.29) together with suitable cosmological boundary conditions reproduces all the electromagnetic effects of the standard Maxwell field theory.

That cosmological boundary conditions are necessary in the action-at-a-distance framework is seen from the following simple illustration. Any retarded signal emitted by particle a will get an advanced reaction back from b, as shown in Figure 5. This signal from b arrives at a at the same time that the original signal left a, no matter how far away b is! Thus electromagnetism ceases to be a local theory: any so-called local effect must take account of the response of the universe, which consists of reactions from **all** such particles b other than a. This was pointed out first by J.A. Wheeler and R.P. Feynman in 1945. Later, between 1962 and 1963, J.E. Hogarth, F. Hoyle, and the author showed that this response depends on the model of the universe. A "correct" response eliminates all advanced effects except those present in the radiation reaction.

It is interesting (and significant) that the steady state model discussed in Chapter 9 generates the correct response, while all Friedmann models fail to do so.

Inertia and Gravity

Our purpose in the above digression into electromagnetism was to show that a similar approach to inertia leads us to a Machian theory of gravity. In the case of inertia we note that the functions $m^{(b)}(X)$ are scalars and so we have to deal with scalar Green's functions. Thus we write

$$m^{(b)}(X) = \int \lambda_b \, \widetilde{G}(X,B) \, ds_b \tag{III.31}$$

and the inertial action as

$$\mathcal{A} = -\sum_{a} \sum_{<\, b} \int \int \lambda_a \lambda_b \, \widetilde{G}(A,B) \, ds_a \, ds_b . \tag{III.32}$$

What is $\widetilde{G}(A,B)$? Again we proceed by analogy with electromagnetism.

From symmetry considerations we need $\widetilde{G}(A,B) = \widetilde{G}(B,A)$. Further we require $\widetilde{G}$ to be a Green's function of a scalar wave equation. To fix $\widetilde{G}$ completely we use another hitherto undiscussed property of Maxwell's electromagnetic theory known as conformal invariance.

Conformal Invariance

Let us consider the transformation

$$\bar{g}_{ik} = \Omega^2 g_{ik} \qquad \text{(III.33)}$$

where Ω is a twice-differentiable function of coordinates x^i and lies in the range $0 < \Omega < \infty$. Such a transformation is called conformal transformation. Given a spacetime manifold $\mathcal{M}$ with coordinates (x^i) and metric (g_{ik}), we have through (III.33) generated another spacetime manifold $\bar{\mathcal{M}}$ with the same coordinate system (x^i) but with a different metric $(\Omega^2 g_{ik})$. $\mathcal{M}$ and $\bar{\mathcal{M}}$ are said to be conformal to each other. If $\mathcal{M}$ is flat then $\bar{\mathcal{M}}$ is said to be conformally flat.

If we identify the corresponding points (with the same x^i) in $\mathcal{M}$ and $\bar{\mathcal{M}}$, we will find that, in general, distances between two points are stretched or compressed when we go from $\mathcal{M}$ to $\bar{\mathcal{M}}$. However, the null-cones in both the manifolds are unchanged. This invariance of null cones is distinct from the invariance under coordinate transformations. The coordinate transformations preserve the null directions locally, and they are important in field theories that describe physical interactions locally. The action-at-a-distance theories describe interactions globally and must take account of the global structure of null cones. Hence such theories are expected to preserve their form under conformal transformations as well.

It is easy to verify that the scalar curvature changes under the conformal transformation to

$$\bar{R} = \Omega^{-2} \left\{ R + 6 \frac{\Box \Omega}{\Omega} \right\} \tag{III.34}$$

where $\Box$ is evaluated with respect to the metric (g_{ik}). There are, however, certain quantities that do remain the same under a conformal transformation. These are known as conformally invariant quantities. It is easy to see for example that the action describing Maxwell's field theory is conformally invariant. Consider the changes

$$\bar{A}_i = A_i + \phi_i \quad (\phi = \text{a scalar function})$$

$$\bar{F}_{ik} = F_{ik}, \quad \bar{J} = \Omega^{-4} J_i .$$

These changes leave the form of Maxwell's equations intact.

We now fix the form of $\tilde{G}(A,B)$ by demanding that our inertial action (III.32) is conformally invariant. Since under the transformation (III.33)

$$d\bar{s}_a = \Omega(A) ds_a, \quad d\bar{s}_b = \Omega(B)\, ds_b, \tag{III.35}$$

we must have

$$\overline{\tilde{G}(A,B)} = \Omega(A)^{-1} \; \Omega(B)^{-1} \; \tilde{G}(A,B). \tag{III.36}$$

The only scalar wave operator that permits (III.36) is then

$$\Box + \frac{1}{6} R. \tag{III.37}$$

In other words, $\tilde{G}(X,B)$ satisfies the wave equation

$$\left[\Box_X + \frac{1}{6} R(X)\right] \tilde{G}(X,B) = \left[-g(X)\right]^{-1/2} \delta_4(X,B). \tag{III.38}$$

$\delta_4(X,B)$ is the four-dimensional Dirac delta function, which vanishes unless $X \equiv B$. Thus we have ensured that the action-at-a distance theory given by (III.32) does not change under conformal transformations.

III.5 THE GRAVITATIONAL EQUATIONS OF HN THEORY

The action of HN theory is given by (III.32), and with the help of definitions (III.26) and (III.31) we may write it as

$$\mathcal{A} = -\sum_a \int m_a \, ds_a . \tag{III.39}$$

Written in this form this action appears to have only the inertial term describing free particles. How can such an action yield any gravitational equations?

The answer to this question lies in the fact that the m_a's in (III.39) are not constants but depend on spacetime coordinates as well as on spacetime geometry. For they are defined with the help of Green's functions, which in turn are defined in terms of spacetime geometry. Thus if we make a small variation

$$g_{ik} \to g_{ik} + \delta g_{ik},$$

the wave equation (III.38) will change and so will its solution. Thus we will have

$$\tilde{G}(A,B) \to \tilde{G}(A,B) + \delta\tilde{G}(A,B)$$

and hence $\mathcal{A} \to \mathcal{A} + \delta\mathcal{A}$. We therefore have a nontrivial problem whose solution may be expressed in the following way. To simplify matters we will take all λ_a to be equal to unity. (Later we will relax this assumption.)

Define the following functions:

$$m(X) = \sum_a m^{(a)}(X) = \frac{1}{2}\,[m^R(X) + m^A(X)]. \qquad \text{(III.40)}$$

$$\phi(X) = m^R(X)\, m^A(X), \qquad \text{(III.41)}$$

$$N(X) = \sum_a \int \delta_4(X,A)\,[-g(X)]^{-1/2}\, ds_a. \qquad \text{(III.42)}$$

As in the electromagnetic case, we have chosen the symmetric (half R + half A) Green's function. The gravitational equations then become (with $m_k \equiv m_{,k}$):

$$R_{ik} - \frac{1}{2} g_{ik} R = -\frac{6}{\phi} (T_{ik} - \frac{1}{6} (g_{ik} \Box \phi - \phi_{ik})$$

$$- \frac{1}{2} (m_i^R m_k^A + m_k^R m_i^A - g_{ik} g^{pq} m_p^R m_q^A), \qquad \text{(III.43)}$$

together with the "source" equation for m(X)

$$m + \frac{1}{6} Rm = N. \qquad \text{(III.44)}$$

This derivation leading to the final set of equations of the theory may appear somewhat long-winded to anybody unfamiliar with the techniques of direct interparticle action. We have followed here the method used by Hoyle and the author, who arrived at this theory via their earlier work on electromagnetism. As in the electromagnetic case, the universe responds to a local event. To ensure causality and to eliminate advanced effects, the correct response should be given by

$$\sum_a m^{(a)A} (X) \equiv \sum_a m^{(a)R}(X) = m(X). \qquad \text{(III.45)}$$

Under these conditions the equations (III.43) further simplify to

$$R_{ik} - \frac{1}{2} g_{ik} R = - \frac{6}{m^2} \Big[T_{ik} + \frac{1}{6} (g_{ik} \Box m^2 - m^2_{ik})$$

$$+ \quad m_i m_k - \frac{1}{2} g_{ik} m^{\ell} m_{\ell} \Big] \qquad \text{(III.46)}$$

Had we adopted the standard field theoretical approach and introduced a scalar inertia field $m(X)$, we could have arrived at (III.44) and (III.46) from the action given by

$$\mathcal{A} = \int \left\{ \frac{1}{12} Rm^2 + m^i m_i \right\} \sqrt{-g}\, d^4x - \sum_a \int m ds_a . \qquad \text{(III.47)}$$

The action-at-a-distance approach, although unfamiliar to a typical theoretical physicist, is useful in that it gives direct expression to Mach's principle. The physical interpretation of the field theoretical term (III.47) is not so easy to see. For this reason, we have discussed the former approach at some length.

Notice that in the former approach our action (III.39) contained only the last term of (III.47), but there m was made up of nonlocal two-point functions. Here m is a straightforward field with sources in matter whose dynamic properties are defined through the first term in the above action.

Since the property of conformal invariance was used in the formulation of the theory, we expect the final equations (III.44) and (III.46) to exhibit conformal invariance. This expectation is borne out. If (g_{ik}, m) are a solution of these equations, then so are

$$\bar{g}_{ik} = \Omega^2 g_{ik}, \quad \bar{m} = \Omega^{-1} m. \tag{III.48}$$

Thus apart from coordinate invariance of general relativity, this theory also shows conformal invariance.

It is well known that the coordinate invariance of the action leads to a conservation law for the energy momentum tensor. In this case the conformal invariance of the action leads to a vanishing of trace of the field equations. It may be easily verified that the trace of (III.46) vanishes in view of (III.44). The vanishing of trace represents the fact that the problem is underdetermined. Just as the vanishing of $T^{ik}_{;k}$ in general relativity shows that more solutions can be generated from any given solution by coordinate transformations, so we can generate more solutions through (III.48). All these solutions are physically equivalent provided we stick to the rule that Ω does not vanish or become infinite.

Suppose we are allowed to choose an Ω in the above range that ensures that

$$\bar{m} = \Omega^{-1} m = \text{constant} = m_a. \tag{III.49}$$

This choice of Ω is possible provided m does not vanish or become infinite. This conformal frame is called the Einstein frame, for from (III.49) we get a simplified form for (III.46).

$$R_{ik} - \frac{1}{2} g_{ik} R = - \kappa T_{ik}, \tag{III.50}$$

with the constant κ given by

$$\kappa = \frac{6}{m_o^2} \ . \tag{III.51}$$

Thus we have arrived at Einstein's equations! At first sight we don't seemed to have gained anything. We have no new theory and hence no new predictions, as in the Brans-Dicke theory. Closer examination, however, reveals several ways in which this theory goes beyond relativity.

1. Our starting point was based on Mach's principle. It is only in the many particle approximation, when the response condition (III.45) is satisfied, that we arrive at the final Einstein like field equations. An empty universe in relativity is given by

$$R_{ik} = 0.$$

which can have well-defined spacetimes as solutions. Test particles in such spacetimes will have well-defined trajectories. Such trajectories would not make any sense according to Mach, since we no longer have a material background against which to measure the motion of these particles. These solutions in fact correspond to the $\underline{f} = 0$ solutions of (III.3). In the HN theory an empty universe corresponds to

$$m = 0, \quad \text{indeterminate } g_{ik},$$

in accord with the Machian $m = 0$ solution of (III.3).

2. The sign of κ is fixed arbitrarily in general relativity. Neither in the heuristic derivation of Einstein nor in the Hilbert action principle is κ required to be positive. It is only when κ is determined by reference to Newtonian gravity in the weak field approximation that we conclude that $\kappa > 0$. In the HN theory *shows that κ must necessarily be positive.* (This conclusion does not depend on our assumption of $\lambda_a = 1$; the result follows whatever sign the λ_a are given.)

3. In the direct interparticle approach it is not possible to accommodate the λ-term of cosmic repulsion. Thus Occam's razor automatically comes into play. In relativity the λ-term is still possible.

4. The transition from (III.46) to (III.50) is possible provided $0 < \Omega < \infty$. What happens if we break this rule? Suppose in the solution of (III.46) we had a hypersurface on which $m = 0$. If we insist on the transformation (III.49) in a region that contains such a hypersurface, we have to pay the price of $\Omega \to 0$, which in turn produces spacetime singularities. The work of A.K. Kembhavi in 1979 showed that the well-known cases of spacetime singularities of relativity arise because of the occurrence of zero mass hypersurfaces

in the solution of the equations (III.46). For a simple example of this conclusion let us look at the standard big bang singularity of relativity.

Consider the Minkowski line element (with $c = 1$)

$$ds^2 = d\tau^2 - dx^2 - dy^2 - dz^2 \tag{III.52}$$

as a solution of (III.46). It is easily verified that the mass function satisfying both (III.44) and (III.46) for a uniform number density N of particles is

$$m = \tau^2 \, . \tag{III.53}$$

This is the simplest possible cosmological solution in this theory.

If we now insist on going over to a frame with constant mass $\bar{m}$, then from (III.48) we see that the appropriate Ω must be given by

$$\Omega = \tau^2 . \tag{III.54}$$

However, Ω vanishes on the hypersurface $m = 0$. The transformation to the Einstein conformal frame is therefore "illegal." The price paid for insisting that $\bar{m}$ = constant is that the resulting model has a geometrical singularity at $\tau = 0$. In fact it is easily verified that the new model is none other than the singular Einstein-de Sitter model. (Make the time transformation $\tau \propto t^{1/3}$ to demonstrate this result explicitly.)

5. It is instructive to see how the phenomenon of Hubble redshift is explained in the flat spacetime model of (III.52) and (III.53). Clearly, a light photon traveling in Minkowski spacetime does not undergo redshift. Consider, however, what happens to a light photon arriving at the observer at the present epoch τ_0 from a galaxy at a distance r. This photon originated in an atomic (or molecular) transition at time $\tau_0 - r$.

From atomic physics, the wavelength of a photon so transmitted varies inversely as the mass of the electron (making the atomic transition). From (III.53) we see that if λ is the wavelength of this photon and λ_0 is the wavelength of a photon emitted in a similar transition at τ_0 at the observer, then

$$1 + z \equiv \frac{\lambda}{\lambda_0} = \frac{m(\tau_0)}{m(\tau_0 - r)} = \frac{\tau_0^2}{(\tau_0 - r)^2} \quad . \qquad \text{(III.55)}$$

Thus the redshift in the above HN cosmology arises from the variation of particle masses.

6. A variable gravitational constant arises in the HN cosmologies if we relax the assumption that λ_a are constants. If λ_a change with time it is possible to generate cosmological models in which G changes with time. We will not discuss such models in detail. The result may be stated in the form

$$\frac{\dot{G}}{G} = - \alpha H, \qquad \text{(III.56)}$$

where H is the Hubble constant of the epoch of measurement and α is a constant of order unity.

It was shown by Hoyle and the author in 1972 that λ_a increasing with time may be interpreted as creation of new particles in the universe. They did not give a dynamic theory of matter creation (like the C-field theory), but instead fixed the time dependence of λ_a by an appeal to the Large Numbers Hypothesis. Since this hypothesis is discussed next, we will postpone further discussion of the HN cosmology.

IV THE LARGE NUMBERS HYPOTHESIS

IV.1 THE LARGE DIMENSIONLESS NUMBERS OF PHYSICS

Physics is riddled with units of various kinds and with experimentally determined quantities of various magnitudes. From this vast collection certain constants emerge as having special significance in the framing of basic physical laws; for example, the constant of gravitation G, the charge of the electron e, and so on. The numbers expressing the magnitudes of G, e, and so on depend on the units used. For example

$$e = 4.80325 \times 10^{-10} \text{ electrostatic units}$$
$$= 1.60207 \times 10^{-20} \text{ electromagnetic units.}$$

Clearly these numbers by themselves cannot have absolute significance.

However, certain combinations of these physical constants have no units at all. For example, the combination of $\hbar$, c, and e

$$\frac{\hbar c}{e^2} = 137.03602 \qquad \text{(IV.1)}$$

does not depend on the units used. It must therefore express some physical fact of absolute significance. Indeed, its reciprocal $e^2/\hbar c$, known commonly as the fine structure constant, expresses the strength of the electromagnetic interaction, which we believe to be an intrinsic property of nature. A future more complete theory may well give a reason why this constant has this particular value.

Given e, G, and the masses of proton and the electron m_p and m_e, we can construct another dimensionless constant (that is, a constant with no units):

$$\frac{e^2}{Gm_p m_e} = 2.3 \times 10^{39} \sim 10^{40}. \qquad \text{(IV.2)}$$

This constant measures the relative strength of the electrical and the gravitational forces between the electron and the proton. Like (IV.1) this constant reflects an intrinsic property of nature. However, unlike (IV.1), the constant in (IV.2) is enormously large! Why such a large number?

Perhaps the appearance of a large dimensionless constant might be dismissed as some quirk on the part of nature. The mystery deepens, however, if we consider another dimensionless number. This is the ratio of the length scale associated with the universe, c/H_o, and the length associated with the electron, $e^2/m_e c^2$. This ratio is

$$\frac{m_e c^3}{e^2 H_o} = 3.7 \times 10^{40} h_o^{-1} \sim 10^{40}. \qquad \text{(IV.3)}$$

Not only do we have another large dimensionless number in (IV.3), but it is of the same order as in (IV.2).

We can generate another large number of special significance out of particle physics and cosmology. Assuming the closure density

ρ , Let us calculate the number of particles in a Euclidean sphere of radius c/H_o, the mass of each particle being m_p. The answer is

$$\mathcal{N} = \frac{4\pi}{3m_p} \left(\frac{c}{H_o}\right)^3 \cdot \frac{3H_o^2}{8\pi G} = \frac{c^3}{2m_p GH_o}$$

$$= 4 \times 10^{79}\, h_o^{-1}$$

$$\sim 10^{80} \quad . \tag{IV.4}$$

Thus taking $\mathcal{N}$ as a standard we see that the large dimensionless numbers of (IV.2) and (IV.3) are both of the order $\mathcal{N}^{1/2}$.

Reactions among physicists have varied as to the significance of all these numbers. Some dismiss it as a coincidence with the rejoinder: So what? Others have read deep significance in these relationships. The latter class includes such distinguished physicists as A.S. Eddington and P.A.M. Dirac.

Dirac pointed out in 1937 that the relationships (IV.3) and (IV.4) contain the Hubble constant H_o, and therefore the magnitudes computed in these formulae vary with the epoch in the standard Friedmann model. If so, the near equality of (IV.2) and (IV.3) has to be a coincidence of the present epoch in the universe, unless the constant in (IV.2) also varies in such a way as to maintain the state of near equality with (IV.3) at all epochs. With this proviso, the equality of (IV.2) and (IV.3) is not coincidental but is characteristic of the universe at all epochs. This proviso also implies that at least one of the so-called constants involved in (IV.2), e, m_p, m_e and G, must vary with the epoch.

This proviso has been generalized by Dirac to what he calls the Large Numbers Hypothesis (LNH). To understand this hypothesis we rephrase the ratio (IV.3) as that between the time scale associated with the universe, $T_o = H_o^{-1}$, and the time taken by light to travel a distance of the order of the classical electron radius, $t_e = e^2/m_e c^3$. The LNH then states that any large number that at the present epoch is of the order

$$\left\{\frac{T_o}{t_e}\right\}^k ,$$

where k is of order unity, varies with the epoch t as $(t/t_e)^k$, with a constant of proportionality of order unity.

Applied to (IV.2), therefore, the LNH implies that the ratio $e^2/Gm_p m_e$ must vary as $(t/t_e)^{-1}$. Dirac made the distinction between e, m_e, m_p on one side and G on the other in the sense that the former are atomic (small-scale quantities) which G has macroscopic significance. In the Machian cosmologies G was in fact related to the large-scale structure of the universe. Dirac therefore assumed that if we use "atomic units" that always maintain fixed values for atomic quantities, then t, will be constant and $G \propto t^{-1}$. That is, in terms of atomic time units the gravitational constant must vary with the epoch t, with $|\dot{G}/G| \sim H$.

We will now explore the implications of LNH for cosmology.

IV.2 THE TWO METRICS

Clearly the variation of G predicted by the LNH goes against Einstein's theory of gravitation, which demands a constant G. As in the Brans-Dicke theory of Chapter 10, we are forced to modify the relativistic framework to accommodate a varying G. Dirac approached this problem in the following way.

First he took note of the many Solar System tests that are in favor of general relativity and argued that the theory should not be abandoned altogether. Instead, Dirac proposed two scales of measurement, one holding in atomic physics and the other in gravitation physics. If we choose the atomic system, we will be able to describe atomic physics in the usual way, that is, with constant values for the atomic constants like e, $\hbar$, m_e, m_p, and so on. However, in this system G will be variable, since Dirac considers it a constant belonging to gravitation physics. If on the other hand we use gravitational units, then according to Dirac G will be constant and atomic quantities will be found variable. And in these latter units the gravitational phenomena can be described by the Einstein equations.

These two units can be specified in Dirac's framework by having two different spacetime metrics. We will denote these by ds_A^2 and ds_E^2 respectively for the atomic and the gravitational systems (the subscript E in the latter case committing us to Einstein's equations of gravity). We will use these subscripts in general on any physical quantity to indicate what system of measurement is being used. Thus according to Dirac

$$G_E, \quad (m_e)_A, \quad (m_p)_A$$

are constants, while

$$G_A, \quad (m_e)_E, \quad (m_p)_E$$

are variable.

Returning to the astronomical tests of general relativity, we note that the mass of the gravitating body (for example, the Sun) occurs in the Schwarzschild solution. Clearly this mass, which is the gravitational mass, must be a constant in the gravitational units. We denote this mass by M_E. Any measurements made on the Earth, however, use atomic systems (such as spectrometers and atomic clocks), and before we interpret any experimental result we must make sure that all observable *quantities are transformed* to atomic units.

This argument tells us how necessary it is to know the ratio

$$\beta = \frac{ds_E}{ds_A} \tag{IV.5}$$

and how the transformation is to be made of any physical quantity from one system of units to another. Here we need a quantitative theory to guide us, a theory that goes further than the above qualitative arguments have so far taken us. We will come to this problem soon.

We also note another outcome of our Solar System example. If we assume that our astronomical body has N_E nucleons, each of mass m_E, then we may write

$$M_E = m_E N_E = m_E N \qquad \text{(IV.6)}$$

where we have dropped the suffix E on N because it is a pure number. Whatever metric we use, we will count the same number of particles in the gravitating body. In (IV.6) we have M_E = constant, $m_E \neq$ constant, since the latter is an atomic quantity. Thus $N \neq$ constant. In other words, we are forced to conclude that the number of nucleons in the body must change with time. Again we need a quantitative theory to tell us how N changes; but creation (or destruction) of nucleons in a macroscopic object is demanded by Dirac's argument.

So far we have not used the LNH, which started us on the two-metric theory. Let us now see how it helps us in deciding how the nonconservation of nucleon number in the body is regulated.

The Creation of Particles

If we go back to (IV.4) and apply the LNH to $\mathcal{N}$, we easily find that $k = 2$, that is,

$$\mathcal{N}(t) \sim \left(\frac{t}{t_e}\right)^2 \propto t^2. \qquad \text{(IV.7)}$$

In other words, the number of particles in the universe in the sense defined in §IV.1 increases with t. Dirac has taken this result to imply that particles are being continually created in the universe.

The creation can occur, according to Dirac, in two possible ways. In additive creation the particles are created uniformly throughout space, while in multiplicative creation the new particles occur preferentially where matter already exists. Thus in the former mode creation occurs mostly in intergalactic space, while in the latter mode creation occurs mostly in the vicinity of existing astronomical objects.

Using these ideas we return to (IV.6). In additive creation the astronomical body will not acquire any significant number of new particles and thus N = constant, giving

$$m_E = \text{constant (additive creation)} \qquad \text{(IV.8)}$$

In multiplicative creation N must increase as t^2 and hence

$$m_E \propto t^{-2} \text{ (multiplicative creation).} \qquad \text{(IV.9)}$$

The Determination of β

The connection between ds_A and ds_E can be fixed by considering the motion of a planet (such as the Earth) around a star (the Sun). The dynamic equation in the Newtonian approximation is

$$GM = v^2 r \qquad \text{(IV.10)}$$

where M = mass of the star, v = speed of the planet, and r = radius of the orbit. The above relation is expected to hold in either of the two systems of units, since GM/v^2r is a dimensionless quantity. Also, with $c = 1$ the speed v is dimensionless. Thus v = constant in either units. Next, in gravitational units M_E = constant, G_E = constant, hence r_E = constant.

If (IV.2) is used with atomic units, we have

$$G_A \sim t^{-1}. \tag{IV.11}$$

Also, in multiplicative creation $M_A = t^2$ while for additive creation M_A = constant. Hence in these units

$$r_A \sim t \quad \text{(multiplicative creation)}, \tag{IV.12}$$

$$r_A \sim t^{-1} \quad \text{(additive creation)}. \tag{IV.13}$$

thus we have

$$\frac{r_A}{r_E} \sim t \quad \text{(multiplicative creation)} \tag{IV.14}$$

and

$$\frac{r_A}{r_E} \sim t^{-1} \quad \text{(additive creation)}. \tag{IV.15}$$

In other words, measured in atomic units, the distance of the planet from the star increases with t if the universe has multiplicative creation of matter, and the distance decreases with t (as t^{-1}) for additive creation. We will consider in Chapter 13 the observable consequences of such an assertion.

From (IV.14) and (IV.15) we get the behavior of β defined in (IV.5). This ratio of ds_E to ds_A behaves as t^{-1} or t, depending on whether we have multiplicative creation or additive creation in the universe.

IV.3 COSMOLOGICAL MODELS

Using the LNH Dirac constructed cosmological models in both the circumstances discussed above, namely for multiplicative and additive creation. As in the case of standard cosmologies, the assumptions of homogeneity and isotropy lead to the Robertson-Walker line element in atomic units:

$$ds_A^2 = c^2 dt^2 - S^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2). \right] \quad \text{(IV.16)}$$

How does the LNH determine k and $S(t)$? We reproduce below the argument given by Dirac.

First we note that the metric proper distance at time t between a galaxy G at $r = 0$ and a galaxy at $r = r_1$ is given by

$$d \equiv S(t) \int_0^{r_1} \frac{dr}{\sqrt{1 - kr^2}} \equiv S(t) \, f \, (r_1). \quad \text{(IV.17)}$$

According to LNH, for large t (that is, for $t \gg t_e$) the expression for $S(t)$ should be $\sim (t/t_e)^n$ or $\ln(t/t_e)$. The (metric) recession velocity corresponding to (IV.17) will therefore be given by

$$\dot{d} \sim n t_e^{-n} f(r_1) t^{n-1} \quad \text{or} \quad \dot{d} \sim t^{-1} f(r_1). \qquad \text{(IV.18)}$$

The constants multiplying $(t/t_e)^n$ or $\ln(t/t_e)$ in $S(t)$ must be on the order of unity, and hence the constants implied in the $(\sim)$ relation above are also on the order of unity. It is then easy to verify that except for $n = 1$, there exists an epoch either in the past (for $n < 1$ or for $S \sim \ln t$) or in the future (for $n > 1$) when $\dot{d} = c$ for any galaxy with $r_1 > 0$. For example, for $n = 1/2$ we find that for a galaxy that at present has $\dot{d} \sim 10^{-3}c$, the condition $\dot{d} = c$ occurred in the past epoch given by

$$t_p = \left\{ \frac{T_o}{t_e} \right\} \cdot 10^{-6}\, t_e \sim 10^{34}\, t_e.$$

That is, t_p/t_e is a large number. However, by the LNH, t_p is a constant epoch when a significant event took place for galaxy G_1, its recession speed became equal to c. Hence such a constant epoch should not generate a large number. Therefore only the case

$$S(t) \sim (t/t_e) \qquad \text{(IV.19)}$$

is permitted by the LNH.

The arguments given above could be criticized on the following grounds. The epoch when $\dot{d} = c$ is not unique to the model as a whole; it depends on $f(r_1)$ and hence on the galaxy chosen. So it is not necessary that LNH should apply to this epoch. Nor is it clear why $\dot{d} = c$ should be considered significant. Nothing special happens to the galaxy in question when its metric velocity of recession becomes equal to c for the observer at $r = 0$. No global property like the event horizon or the particle horizon enters the argument.

Nevertheless if we follow the argument further, then we can write our cosmological line element as

$$ds_A^2 = c^2\, dt^2 - (At)^2 \left[\frac{dr^2}{1 - kr^2} + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2)\right], \quad \text{(IV.20)}$$

where A is a constant. We next consider multiplicative creation. Since in this case from §IV.2

$$ds_E = t^{-1}\, ds_A, \quad \text{(IV.21)}$$

it is easy to see that a transformation

$$dt_E = \frac{dt}{t} \quad \text{(IV.22)}$$

gives us

$$ds_E^2 = c^2\, dt_E^2 - A^2 \left[\frac{dr^2}{1 - kr^2} + r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2)\right]. \quad \text{(IV.23)}$$

Now we recall that the above line element must be a solution of Einstein's equations. In fact Einstein did obtain such a static solution for homogeneous and isotropic dust with the use of the λ-term.

Thus provided we admit the λ-term, the spacetime of (IV.23) is none other than the spacetime of the Einstein universe with $k = +1$. With a suitable scaling of the r-coordinate we can express (IV.23) in the standard form. Notice, however, that unlike the Einstein universe this Dirac universe does show the phenomenon of redshift of galaxies. For redshift measurements involve comparisons of the rates at which atomic clocks run at the emitting and receiving galaxies: and for such comparisons the line element (IV.20) instead of (IV.23) must be used.

For additive creation the situation is more complicated. In the multiplicative creation case the gravitational mass of an astronomical object was held constant in the gravitational units in spite of creation of new particles, by letting the particle masses decrease with time. In the additive creation case the particle masses remain constant even though their number increases [see (IV.8)]. Dirac was therefore faced with apparent nonconservation of energy. To conserve energy Dirac proposed that along with positive mass particles an equal number of negative mass particles is also created. The negative mass distribution is homogeneous and remains undertectable by standard astronomical observations. In a completely homogeneous situation the positive and negative mass distributions compensate gravitaionally to produce flat Minkowski spacetime. The formation of stars and galaxies by the accumulation of positive mass particles in the actual universe is a result of small departures from this completely homogeneous situation.

It is worth pointing out that when Dirac first proposed a cosmological model based on the LNH between 1937 and 1938 he assumed no matter creation. In this model the number of particles per unit coordinate volume is constant, as in standard cosmologies. Hence the number of particles per unit proper volume goes as S^{-3}, and since the proper volume of the universe goes as $(c/H)^3$, the number of particles in the universe denoted earlier by goes as

$$S^{-3}\left(\frac{c}{H}\right)^3 \propto (\dot{S})^{-3}.$$

However, by LNH we know that

$$\mathcal{N} \propto t^2. \qquad \text{(IV.24)}$$

Therefore we have

$$t^2\,\dot{S}^3 = \text{constant},$$

that is,

$$S \propto t^{1/3}. \qquad \text{(IV.25)}$$

Thus for no particle creation S increases much more slowly with t. (Of course, this solution is ruled out if we apply the LNH to the function S, as we did in the beginning of this section.)

IV.4 THE SCALE COVARIANT THEORY OF GRAVITY

In 1977 V.M. Canuto and S.H. Hsieh proposed a field theory to incorporate Dirac cosmology. As we have seen in the case of the steady state theory, the C-field theory gave a field-theoretic description of creation of matter that was demanded by the deductive approach of the perfect cosmological principle. In the same way this theory attempts to provide a quantitative framework for the deductions of LNH.

It is clear from the preceding discussion that the crucial function needed to quantify Dirac's arguments is the function β relating the two metrics ds_A and ds_E. How do we determine β? In the theory proposed by Canuto and Hsieh, called by these authors the Scale Covariant Theory of Gravity (SCTG), the first step is to note that a physicist making measurements on the Earth uses the atomic metric rather than the gravitational metric. How would the gravitational equations look in the atomic metric? The answer to this question is straightforward when we note that the two metrics are conformal to each other.

For convenience let us drop the suffix A on the atomic metric and on all other quantities measured with this metric. Write therefore

$$ds^2 = g_{ik}\, dx^i\, dx^k = \beta^{-2}\, g_{ikE}\, dx^i\, dx^k. \qquad \text{(IV.26)}$$

Then, since we know that the metric tensor g_{ikE} satisfies Einstein's equations, we can work out the equations satisfied by the atomic metric tensor g_{ik}. A straightforward manipulation gives the answer as

$$R_{ik} - \frac{1}{2} g_{ik} R = - \kappa T_{ik} - \frac{2}{\beta} \beta_{ik} + \frac{4}{\beta^2} \beta_{,i}\beta_{,k}$$

$$+ \frac{1}{\beta^2} g_{ik} (2\beta \Box \beta - \beta^{,l} \beta_{,l}). \qquad \text{(IV.27)}$$

Here everything is expressed in terms of the metric tensor g_{ik}. The energy momentum tensor T_{ikE} of the rest of the physics in the gravitational metric becomes T_{ik} in the atomic metric. So long as scale covariance (that is, conformal invariance) is maintained, we have definite prescriptions for obtaining T_{ik}. Thus if T_{ikF} is the perfect fluid energy tensor derivable from the action

$$\left(- \sum_a c \int m_a \, ds_a \right)_E \qquad \text{(IV.28)}$$

then scale covariance of the theory demands that to preserve (IV.28) in going from $ds_E \to ds$ we must have masses transforming as

$$m_E \to m = m_E \beta \qquad \text{(IV.29)}$$

Notice, however, that if the masses m_a in (IV.28) are constant, the SCTG gives nothing different from general relativity!

For the particle masses m_a must vary in the gravitational framework (multiplicative creation) and stay fixed in the atomic framework. Clearly then we need some prescription to tell us how the m_a vary in the former framework. Or, alternatively, the standard form (IV.28) must be replaced by something different involving β, in such a way that the result m_a = constant emerges in the atomic framework. Although the SCTG does not specify how this is done, we will assume henceforth that this is the case. (This is one of the two aspects in which the SCTG is not yet complete. The second aspect is discussed shortly.)

What is $\dot{\beta}$? If we take the trace of the equations (IV.27) we get a wave equation for β:

$$\beta + \frac{1}{6} R\beta = \frac{\kappa}{6} T. \qquad \text{(IV.30)}$$

However, this wave equation is not independent of the field equations and therefore does not supply any new information. The situation here is similar to that for the Hoyle-Narlikar theory of the last chapter. Indeed, the unknown β can be eliminated by transforming back to the gravitational metric.

Clearly, to say what β is, that is, how the atomic units are related to gravitational units, a new and independent source equation for β is needed. Canuto and Hsieh have so far not been able to suggest how this equation is to be obtained. As they have pointed out, the scale covariance must be broken by such an equation, hence until this symmetry-breaking term is introduced the SCTG remains incomplete.

So in the absence of a true equation for β we have to fall back on the LNH to determine β. It can be shown, for example, that G is a definite function of β corresponding to the cases of no creation and multiplicative creation:

$$G\beta = \text{constant (no creation)} \qquad \text{(IV.31)}$$

$$G\beta^{-1} = \text{constant (multiplicative creation)} \qquad \text{(IV.32)}$$

Canuto and Hsieh have called β the gauge function, and (IV.31) and (IV.32) are examples of fixing specific gauges. A complete theory should be able to tell whether either of the two gauges (or some other gauge) is the correct one. It is also worth noting that in this framework the variation of the gravitational constant depends on the variation of β and the gauge chosen. Thus if experiments give $\dot{\beta}/\beta > 0$ then (IV.31) gives $\dot{G}/G < 0$ while (IV.32) gives $\dot{G}/G > 0$.

IV.5 HN COSMOLOGY REVISITED

Some of the ideas of Dirac and of Canuto and Hsieh are found in a version of HN cosmology proposed by its authors between 1971 and 1972. In our earlier discussion of the HN cosmology we considered the case where $\lambda_a, \lambda_b\ldots$, the constants that denote the strength of the inertial interaction, are true constants. If, however, these constants vary with time, new cosmological models emerge. In these models the following properties hold: (1) there is particle creation at all epochs in such a way that the LNH is satisfied, (2) in atomic units G varies, while (3) in the gravitational units G is constant and particle masses vary. Thus this model is like the multiplicative creation model later proposed by Dirac, although its motivation and quantitative details are different. We briefly illustrate how this model works.

Consider a homogeneous and isotropic Minkowski universe given by

$$ds_M^2 = d\tau^2 - dr^2 - r^2\,(d\theta^2 + \sin^2\theta\, d\phi^2), \qquad \text{(IV.33)}$$

where we have put $c = 1$ for convenience. Let $n(\tau)$ be the particle number density and $\lambda(\tau)$ the time-varying inertial coupling constant of (IV.31). The functions $n(\tau)$ and $\lambda(\tau)$ vary in such a way as to compensate each other's effect; that is, to maintain

$$\lambda n = \text{constant}. \qquad \text{(IV.34)}$$

Thus the mass function $m(\tau)$ is the same as if we had a universe of uniform particle number density n = constant and fixed λ. As in (IV.53) we then get

$$m(\tau) \doteq \tau^2. \qquad \text{(IV.35)}$$

Since Gm^2 = constant, we get the gravitational constant in the Minkowski framework as

$$G_M = \tau^{-4}. \qquad \text{(IV.36)}$$

The mass of a typical particle is not, however, $m(\tau)$ but $\lambda m(\tau)$. To determine it we need to know $\lambda(\tau)$. Hoyle and the author determined $\lambda(\tau)$ from a requirement that the universe is opaque to electromagnetic radiation along the future light cone. A thorough discussion of this issue will take us to the absorber theory of radiation, which lies beyond the scope of this book. We simply quote the result.

This requirement fixes $\lambda(\tau) = \tau^{-1}$ and $n(\tau) = \tau$. It is then verified that the LNH is incorporated by the fact that the dimensionless number

$$\lambda^2 \, (\tau^3 \, n)^{1/2} = \text{constant} = O(1). \tag{IV.37}$$

A conformal transformation

$$ds_E = \Omega_E \, ds_M, \quad \Omega_E = \tau^2 \tag{IV.38}$$

then takes us to the gravitational framework in which G_E = constant. Also, the gravitational mass of an astronomical body remains constant. Thus as in Dirac's multiplicative creation theory, the local solar system tests give the same answer as in relativity.

To transform to atomic framework we need another conformal transformation:

$$ds_A = \Omega_A \, ds_M, \quad \Omega_A = \tau. \tag{IV.39}$$

By writing $t = \tau^2$ the line element now becomes

$$ds_A^2 = d\tau^2 - 2H_o t \, [dr^2 + r^2 \, (d\theta^2 + \sin^2\theta \, d\phi^2)]. \qquad \text{(IV.40)}$$

In this framework the gravitational constant varies as

$$G_A = t^{-1}. \qquad \text{(IV.41)}$$

There is therefore considerable similarity between this theory and the model proposed by Dirac a few years later.

IV.6 CONCLUSION

This brings us to the end of our excursion through some of the better-known parts of nonstandard cosmology. Our survey is by no means exhaustive. We have not discussed such important models as the matter-antimatter symmetric cosmology of Alfven and Klein, the Einstein-Cartan cosmologies, or Milne's kinematic relativity; nor have we discussed such unusual ideas as Segal's chronometric cosmology or McCrea's notion of cosmological uncertainty.

Our purpose was to summarize a few nonstandard cosmologies in order to show that "respectable" cosmology has not been confined to Friedmann models only. Indeed, if the history of science (and astronomy in particular) is anything to go by, it would be premature to conclude that the problem of the universe has been basically solved. To what extent do the theoretical ideas (standard or non-standard) presented so far in this book stand up to observations available today? We briefly review this question next.

V OBSERVATIONAL TESTS

Cosmological tests are of two kinds: (1) those which look at the remote parts of the universe and (2) those which examine nearly regions. Because light (our mode of observation) travels with a finite speed, the tests of the first kind tell no about what the universe was like long time ago. A comparison of those conditions with what exist now can tell us how (if at all) the universe has changed over the cosmological time scales. In principle such tests can be well focussed on predictions of specific models. In practice, however, observational uncertainties predominate over the predicted differences in specific cosmologies. Tests of the second kind are less prone to observational uncertainties. Their value lies in telling us what relics exist today that carry signatures of past events in the universe. However, the uncertainty here comes from the specific scenarios that links those relics with the early epochs.

These shortcomings should be taken into account while assessing the performance of any cosmological theory vis-a-vis observations.

V.1 OBSERVATIONS OF THE DISTANT PARTS OF THE UNIVERSE

The Redshift Magnitude Relation

The steady state cosmology requires $q_o = -1$ while standard cosmologies generally predict q_o in the range o to 1. The measurement of the Hubble relation down to large redshifts ($z > 0.2$)

and faint magnitudes should in principle distinguish between the predictions of different models. Figure 6 illustrate how the (m-z) curves begin to diverge from one another for different values of q_o as z increases beyond 0.2.

Unfortunately the numerous observational uncertainties tend to smear out these differences. These include the aperture correction, the K-correction, the Scott effect, the Malmquist bias, intergalactic absorption and possible luminosity evolution (not, of course, in the steady state model). Details of these may be found in literature cited at the end. Figure 6 illustrates the scatter of actual points around the predicted curves.

The nature of K-correction was not understood in the early 1950s and this had resulted in the erroneous claim by Stebbins and Whitford that the steady state theory was disproved. Likewise, the effect of G-variation on galactic magnitudes in the Hoyle-Narlikar cosmology was incorrectly estimated by Barnothy and Tinsley in 1972. At present assessment both the steady state cosmology and the H-N cosmology are consistent with the data on m-z relation.

Radio Source Counts

The counting of radio sources is a potential test of space-time geometry. In a complete survey the radio astronomer counts all sources brighter than a specified flux density S. For details of this test see reference [4] at the end.

If quasars are at cosmological distances then counts of quasars indicate evolution and disprove the steady state theory. If, however, the quasars are not so far away, their counts are not of any value in deciding cosmology.

For radio galaxies also strong evolution was claimed, most recently by Peacock in 1985. However, after a reanalysis of the data taking into account redshifts of all sources, P. DasGupta, G. Burbidge and the author have claimed that the hypothesis of 'no evolution' cannot be ruled out.

So far as G-varying H-N cosmology or the Dirac cosmology are concerned the source count does not rule them out.

Angular sizes

Here again some studies have claimed evidence for evolution (see reference [4] at the end) but the uncertainty of the data together with a suitable luminosity-size anticorrelation have been claimed by DasGupta and the author to 'save' the steady state theory.

V.2 OBSERVATIONS OF LOCAL SPACE TIME

The Microwave Background

It is usually assumed that the microwave background provides a strong evidence for the big bang. The standard model does, however run into difficulty in explaining the extraordinary smoothness of the background on the small scale. So far no satisfactory theory of galaxy/cluster/supercluster formation has been able to account for this effect. It is tempting to suggest that dust grain thermalization as required in the steady state theory took place after these discrete structures formed.

The horizon problem of the standard model and its implication for anisotropy of the microwave background have led to the various inflationary scenarios. These solve one set of problems but raise others! It is too early to pass judgement on this approach. For a detailed discussion of these ideas see reference [5].

The Age of the Universe

The standard big bang cosmology runs into severe 'age problem' if $h_o = 1$ and marginal age problem if $h_o = 1/2$. This is because the total age of the standard model lies between $\sim 9h_o^{-1/2}$ billion years for $q_o = 0$ to $\sim 5h_o^{-1/2}$ billion years for $q_o = 1$. Compared to the ages of globular clusters (> 12 - 18 b.y), this age is not quite adequate.

The problem is not so severe for the steady state model but continues to be so for other nonstandard models discussed here. For a fuller discussion see references at the end.

The Abundance of Light Nuclei

The big bang standard models run into the over-abundance problem of ^4He, especially if the number of neutrino species turns out to be 3 or more. There also may be a problem of getting both deuterium and ^4He right in the same standard model. We have already discussed the Brans-Dicke cosmology which also tends to run into the same problem.

In steady state model ^4He and ^2H must be produced locally in supermassive stars. Here the details work out reasonably well for ^4He but a stellar scenario for ^2H seems hard to achieve.

The variation of G

As yet no laboratory experiment is sensitive enough to measure $\dot{G}/G$ of the order of 10^{-11}/yr. Such evidence as is available comes from radar measurements of planetary distances and from anomalies in the motions of moon and the planets. Thus Hellings et al have argued that range measurements to Viking-landers and Mariner spacecrafts around mass rule out $\dot{G}/G$ of the above order that is required by most cosmological theories. Van Flandern on the other hand claims that $\dot{G}/G \sim (-6.9 \pm 2.4) \times 10^{-11}\ yr^{-1}$ from the data of Moon's motion. In the case of Moon's motion tidal forces also contribute a term of the same order as the $\dot{G}/G$ term.

Thus the issue seems somewhat uncertain.

V. 3 CONCLUSIONS

The present situation in cosmology may be dignosed under the following sentence: 'The sophistication of theories is not matched by the accuracy of observations'. This sentence neither reflects favourably on theories nor adversely on observations. In standard cosmology ideas from particle physics have brought in numerous speculative scenarios for the early universe. These have yet to pay dividend in terms of explaining any of the observed relics like galaxies, photon to baryon ratio, the nature of matter etc. or in terms of predicting any relics that can be (and have been) observed. Although sophisticated electronics has made extra-galactic observations much more precise compared to the days of Hubble and Eddington, we are still far from appreciating the nature of errors which can easily vitiate any cosmological deduction.

In such a situation nonstandard cosmologies have a useful role to play in offering alternatives that can stimulate the observers into devising discriminatory tests, as the steady state theory did. In any case cosmology can remain a healthy science only until dissenting alternatives to the standard picture are freely aired.

REFERENCES

Rather than give references to individual papers I list below some text books which provide these in great detail:

1. S. Weinberg (1972) Gravitation and Cosmology, New York, Wiley (For general relativity and relativistic cosmology).
2. J.V. Narlikar and A.K. Kembhavi (1980) Non-standard Cosmologies, Fundamentals of Cosmic Physics 6, Nos. 1,2. (for non-standard cosmologies).
3. F. Hoyle and J.V. Narlikar (1974) Action at a Distance in Physics and Cosmology, New York, W.H. Freeman.
4. J.V. Narlikar (1983) Introduction to Cosmology, Boston, Jones and Bartlett (for standard cosmology, nonstandard cosmologies and observational tests of cosmologies).
5. J.V. Narlikar and T. Padmanabhan (1986) Gravity, Gauge Theories and Quantum Cosmology, Dordrecht, D. Reidel (for standard cosmology, particle physics and the early universe).

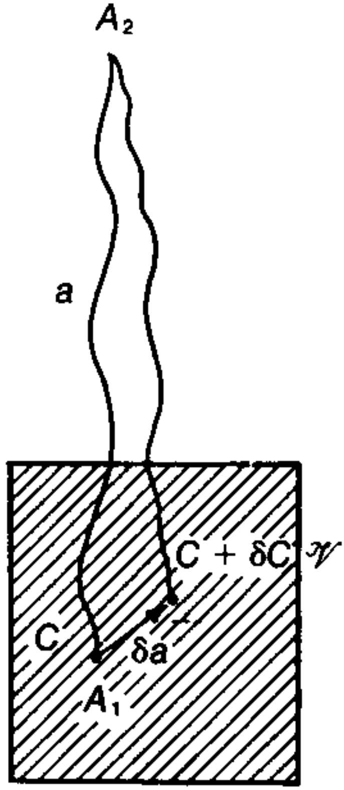

FIGURE 1 : The world line of a begins at A_1 and ends at A_2. If we consider variations in the shaded region, the point A_1 shifts by δa^i. This shift produces a change in the C-field interaction term by an amount $-\delta C = C_i \ \delta a^i$. The change in the inertial part of the action similarly makes a contribution at A_1 of $p_i^{(a)} \delta a^i$ where $P_i^{(a)}$ is the 4 momentum of the particle a. The result (II.7) follows by equating the net contribution of $\delta \mathcal{A}$ at A_1 to zero.

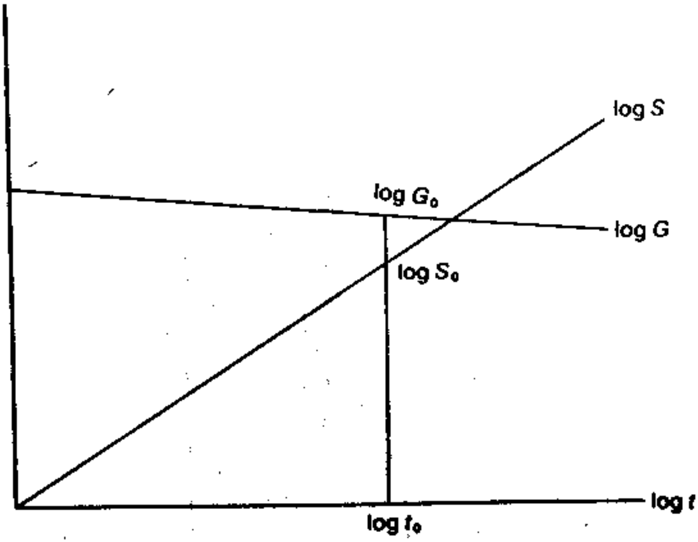

FIGURE 2 : The temporal behavior of S and G. Both are plotted on a log-log plot for $\omega = 6$. The scales are arbitrary.

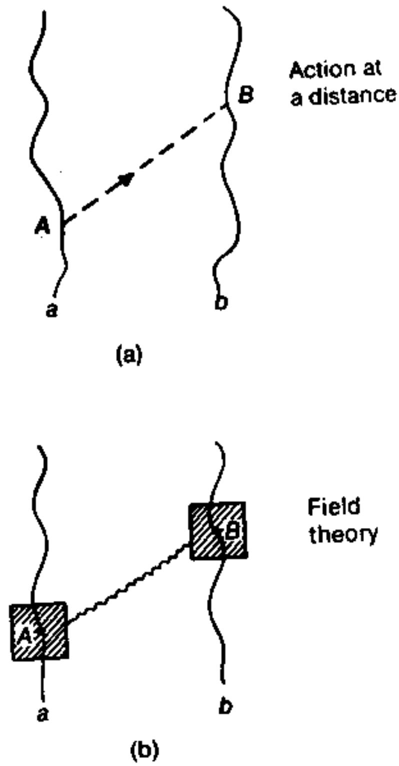

FIGURE 3 : (a) In the action-at-a-distance picture the influence from the point A on the world line of particle a is transmitted directly across spacetime (along the dotted track) to the point B on the world line of particle b. (b) In field theory the field in the neighborhood of A (shown by the shaded region) is disturbed: the disturbance propagates across spacetime as a wave in the ambient field and reaches the neighborhood of B. The disturbance then exerts a force on b at B. This is how the influence propagates from a to b.

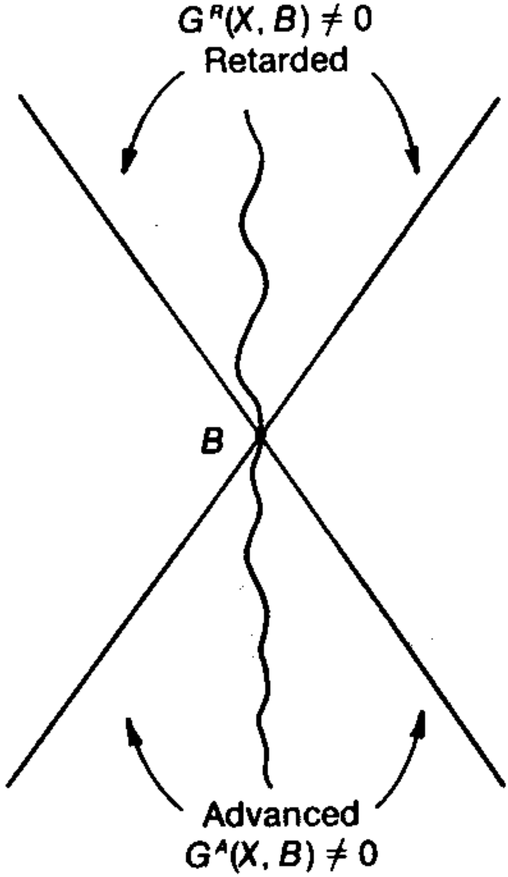

FIGURE 4 : The retarded Green's function of B is nonvanishing only in the future light cone of B, while the advanced Green's function is nonvanishing only in the past light cone.

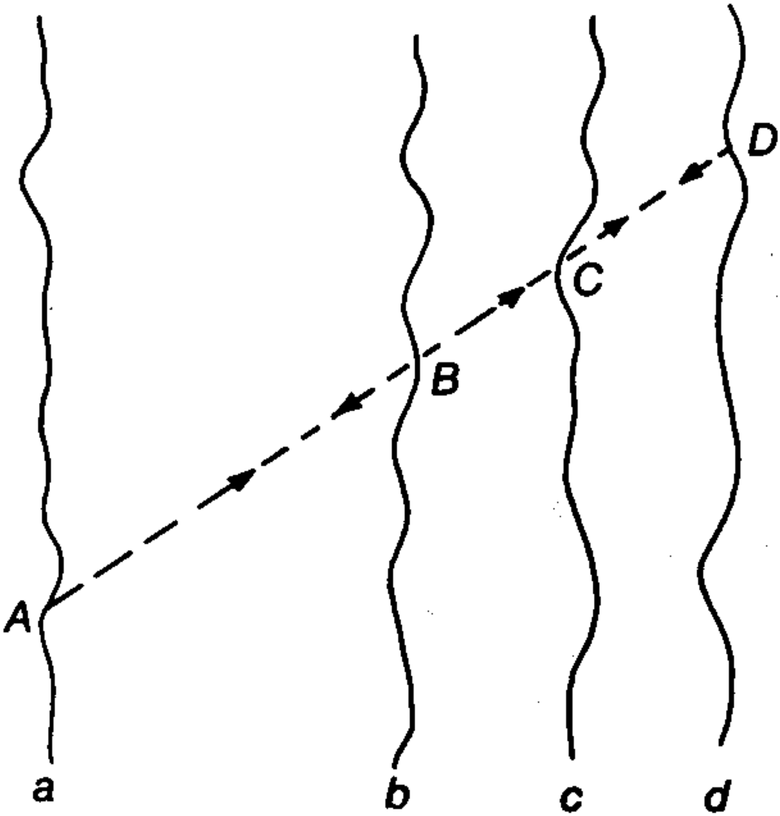

FIGURE 5 : A retarded signal (shown by dotted line) leaving point A on the world line of a hits particles b,c,d,... on points B, C, D,... Their advanced response returns to A along the same dotted track, no matter how far these particles are from a. Thus even the remote parts of the universe generate instantaneous responses to the retarded disturbance leaving A.

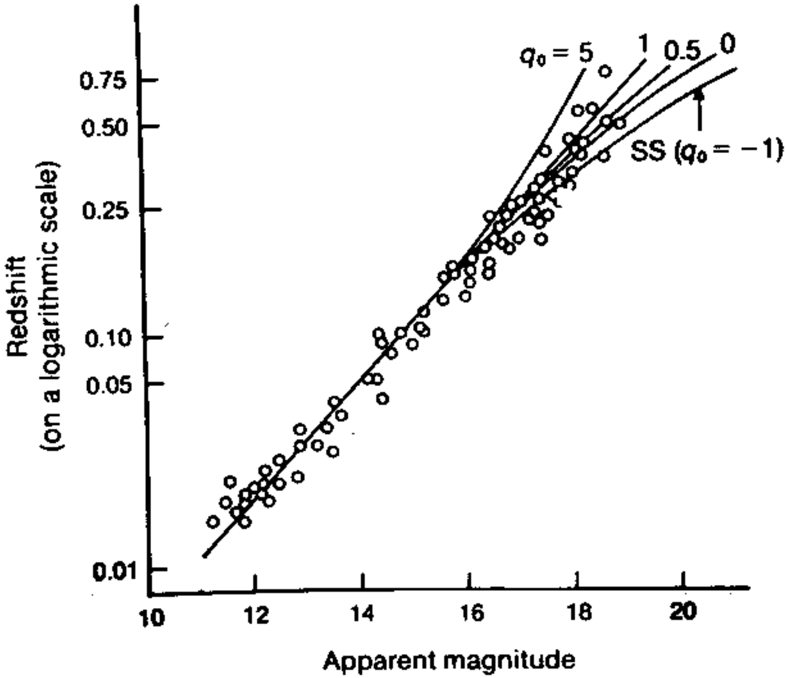

FIGURE 6 The redshift magnitude relation for the brightest cluster members. A number of theoretical curves (q_o = 5,1,0.5,0, - 1) are superposed on the data. SS stands for the steady state model. [Based on J. Kristian, A. Sandage, and J.A. Westphal, 1978, "The extension of the Hubble diagram-III," Ap. J.221, 383.]