PROBLEMS IN RELATIVITY AND COSMOLOGY

Thesis submitted to the
University of London

By

WILLIAM DAVIDSON M.Sc.
For the Degree of Doctor of Philosophy

Royal Holloway College,
(University of London)
Englefield Green,
Surrey.

Date: March 1958

ABSTRACT

The thesis consists of three parts, dealing with three problems of relativity and cosmology, whose titles and abstracts are as follows.

PART I: THE THEORY OF OBSERVABLE CRITERIA FOR THE SOLUTION OF THE COSMOLOGICAL PROBLEM.

A systematic derivation and critical analysis of observable relations in cosmology is presented. Many new formulae are obtained by the author, applicable to both an evolutionary universe and the steady state universe. It is demonstrated how the observable relations can, in principle, decisively distinguish between these two types of universe. The results are used to analyse the redshift data published by Humason, Mayall and Sandage in 1956.

The author has also derived formulae to utilise the developing techniques of colour photometry and radio astronomy. These are applied to the Stebbins-Whitford type of analysis of distant spectra, and to Ryle's phenomenon of colliding galaxies respectively.

PART II: GENERAL RELATIVITY AND MACH'S PRINCIPLE.

Believing that cosmological solutions involving empty space at infinity are logically inadmissible in general relativity, the author has generalised, with certain restrictions, an analysis by Einstein of Mach's Principle in quasi-Galilean space-time to an arbitrary space-time. It is shown that by adopting a steady state cosmology, which recommends itself on

the grounds of logical simplicity, general relativity can fully account for inertia. The presentation is considered to be a substantial improvement on a previous publication by the author. Additional supporting material is advanced together with a comprehensive historical and critical analysis of the problem.

PART III: THE MECHANISM OF STEADY STATE COSMOLOGY ACCORDING TO GENERAL RELATIVITY.

The stationary exterior and interior solutions for a spherically symmetric concentration or rarefaction of mass in the steady state universe are obtained and analysed. It is shown that in a rarefaction the density of inertia becomes negative.

These solutions are utilised in the development of a theory of the steady state as one of dynamic equilibrium, between matter on the one hand and radiation of negative average energy density on the other. It is shown that the creation of matter leads to the emission of gravitational waves carrying negative energy, and the view is put forward that a radiation field must be identified with a quantised gravitational field.

Acknowledgments

I wish to acknowledge here the patient training in research methods, as well as the education and proper perspective in subject matter, which I have received from my supervisor of research, Professor W. H. McCrea. For his stimulating criticism, encouragement, and friendly advice at all times

I express most grateful appreciation.

I should like to thank also the Principal of Royal Holloway College, Dr. E.C. Batho, and the Principal of the Battersea College of Technology, Dr. R. West, for facilities to perform this research.

CONTENTS

PART I: THE THEORY OF OBSERVABLE CRITERIA FOR THE SOLUTION OF THE COSMOLOGICAL PROBLEM

PART II: GENERAL RELATIVITY AND MACH'S PRINCIPLE

PAGE NO.

PART III: THE MECHANISM OF STEADY STATE COSMOLOGY ACCORDING TO GENERAL RELATIVITY

PAGE NO.

PART I

THE THEORY OF OBSERVABLE CRITERIA FOR THE SOLUTION OF THE COSMOLOGICAL PROBLEM

INTRODUCTION

A comprehensive review of the literature on cosmology, and of papers on the theory of observable relations in particular, will not be given in this introduction since it has been thought best to deal with the work of previous authors in the context of the subject matter dealt with progressively in this thesis. Suffice it to say here that, invaluable as the pioneer work of these authors has been, much of it is inadequate to deal with the cosmological observational programme of the present day, both by virtue of the new theoretical problems that have arisen - the steady state theory, the so called Stebbins Whitford effect, radio stars, to mention but a few - and also because of new instrumental techniques, such as photoelectric photometry and radio astronomy. A systematic presentation of the observable relations, with a thorough investigation of their significance for present day cosmology, has been wanting in recent years and this part of the present thesis was initiated early in 1955 to help to fill this gap.

Since then the only publication which has gone some way to answering the same need is a book by G. C. McVittie (<u>16</u>) which includes a discussion of the observable relations with attention to some of the present day problems. However, because of its less exhaustive analysis and considerably smaller range of investigation, it has by no means dispensed with the need for this thesis.

All the formulae to follow have been derived independently by the present author but manifestly some will have been obtained by previous writers. Where the author has been aware of this reference is made to the original derivation - although it will not necessarily be the same as that given in the thesis. Likewise it is stated when the author's work is believed to be original.

CHAPTER I: OBSERVABLE QUANTITIES RELATED TO THE THEORETICAL PARAMETERS OF THE COSMOLOGICAL METRIC

(i) The cosmological metric

The general theory of relativity assumes, as a natural generalisation of special relativity, that the metrical properties of 4-dimensional space-time are characterised by the 'interval' S between any two events in space-time. This quantity is related to the space-time coordinates according to the differential quadratic form (invariant)

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

with summation over μ and ν which take the values 1,2,3,4, covering the coordinates of space and time. Thus the geometry of space-time is Riemannian.

The $g_{\mu\nu}$ are functions of the x^μ and, according to general relativity, are related to the distribution of matter and energy in the reference frame used (in principle arbitrary) according to Einstein's gravitational field equations (1). Thus equation (1.1) gives the 'line element' or 'metric' of space-time corresponding to a given distribution of mass.

The interval ds between adjacent events in space-time is physically measurable by rigid metre stick or by a clock according to whether the interval is space-like ($ds^2 < 0$, taking the signature of the quadratic form as negative), or time-like $(ds^2 > 0)$ respectively. This is because by a suitable transformation of coordinates the invariant form (1.1) can be reduced, in the neighbourhood of any assigned event, to the canonical form of special relativity (without change of signature), where the physical identification of the coordinates is well known from that theory. We shall return to this matter in Section (ii).

The gravitational equations imply that where there is, for example, spherical symmetry about the spatial origin in the distribution of mass then this will be reflected in the $g_{\mu\nu}$ in that the metric will be unaltered for pure rotation of the axes about the spatial origin. This principle implied by the field equations has been called the 'Principle of the geometrization of physics'.

On the basis of this principle, without reference to the field equations as such, it is possible from purely geometric and kinematic considerations to deduce the form of metric corresponding to a given physical situation, possessing a degree of symmetry or uniformity, up to a minimum number of residual undetermined parameters.

For the case of the cosmic metric of the expanding universe, with which we are concerned in this thesis, this has been achieved by H. P. Robertson by two derivations. The first (2) was without regard to the physical definition of the coordinates and based on the point of view of the invariant properties of a Riemannian metric under certain group transformations which expressed the desired geometrical symmetry. The second (3) was from the point of view of the space-time (Riemannian) geometry determined by the measurements of a 3 parameter set of 'equivalent observers' whose physical observations had the desired symmetry, the space time coordinates being defined in terms of these measurements according to the operational methodology of light signals, theodolites and clocks only, that was first employed by E. A. Milne in his theory of kinematic relativity. A derivation similar to this last method was given independently by A. G. Walker (4).

The symmetry properties referred to include a certain body of assumptions regarding the universe as whole which in various forms is known as the 'cosmological principle'. The form which is most economical with respect to the derivation of the cosmological metric is that adopted by R. C. Tolman (5). Tolman derives the metric based on the single inference from astronomical observations, that as regards the large scale festures of the universe there is spherical symmetry about every observer with respect to whom the average motion of matter in his (sufficiently large) neighbourhood is zero. Such observers will later be discussed as 'fundamental observers'. Adopting 'co-moving' coordinates, i.e. so that an element of matter has permanent spatial coordinates, Tolman shows that this consideration, if local irregularities are neglected and ideal spatial isotropy assumed for every point, leads to homogeneity in the spatial metric and to the existence of a cosmic time. In fact the cosmological metric determined in this way has the form

$$ds^2 = c^2 dt^2 - \frac{R^2(t)\left(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2\right)}{(1 + kr^2/4)^2} \quad \ldots\ldots\ldots \ (1.2)$$

The function $R(t)$ is a disposable function of the cosmic time t which, interpreted geometrically, is such that $k/R^2(t)$ is the curvature of the 3-space $t=$ constant at the cosmic time t. The parameter k is a constant which takes the values 1, 0, -1 according as this curvature is positive, zero, or negative.

In deducing this metric the assumed spatial isotropy implies that the contents of the universe are smoothed out to an ideally homogeneous state. Such a metric is therefore relevant to the study of cosmic properties in which small scale irregularities are unimportant, and particularly useful as a theoretical background against which may be plotted the now far reaching observations of the expanding universe.

In the rest of this chapter the theoretical parameters occurring in the universes of metric given by (1.2) will be related to the observable quantities employed by astronomers. The object of this programme will be to develop relations between the observable quantities corresponding to different values of $R(t)$ and k with which the observed relations may be compared, and so to provide the means to determine those values of $R(t)$ and k which fit the actual universe.

(ii) Locally measured time and distance by fundamental observers

It is an essential feature of general relativity that at any point of space-time the differential quadratic form (1.1) linking adjacent events in its immediate neighbourhood can, by a suitable transformation of coordinates, be reduced to the canonical form of special relativity. As stated in Section (i) this is a consequence of the invariant character ascribed to the interval s. These 'local' coordinate systems have therefore the metric

$$ds^2 = c^2 dT^2 - dx^2 - dy^2 - dz^2 \qquad \text{.........(1.3)}$$

In general the first derivatives of the $g_{\mu\nu}$ in these local systems do not vanish at the point in question. It is a consequence of the principle of equivalence that general relativity theory always permits the setting up of 'natural' local coordinate systems in which the first derivatives of the $g_{\mu\nu}$ do vanish at the point in question. Such systems are characterised by the fact that in them the gravitational field vanishes at the point in question, that is they are 'freely falling' local coordinate systems. In these systems a free particle (stress absent) moves locally without acceleration and the local portion of its world line satisfies the differential equations of the straight line path of a free particle in special relativity. Thus natural coordinate systems are locally inertial.

In a general space-time the world lines of freely moving particles in general relativity are geodesic curves, on which the interval between adjacent points is time-like $(ds^2 > 0)$ and which satisfy differential equations which are the covariant generalisation of the equations of the straight line paths of special relativity.

It follows that at any event we can imagine a local 'free falling' observer moving on a geodesic through the point making local measurements in a natural coordinate system for which the metric is (1.3). For a time-like interval ds between adjacent events, given in a general reference system by (1.1), we can therefore provide a direct physical interpretation of ds in principle by imagining an observer, moving on a suitable chosen geodesic world line, in fact through the events, who registers

ds/c as $d\tau$ on his clock, while for him $dx = dy = dz = 0$.
Similarly for a space-like interval ds a suitably chosen observer will measure $i\,ds$ $(i = \sqrt{-1})$ as the spatial length dl where $dl^2 = dx^2 + dy^2 + dz^2$, while for him $d\tau = 0$.
The time $d\tau$ registered on a clock moving with such an observer, between neighbouring events on his world line, may be called the 'proper' time between such events, while the distance $dl = (dx^2 + dy^2 + dz^2)^{\frac{1}{2}}$ registered on his metre stick between events in his neighbourhood which are simultaneous for him $(d\tau = 0)$ may be called the proper distance between such events.

In the case of the cosmological metric (1.2) it is easily shown that the curves in space-time given by $r = \text{const.}$, $\theta = \text{const.}$, $\phi = \text{const.}$, are time-like geodesics. Thus this reference frame which was designed to be co-moving with the contents of the universe can, at all points of space-time, be locally identified with a co-moving local inertial system of metric given by (1.3) whose space origin has constant r, θ, ϕ coordinates. Observers at the space origin of such locally inertial reference systems will be called 'fundamental observers' since they are permanently attached to freely moving material of the universe. That the world lines of matter are geodesics here does not depend on the absence of stress, since such stress must be isotropic and a function of cosmic time only, because of the isotropic and homogeneous character of the geometry.

By a comparison of (1.2) and (1.3) it is seen that the cosmic time t measures proper time for fundamental observers everywhere. It follows also that the spatial part of the metric (1.2) is to be identified with the local proper distances, measured by the fundamental observer at (r, θ, ϕ) at epoch t, corresponding to coordinate variations $dr, d\theta, d\phi$.

(iii) Redshift or Doppler effect

It is another natural generalisation of special relativity that in general relativity the world lines or trajectories in space-time of light rays are, in the absence of matter (empty space), null geodesics characterised by certain covariant differential equations and for which the interval ds between neighbouring points vanishes. In terms of the locally inertial metrics as given by (1.3) this means that the locally measured velocity of light is the constant c.

Although in our smoothed out cosmological medium space cannot be regarded as empty, we shall nevertheless suppose in agreement with observation that the smoothed out density would be so small as not to affect the velocity of light. For the cosmological metric (1.2) it is easily shown that the null geodesics through the space origin are the lines $\theta = \text{const.}, \phi = \text{const.}$

Accordingly consider the radially moving light pulse, emitted from a 'particle' at (r, θ, ϕ) at epoch t which reaches the origin $r = 0$ at epoch t_0. Since $ds = 0$ we have $c^2 dt^2 - R^2(t)\, dr^2 / (1 + kr^2/4)^2 = 0$ along the light path so that on integration we obtain

$$c\int_t^{t_0} \frac{dt}{R(t)} = \int_0^r \frac{dr}{1 + kr^2/4} \qquad \text{.................................} \quad (1.4)$$

If the emission is of duration dt and the corresponding reception is of duration dt_0, then differentiating (1.4) keeping r constant we get

$$-c\frac{dt}{R(t)} + c\frac{dt_0}{R(t_0)} = 0$$

whence
$$\frac{dt_0}{dt} = \frac{R(t_0)}{R(t)} \qquad \text{.......................} (1.5)$$

If ν is the frequency of emission and ν_0 that of the observed received light then, counting wave crests, we must have

$$\nu_0\, dt_0 = \nu\, dt \qquad \text{.......................} (1.6)$$

so that
$$\frac{\lambda_0}{\lambda} = \frac{\nu}{\nu_0} = \frac{R(t_0)}{R(t)} \qquad \text{.......................} (1.7)$$

where λ, λ_0 are the locally observed wavelengths of emission and reception respectively.

For an expanding universe the proper distance between neighbouring particles of fixed r, θ, ϕ must increase with epoch so that it follows from the analysis in Section (ii) that $R(t_0) > R(t)$. Equation (1.7) therefore indicates that there will be a shift of the spectrum of distant sources towards the longer wavelengths, due to the Doppler effect of recession, by a fraction $\delta \equiv \delta\lambda/\lambda = R(t_0)/R(t) - 1$ which is independent of wavelength. In deducing this result it is assumed that the characteristic emissions of an atom are of frequency independent of epoch and of location.

The redshift formula (1.7) is now well known but seems to have been first derived by G. Lemaitre (6) for expanding cosmological models. Alternative explanations of the observed systematic redshift, other than by the Doppler effect due to a genuine expansion of the universe, have been put forward in the past, so that observational tests which could discriminate between these different theories would be a valuable asset. This matter will be dealt with in Chapter II Section (viii).

(iv) Local redshift - distance relation

If we characterise the redshift by the fraction $\delta\lambda/\lambda \equiv \delta$ then

$$\delta = \frac{R(t_0)}{R(t)} - 1 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.8)}$$

Expanding the right hand side in powers of $\tau \equiv t_0 - t$ by Taylor series we get as a first approximation for near sources

$$\delta \simeq \frac{\dot{R}(t_0)}{R(t_0)}\,\tau \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.9)}$$

where $\dot{R}(t_0) \equiv d\{R(t_0)\}/dt_0$.

The integrated proper distance, as measured by rigid measuring rods, by a succession of fundamental observers in the simultaneity of their clocks when they all register cosmic proper time t, from $r=0$ to a point (r, θ, ϕ), we shall call the total proper distance between these points at epoch t. By our analysis of elementary proper distance this is clearly, for epoch t_0,

$$\ell = R(t_0)\int_0^r \frac{dr}{1+kr^2/4} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.10)}$$

By equation (1.4), to a first approximation for small τ we can write

$$\int_0^r \frac{dr}{1+kr^2/4} \simeq \frac{c\,\tau}{R(t_0)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.11)}$$

whence $\quad \ell \simeq c\,\tau \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots$ (1.12)

Hence by (1.9)

$$\delta \simeq \frac{\dot{R}(t_0)}{c\,R(t_0)}\,\ell \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1.13)}$$

Thus redshift δ is directly proportional to measured distance in the neighbourhood of any fundamental observer.

If we differentiate (1.10) with respect to t_0, keeping r constant, we obtain the local velocity-distance relation characterising the expansion

$$\dot{l} = \{\dot{R}(t_0)/R(t_0)\}l \qquad \text{.............. (1.14)}$$

$$\simeq c\delta \qquad \text{.............. (1.15)}$$

on using (1.13). Thus we see that the rate of change of proper distance as the epoch advances is equal, in the first approximation, to the Doppler velocity $c\delta$ given by the classical formula and both are proportional to proper distance at the epoch of observation.

The distance assigned to luminous objects by astronomers is in general different from l as given by (1.10) but, as will be shown in the next section, it is identifiable with l for sources near the observer. Accordingly we recognize from (1.14) and (1.15) that the constant $\dot{R}(t_0)/R(t_0)$ has to be identified with the so-called Hubble parameter in the redshift-distance law obtained by E. Hubble (7), who found that the Doppler velocity of recession of the nebulae was proportional to astronomical distance. As we shall see in the next section this relation is only linear in the first approximation i.e. for nearby nebulae only.

(v) Astronomical distance, bolometric luminosity, bolometric magnitude

The distance D assigned to a luminous object by astronomers is determined by its apparent luminosity L^*_{bol}, assuming an inverse square law of diminution of intensity, provided the absolute power E^* of the source is known.

Thus if E^* is the energy radiated in unit time from the source over all frequencies, while L^*_{bol} is the received energy flux per unit area, then D is given by the equation

$$L^*_{bol} = \frac{E^*}{4\pi D^2} \qquad \text{.....................} \quad (1.16)$$

The distance D defined in this way we shall call the 'luminosity distance'.

The quantity E^* is measured by absolute bolometric magnitude M_{bol}, while L^*_{bol} is measured by apparent bolometric magnitude m_{bol}. M_{bol} is the apparent magnitude at a distance of 10 parsecs. A source of apparent magnitude m is (apparently) $10^{2/5}$ times as bright as one of apparent magnitude $m+1$ by definition. The term bolometric indicates total energy radiated over all frequencies when applied to absolute intensity E^*, and when applied to apparent luminosity it indicates total flux of energy, received per unit area, over all frequencies, at a point outside the earth's atmosphere (that is before the radiation suffers heterochromatic absorption).

If $L^*_{0\,bol}$ is the luminosity at 10 parsecs it follows that

$$L^*_{0\,bol} = \frac{E^*}{4\pi 10^2} \qquad \text{...............} \quad (1.17)$$

By definition of magnitudes

$$\frac{L^*_{bol}}{L^*_{0\,bol}} = 10^{\frac{2}{5}(M_{bol} - m_{bol})} \qquad \text{...............} \quad (1.18)$$

Hence if D is in parsecs it follows that

$$\log_{10} D = \cdot 2(m_{bol} - M_{bol}) + 1 \qquad \text{...............} \quad (1.19)$$

The introduction of a theoretical distance arising from the assumption of an inverse square law of diminution of luminosity, in agreement with the practice of astronomers, is due to E. T. Whittaker (8) who defined the astronomical distance precisely as we have defined the luminosity distance D. In the paper quoted Whittaker obtained the expression for astronomical distance in the de Sitter universe, although for an unusual form of the metric. The expression for D was implicit in a formula for luminosity derived in an earlier paper (9) by R. C. Tolman for the general cosmological metric, taken in a form related by a simple change of radial coordinate to (1.2), although Tolman did not concern himself with a formula for D as such. The formula for D appropriate to our metric (1.2) was derived by W. H. McCrea in a later publication (10). As the result requires special interpretation for the recent developments in cosmological theory and in observational technique, to be dealt with in this thesis, we shall give our own derivation here.

Suppose that from the source at (r, θ, ϕ) the energy of radiation in the waveband $\lambda, \lambda+d\lambda$ which is emitted in the interval $t, t+dt$ is

$$E(\lambda, t)\, d\lambda\, dt$$

This radiation will subsequently lie in a spherical shell which will pass over the origin $r=0$ in the interval t_0, t_0+dt_0 say. At time t_0 the integrated proper area of the shell according to fundamental observers on it will be

$$\frac{4\pi R^2(t_0)\, r^2}{(1+kr^2/4)^2}$$

The total integrated energy of the radiation apparent to such observers will be

$$E(\lambda, t)\, d\lambda\, dt\, \frac{R(t)}{R(t_0)}$$

due to the degradation caused by the recession (equation (1.7)). Accordingly the flux of energy per unit proper area at the origin will be

$$L_{bol}(\lambda_0)\,d\lambda_0 = \frac{E(\lambda, t)\, d\lambda \cdot \frac{dt}{dt_0} \cdot \frac{R(t)}{R(t_0)}}{4\pi R^2(t_0)\, r^2 / (1 + kr^2/4)^2}$$

Units of energy per unit area per unit time

By equation (1.5) we have $\frac{dt}{dt_0} = \frac{R(t)}{R(t_0)}$ and if we consider now emission and reception of radiation of all wavelengths and put

$$\int_0^\infty E(\lambda, t)\, d\lambda = E^*(t) \qquad \text{.............. (1.20)}$$

$$\int_0^\infty L_{bol}(\lambda_0)\, d\lambda_0 = L^*_{bol} \qquad \text{.............. (1.21)}$$

we can write

$$L^*_{bol} = \frac{E^*(t)}{4\pi D^2} \qquad \text{.............. (1.22)}$$

where

$$D = \frac{R^2(t_0)\, r}{R(t)(1 + kr^2/4)} \qquad \text{.............. (1.23)}$$

This is the required relation of luminosity distance D to the theoretical parameters of the cosmological metric (1.2), and is the expression for D obtained by McCrea. We note that if r is small, so that $R(t) \simeq R(t_0)$, then $D \simeq R(t_0)\, r$ which agrees with the expression for proper distance ℓ given by (1.10), when r is small. This justifies the assertion made at the end of Section (iv) that the distance assigned by astronomers to luminous objects agrees, for the neighbourhood of a fundamental observer, with the distance which would be obtained by that observer using ordinary trigonometric measurement based on Euclidean geometry. Thus we may write the redshift-distance law expressed by (1.13) in the form

$$c\,\delta \simeq \frac{\dot{R}(t_0)}{R(t_0)} \cdot D \qquad \text{................(1.24)}$$

for observations in the vicinity of a fundamental observer at epoch t_0.

In our derivation of D above we have expressly and formally allowed for the dependence of the absolute intensity of the source on epoch viz. $E^*(t)$. Recent developments in observational technique have made feasible the detection of an evolutionary process in the nebulae, and attempts have been made with this objective by J. Stebbins and A. E. Whitford (11,18) based on the determination of the colour indices of the galaxies. Such a consideration is of great importance as, for instance, it would distinguish between the expanding models of general relativity (12), for which evolution can be expected on the basis of conservation of energy, and the steady state models of H. Bondi and T. Gold (13), and of F. Hoyle (14), where in contrast there must be an absence of systematic evolution due to continual creation of new galaxies. To cast our theoretical formula (1.19), relating distance and apparent magnitude, into a form suitable for the detection of such effects in principle we proceed as follows:

Suppose M^0_{bol} is the absolute bolometric magnitude of nearby nebulae at the time t_0, and $E^*(t_0)$ the corresponding absolute intensity. Then we have the equation

$$m_{bol} - M^0_{bol} = 2.5 \log_{10}\left(\frac{L^*_{0\,bol}}{L^*_{bol}}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots (1.25)$$

where m_{bol} is the apparent magnitude of a distant nebula, L^*_{bol} its corresponding apparent luminosity given by (1.22), and

$$L^*_{0\,bol} = \frac{E^*(t_0)}{4\pi 10^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1.26)$$

is the luminosity of a nearby source at the distance of 10 pc.

Hence we can write

$$\log_{10} D = \cdot 2 \left\{ m_{bol} - M_{bol} + 2\cdot 5 \log \left(\frac{E^*(t)}{E^*(t_0)} \right) \right\} + 1 \quad \ldots\ldots \quad (1.27)$$

Since, as will be shown in the next chapter, D can be expanded in powers of δ with coefficients depending on the model concerned, while $E^*(t)$ can also be expanded relative to $E^*(t_0)$ in powers of δ , equation (1.27) demonstrates the possibility in principle of detecting evolutionary effects. However we shall see in the next section that the use of the heterochromatic magnitudes actually measured by astronomers is more conducive to the practical determination of such results, rather than the bolometric magnitudes only obtainable indirectly.

(vi) Heterochromatic magnitudes

The magnitudes directly measured by astronomers are not bolometric corresponding to total energy received outside the earth's atmosphere, but 'heterochromatic' corresponding to the selective effects of the atmosphere and sensitivity of photographic plates and other detective devices. Accordingly these heterochromatic magnitudes must be 'corrected' to obtain bolometric magnitudes. If bolometric magnitudes are used to theoretically investigate the results of the astronomical observations, as has been the custom in the literature, it is important to apply this correction as accurately as possible. As the existing theory of the correction given in the literature does not appear to be in the best form for dealing with the theoretical problems, such as the determination of a possible evolutionary effect, we shall examine this matter here.

We shall find that no reference at all to bolometric magnitudes is necessary in the formulae connecting observational quantities, which in fact find their most natural expression in terms of the magnitudes actually measured.

A brief theoretical analysis of the magnitude correction was included in the important report on the latest redshift measurements by Humason, Mayall, and Sandage (15). However the analysis there is unsatisfactory and ambiguous for reasons that will be pointed out in Chapter II in connection with equation (2.30). The following analysis by the present writer is similar in its general form to a recent exposition in a book by G. C. McVittie (16), but follows independent lines both in the derivation and detail of the resulting 'correction' formula which, for reasons that will be given later, is considered to be more satisfactory than McVittie's analogous result. It will form the basis for much original application and critical analysis in this thesis.

Corresponding to the emission as in Section (v) of an amount of radiation $E(\lambda, t)\,d\lambda\,dt$ from the source at (r, θ, ϕ) in time dt the energy, received now per unit area at the surface of the earth by the detective device, will be given by multiplying the previous result for the energy arriving outside the earth's atmosphere by a sensitivity function of received wavelength $s(\lambda_0)$ viz.,

$$L_{net}(\lambda_0)\,d\lambda_0 = \frac{s(\lambda_0)\,E(\lambda, t)\,d\lambda}{4\pi D^2}$$

where D is given by (1.23). Now by equation (1.7)

$$\lambda = \frac{\lambda_0}{1+\delta} \quad , \quad d\lambda = \frac{d\lambda_0}{1+\delta} \qquad \text{............(1.28)}$$

Hence

$$L^{*}_{het} = \int_0^\infty L_{het}(\lambda_0)\, d\lambda_0 = \left\{ \frac{1}{1+\delta} \int_0^\infty s(\lambda_0)\, E\left(\frac{\lambda_0}{1+\delta}, t\right) d\lambda_0 \right\} \Big/ \left\{ 4\pi D^2 \right\} \qquad \text{.....(1.29)}$$

The corresponding observed luminosity of a <u>local</u> source of the same general type at a distance of 10 parsees will be

$$L^{*}_{\substack{o\\het}} = \int_0^\infty L^{o}_{het}(\lambda_0)\, d\lambda_0 = \left\{ \int_0^\infty s(\lambda_0)\, E(\lambda_0, t_0)\, d\lambda_0 \right\} \Big/ \left\{ 4\pi 10^2 \right\} \qquad \text{.....(1.30)}$$

where we have put the parameter t equal to t_0 allowing for the distinction between near and distant nebulae which might arise from an evolutionary effect.

It is to be noted that the difference in the apparent absolute intensity of a near and distant source (comparing (1.29) and (1.30)) arises not solely from a possible evolutionary effect, but includes also an effect on selectivity which is a function of the redshift δ . Mathematically this is because $s(\lambda_0)$ is a function of <u>received</u> wavelength. This gives rise to a term in the relation between heterochromatic and bolometric magnitudes which has been called the K correction. To find the total correction we proceed as follows.

Let m_{het} be the apparent heterochromatic magnitude of a distant source and M^{o}_{het} the absolute heterochromatic magnitude of a nearby source. Then we must have

$$m_{het} - M^{o}_{het} = 2.5 \log_{10}\left(\frac{L^{*}_{het}}{L^{*}_{het}} \right) \qquad \text{..............(1.31)}$$

It follows from (1.25) and (1.31) that the bolometric magnitude of the distant source is derived from the measured heterochromatic magnitude by the relation

$$m_{bol} - M^{o}_{bol} = m_{het} - M^{o}_{het} - 2.5 \log_{10} \left\{ \frac{L^{*}_{bol}}{L^{*}_{\substack{o\\bol}}} \Big/ \frac{L^{*}_{het}}{L^{*}_{\substack{o\\het}}} \right\} \qquad (1.32)$$

That is from the heterochromatic magnitude modulus on the right we substract the quantity

$$2{\cdot}5\log_{10}\left\{\frac{\frac{1}{1+\delta}\int_0^\infty E\left(\frac{\lambda_0}{1+\delta},t\right)d\lambda_0}{\int_0^\infty E(\lambda_0,t_0)\,d\lambda_0}\right\}-2{\cdot}5\log_{10}\left\{\frac{\frac{1}{1+\delta}\int_0^\infty s(\lambda_0)\,E\left(\frac{\lambda_0}{1+\delta},t\right)d\lambda_0}{\int_0^\infty s(\lambda_0)\,E(\lambda_0,t_0)\,d\lambda_0}\right\}$$

..........(1.33)

If there were no evolutionary effect then we could put $t = t_0$ in this expression, so that its first term would then vanish. The second term would then represent a K correction, arising from the selective effect of redshift only. This has been the only correction attempted until recently in theory or practice. In general however, if we allow for an age effect both terms must be considered. The first term is just

$$2{\cdot}5\log_{10}\left\{\frac{E^*(t)}{E^*(t_0)}\right\}$$

which occurs in equation (1.27). It therefore follows from (1.27), (1.32), (1.33) that

$$\log_{10} D = {\cdot}2\left[m_{het}-M^0_{het}+2{\cdot}5\log_{10}\left\{\frac{\frac{1}{1+\delta}\int_0^\infty s(\lambda_0)\,E\left(\frac{\lambda_0}{1+\delta},t\right)d\lambda_0}{\int_0^\infty s(\lambda_0)\,E(\lambda_0,t_0)\,d\lambda_0}\right\}\right]+1$$

............(1.34)

From this relation is seen the advantage of not considering bolometric magnitudes at all in relating luminosity distance D (which can be expanded, as we shall show in the next chapter, in powers of δ) to the measured heterochromatic magnitudes. The first two terms in the square brackets are measured quantities. The third involves a K 'correction' depending on the energy distribution in the nebular spectrum and on δ, together with a 'correction' due to a possible evolutionary effect; there is also a dependence on the detective device used - via the sensitivity function. A knowledge of the orthodox K correction may be derived by empirical means involving a study of the energy distribution in the spectra of <u>nearby</u> nebulae, for which the evolutionary correction would be zero, and applying this to the mathematically well defined correction term given in (1.34). To separate the K correction from the evolutionary effect for nebulae at arbitrary distance we shall expand the total correction term in powers of δ. It will be sufficient for the purpose of the theory to retain terms as far as first order in δ only. Thus, to order δ,

$$2.5 \log_{10}\left\{\frac{\frac{1}{1+\delta}\int_0^\infty S(\lambda_0)\, E\left(\frac{\lambda_0}{1+\delta}, t\right) d\lambda_0}{\int_0^\infty S(\lambda_0)\, E(\lambda_0, t_0)\, d\lambda_0}\right\}$$

$$= 2.5 \log_{10}\left[\frac{\int_0^\infty S(\lambda_0)(1-\delta)\left\{E(\lambda_0,t_0) - \delta\lambda_0 E'(\lambda_0,t_0) - \tau \dot{E}(\lambda_0,t_0)\right\} d\lambda_0}{\int_0^\infty S(\lambda_0)\, E(\lambda_0,t_0)\, d\lambda_0}\right]$$

where

$$\tau \equiv t_0 - t\ , \quad E'(\lambda_0,t_0) \equiv \frac{\partial}{\partial\lambda}\left\{E(\lambda,t)\right\}_{\substack{t=t_0\\ \lambda=\lambda_0}}\ , \quad \dot{E}(\lambda_0,t_0) \equiv \frac{\partial}{\partial t}\left\{E(\lambda,t)\right\}_{\substack{t=t_0\\ \lambda=\lambda_0}}$$

Now by (1.9), $\tau = \delta/\alpha_1$, to first order, where $\alpha_1 \equiv \dot{R}(t_0)/R(t_0)$.

Thus the correction term becomes to order δ

$$2.5 \log_{10}\left[\frac{\int_0^\infty S(\lambda_0)\left\{E(\lambda_0,t_0) - \delta\left(E(\lambda_0,t_0) + \lambda_0 E'(\lambda_0,t_0) + \frac{1}{\alpha_1}\dot{E}(\lambda_0,t_0)\right)\right\} d\lambda_0}{\int_0^\infty S(\lambda_0)\, E(\lambda_0,t_0)\, d\lambda_0}\right]$$

$$= 2.5 \log_{10}\left[1 - \delta\left\{\frac{\int_0^\infty S(\lambda_0)\left\{E(\lambda_0,t_0) + \lambda_0 E'(\lambda_0,t_0)\right\} d\lambda_0}{\int_0^\infty S(\lambda_0)\, E(\lambda_0,t_0)\, d\lambda_0} + \frac{\frac{1}{\alpha_1}\int_0^\infty S(\lambda_0)\, \dot{E}(\lambda_0,t_0)\, d\lambda_0}{\int_0^\infty S(\lambda_0)\, E(\lambda_0,t_0)\, d\lambda_0}\right\}\right]$$

which to the required order can finally be written, on changing to natural logarithms and expanding the log.,

$$= -1\cdot086\left(K+L\right)\delta \qquad \text{.......................} \quad (1.35)$$

where
$$K = \frac{\int_0^\infty s(\lambda_0)\left\{E(\lambda_0,t_0)+\lambda_0 E'(\lambda_0,t_0)\right\}d\lambda_0}{\int_0^\infty s(\lambda_0)\,E(\lambda_0,t_0)\,d\lambda_0} \qquad \text{.......................} \quad (1.36)$$

and
$$L = \frac{\frac{1}{\alpha_1}\int_0^\infty s(\lambda_0)\,\dot{E}(\lambda_0,t_0)\,d\lambda_0}{\int_0^\infty s(\lambda_0)\,E(\lambda_0,t_0)\,d\lambda_0} \qquad \text{...................} \quad (1.37)$$

The K term arises directly from the selective effect of the redshift and would occur even in the absence of an evolutionary process. The L term would arise solely in the case of an evolutionary universe.

To this approximation in the correction term we can therefore write (1.34) in the form

$$\log_{10} D = \cdot 2\left\{m_{het} - M^0_{het} - 1\cdot086\left(K+L\right)\delta\right\}+1 \qquad \text{..........} \quad (1.38)$$

As we shall show in the next chapter the luminosity distance may also be expanded in powers of δ with coefficients given in terms of the parameters of the model. It will then be indicated how, assuming m_{het}, M^0_{het} and K are known from observations, the parameters of the model and the evolutionary term can be determined.

In contrast the equation (1.27) contains no term that is directly observed. Equation (1.27) is related to an equation given by H. P. Robertson (<u>17)</u> in a short theoretical analysis of observable quantities in cosmology, where he expresses all his results in terms of bolometric magnitudes. For this reason his analysis suffers from the handicap that in terms of it the significance of the observational results is obscure - since the step from measured to bolometric magnitudes is crucial for the determination of the unknowns. Indeed the use of Robertson's equation by Humason, Mayall and Sandage has led to ambiguity in the corrections of the magnitudes as will be shown in Chapter II Section (iv).

An important conclusion to be drawn from our analysis is that if $S(\lambda_0)$ is such that only radiation in a narrow waveband is registered on the detective device, e.g. by the use of filters in colour photometry, so that $S(\lambda_0) = 0$ outside this waveband, then equations (1.36) and (1.37) can be written, with approximation depending on the wavelength and width of waveband used,

$$K = \left\{ E(\lambda_0, t_0) + \lambda_0 \, E'(\lambda_0, t_0) \right\} \Big/ E(\lambda_0, t_0) \qquad \ldots\ldots \; (1.39)$$

$$L = \dot{E}(\lambda_0, t_0) \Big/ \left\{ \alpha_1 \, E(\lambda_0, t_0) \right\} \qquad \ldots\ldots \; (1.40)$$

for the neighbourhood of any given wavelength λ_0. It follows that if λ_0 is chosen so that

$$\left. \begin{aligned} & E(\lambda_0, t_0) + \lambda_0 \, E'(\lambda_0, t_0) = 0 \\ \text{then} \quad & K = 0 \end{aligned} \right\} \qquad \ldots\ldots \; (1.41)$$

so that the K correction due to redshift is then zero. It is clear from (1.39) that this condition is likely to be most nearly satisfied at the red end of the spectrum because of the peaked nature of the light curve.

The advantages of using selected wavelengths and so minimising the K correction have been pointed out by M. F. Walker whose unpublished results of computing the K correction were made available to Humason, Mayall, and Sandage (15). Walker found that the optimum effective wavelengths for minimum K correction for $\delta = 0$ to $\delta = \cdot 3$ were $\lambda_0 = 6300$ (Å.U.) for E nebulae, $\lambda_0 = 6200$ for S_b nebulae and $\lambda_0 = 5500$ for S_c nebulae, for which $K < \cdot 1$ magnitudes. Such a consideration is important for accurate determination of observable relations with a view to fitting any given cosmological model. In the case of the steady state model we can also put $L = 0$ corresponding to the statistical **constancy** of the intensity of the sources. We shall in Chapter III make use of this fact that, for a suitable choice of wavelength, no correction at all is necessary to the heterochromatic magnitude modulus in an exact analysis of the observational results to be expected of the steady state model.

For narrow wavebands for which $K \neq 0$ equations (1.39) and (1.40) will apply. We mention particularly the two colour and six colour photometry by J. Stebbins and A. E. Whitford (11). By comparing the nebular spectrum in the local group with that in the distant clusters they originally found that for the elliptic nebulae (which locally were believed to have a uniform energy curve) there was an excess of red more than double that expected from shifting the local energy curve by an amount corresponding to the redshift. This was tentatively attributed by Whitford to an abundance in the distant clusters of shortlived red supergiants which in the much older local elliptic galaxies have now expired, with no interstellar matter in these galaxies to replace them. Such effects, if substantiated, would provide direct evidence of evolution in the universe.

As reported recently (18) Whitford has now completed examination of the elliptic spectra in four clusters at varying distances by a photoelectric seven colour filter system. After allowing for redshift he now finds the form of the output curves to be intrinsically the same in all clusters, and that the discrepancy originally found in the two colour measurements was spurious and arose from the peculiar properties of the energy curve of M_{32} that was used as a standard.

However it has not perhaps been sufficiently emphasised that the Stebbins Whitford measurements of the spectra are relative measurements between one part of the spectrum and another. An absolute evolutionary change of the output energy which was substantially uniform over the whole spectrum would be undetected by their analysis. In the next chapter we shall show how the use of the number counts, the redshift-magnitude relation, and the Stebbins Whitford type of spectrum analysis can lead to the detection of any systematic evolution of the nebulae that may exist.

Another important application of formulae (1.39), (1.40) is to reception at radio wavelengths where the usual 'observations' are made in narrow wavebands. Methods of utilising the possibilities of this important field of cosmic investigation will also be dealt with in the next chapter.

An analysis of the K correction was given in 1935 by E. Hubble and R. C. Tolman (19) on the assumption that there was no evolutionary effect. Unfortunately their equation which corresponds to (1.38) involves a distance which is not the luminosity distance D so that they have to introduce an extra term depending on δ in the right hand side. However there is now no doubt that D is the distance which is given by the astronomers from their measure of magnitudes.

McVittie's analysis leading to an equation analogous to (1.38), mentioned earlier, obtains the correction term to order δ^2. However his coefficients are evaluated for the epoch of emission so that although he claims to derive a term which he calls the Stebbins Whitford correction, he has not in fact excluded evolutionary effects from his other terms which he seems to regard as the orthodox K correction associated with redshift only. In this thesis we shall show that the second order terms are unnecessary for the solution of the cosmological problem and in any case involve a refinement beyond the means of detection at the present time.

(vii) Distance by apparent size

The luminosity distance D was obtained based on the assumption that the intensity of a light source was inversely proportional to the square of the distance, as in Euclidean geometry. We can also find a distance $\bar{r}$ in terms of which apparent size is related to solid angle as in Euclidean geometry.

Angular measurements of the galaxies have hitherto proved to be technically difficult but have been undertaken again recently in relation to apparent luminosities by W. A. Baum (20).

Suppose we make angular measurements on the celestial sphere in terms of the angles θ and ϕ which recur in the cosmological metric (1.2). Let light be emitted at epoch t from one corner of a rectangle whose angular dimensions are $d\theta$, $d\phi$ the source of light at this corner having coordinates (r, θ, ϕ) in terms of the metric (1.2). The proper area of this rectangle (i.e. as measured by a fundamental observer at the source) will be, at the time of emission,

$$dA = \frac{R^2(t)\, r^2 \sin\theta \, d\theta \, d\phi}{(1+kr^2/4)^2}$$

The solid angle of this area subtended at the origin $r=0$ (i.e. as viewed there by the light from the source) will be

$$d\omega = \sin\theta \, d\theta \, d\phi$$

so that the distance $\bar{r}$, assigned as in Euclidean geometry on the assumption that dA is known to the observer at the origin, will be given by

$$\bar{r}^2 d\omega = dA \qquad \text{............ (1.42)}$$

Hence

$$\bar{r} = \frac{R(t)\, r}{1+kr^2/4} \qquad \text{............ (1.43)}$$

which is the required distance by apparent size in terms of the cosmological parameters.

By comparing (1.43) with (1.23) we see that the relation between the luminosity distance and the distance by apparent size is

$$D = \left\{\frac{R(t_0)}{R(t)}\right\}^2 \bar{r}$$

which by (1.8) can be written

$$D = (1+\delta)^2 \bar{r} \qquad \text{........... (1.44)}$$

Formula (1.43) is well known and was given for instance by W. H. McCrea in 1935 (<u>10</u>). Its application has usually been based on the assumption that the average intrinsic size of a galaxy remained constant with epoch. In this thesis we shall (Chap.II Sect.(vi)) utilise equation (1.43) in an application of the criterion of apparent size particularly as a test for an evolutionary universe.

<u>(viii) Counts of the Nebulae</u>

With the advances now being made in observational technique, the counts of nebulae in the sky corresponding to given limits of apparent magnitude become of increasing theoretical importance. Preliminary to relating these counts to other observable quantities in the next chapter, we shall find here the theoretical number $N(r)$ of nebulae of a given character in the volume of space corresponding to the limits $r=0$ to general r of the cosmological metric (1.2), visible at the time of observation t_0. Although the metric (1.2) corresponds to a homogeneous model contrary to the observed phenomenon of clustering of galaxies we can imagine the region dealt with as so large as to nullify the effect of clustering.

The total locally measured proper volume between coordinate r and $r+dr$, by the discussion in Sect.(ii), is at epoch t

$$4\pi R^3(t)\, r^2 dr \Big/ \left(1+kr^2/4\right)^3$$

If $n(t)$ is the number of the nebulae being counted, per unit proper volume at this epoch, then

$$dN = 4\pi\, n(t)\, R^3(t)\, r^2 dr \Big/ \left(1+kr^2/4\right)^3$$

For the cosmological models of general relativity in which it is assumed galaxies are conserved this number between $r, r+dr$ of the co-moving coordinates must remain statistically constant. Hence

$$n(t)\ R^3(t) = \text{constant} = n_0 R_0^3$$

where suffix 0 indicates values at $t = t_0$. Thus between $r=0$ and general r the number of galaxies is

$$N(r) = 4\pi n_0 R_0^3 \int_0^r \frac{r^2 dr}{(1+kr^2/4)^3} \qquad \text{.............. (1.45)}$$

a well known formula also to be found in the paper by W. H. McCrea (10). We have presented a proof here to emphasise the assumption of conservation and permanency of luminous galaxies as the universe expands. This assumption can be justified by the plausible argument that where there is no creation of matter the intergalaxial density of matter becomes so low due to the expansion that after a certain stage no further condensation into new galaxies is possible; on the other hand existing galaxies are not likely to become non luminous until much later than the present epoch judging from the plentiful supply of gas and dust in most neighbouring (old) galaxies, out of which new stars are still being formed. The elliptic galaxies are an exception to the last remark but even here we can fall back on the argument that all galaxies possess stars with a potential future lifetime comparable to the present age of the universe, as calculated from the Hubble expansion.

It is to be noticed however that the assumption of of conservation would **apply** only to those regions which are not so distant that the epoch of emission of light from them would have to be earlier than the time when galaxies ceased to form. If, as argued above, this time was after the start of the expansion, which is calculated approximately by Hubble's law, such regions would be at a distance which might well come into the range of present instruments, in particular the radio telescopes which seem to have a range considerably greater than the optical (21). Such a sudden decrease in the nebular counts occurring at this range would be direct confirmation of an evolutionary universe. Indeed this is a phenomenon reported in the counts of extra galactic radio sources by M. Ryle and P. Scheuer (22).although its explanation as the limit of the instrument cannot as yet be ruled out (c.f. B. Y. Mills 34). We shall deal further with this matter in Chap. II Section (v).

In contrast with a universe exhibiting systematic evolution we must consider the steady state theory in which there is continuous creation of matter leading to continuous creation of galaxies - according to the authors of the theory, H. Bondi and T. Gold (13), and independently F. Hoyle (14). For this case $n(t)$ is statistically constant, and also the metric is (1.2) for the particular case when $R(t) = e^{t/T}$, T constant, and $k=0$. For the steady state model therefore

$$N(r) = 4\pi n \int_0^r e^{3t/T} r^2 dr \qquad \text{..............} \quad (1.46)$$

where t is the cosmological epoch of emission of the light seen at $r=0$ at the epoch of observation t_0, and n is a constant equal to the average number of nebulae per unit volume. Formula (1.46) is a variant in form of a result given for $N(r)$ by Bondi and Gold (see equation (3.19)). The observable phenomena expected of the steady state will be examined in the next chapter with those of other models, and also separately in Chapter III.

CHAPTER II: RELATIONS BETWEEN OBSERVABLE QUANTITIES IN WORLD MODELS BY THE METHOD OF EXPANSION IN SERIES

(i) Introduction

In Chapter I we obtained expressions for the observable quantities in cosmology in terms of the theoretical parameters r, t, $R(t)$ and k occurring in the metric (1.2). In principle we might eliminate r, and t corresponding to r by equation (1.4), and obtain exact relations between these observational quantities involving the unknown parameters R and k ; these last parameters, which are required to specify completely the metric of class (1.2) relevant to the actual universe, could then be obtained in principle by a comparison with observation. The mathematical difficulty is however prohibitive of this, as well as the fact that exact relations would not lend themselves to practical tests.

The method of expansion in series has been adopted by previous writers, notably G. C. McVittie (23), O. Heckmann (24) and more recently H. P. Robertson (17). This method will also be adopted in this thesis, although for the particular case of the steady state model exact relations will be obtained in Chapter III because of the special character of that model. The work of the present chapter is entirely independent as regards the nature and derivation of the observable relations obtained, unless otherwise stated, and also as regards their interpretation and application to the cosmological problem. Comment on work by previous writers will be made as necessary and the original features derived by the present author will be indicated.

(ii) Expansion in series

It is convenient to replace coordinate r in the cosmological metric (1.2) by means of the substitution $q = R_0 r$, where $R_0 \equiv R(t_0)$ and t_0 is the present epoch of observation; the new coordinate q has then the dimension of distance. We then have

$$ds^2 = c^2 dt^2 - \frac{R^2(t)}{R_0^2} \cdot \frac{(dq^2 + q^2 d\theta^2 + q^2 \sin^2\theta \, d\phi^2)}{(1 + q^2/a^2)^2} \quad \ldots\ldots\ldots (2.1)$$

where $a^2 = 4R_0^2/k$. Thus a^2 is infinite, positive finite, or negative finite according as the 3-space $t =$ constant is flat, of positive curvature, or of negative curvature respectively. Putting $t_0 - t \equiv \tau$ we write

$$\begin{aligned} R(t) &= R(t_0 - \tau) \\ &= R_0 - \tau \dot{R}_0 + \frac{\tau^2}{2} \ddot{R}_0 + \cdots \\ &= R_0\left(1 - \alpha_1 \tau + \frac{\alpha_2}{2} \tau^2 + \cdots\right) \end{aligned} \quad \ldots\ldots(2.2)$$

where $\alpha_1 = \dot{R}_0/R_0$, $\alpha_2 \equiv \ddot{R}_0/R_0$, and so on ....(2.3)

Now the light path from the time of emission t to that of reception t_0 at the origin $q = 0$ is given by

$$c R_0 \int_t^{t_0} \frac{dt}{R(t)} = \int_0^q \frac{dq}{1 + q^2/a^2} \quad \ldots.(2.4)$$

Hence by (2.2), (2.3), (2.4),

$$c \int_0^\tau \frac{d\tau}{1 - \alpha_1 \tau + \frac{\alpha_2}{2}\tau^2 + \cdots} = \int_0^q \left(1 - \frac{q^2}{a^2} + \cdots\right) dq$$

$$\therefore c\tau\left\{1 + \frac{\alpha_1}{2}\tau + \frac{1}{6}(2\alpha_1^2 - \alpha_2)\tau^2 + \cdots\right\} = q - \frac{q^3}{3a^2} + \cdots$$

A first approximation to this relation is

$$q = c\tau \quad \ldots\ldots\ldots\ldots(2.5)$$

A second which will be sufficient for our purposes is

$$q = c\tau\left(1 + \frac{\alpha_1}{2}\tau\right) \quad \ldots\ldots\ldots (2.6)$$

Thus to this approximation the curvature is irrelevant and observational methods are not sufficiently refined to obtain it by taking account of higher more complicated approximations. An independent method of obtaining the curvature by utilising the field equations of general relativity will be given in Chapter IV.

We may now expand the luminosity distance D in powers of τ. By equation (1.23) we may write putting $q = R_0 r$

$$D = \frac{R_0 q}{R(t)(1+q^2/a^2)} \qquad \text{................ (2.7)}$$

Substituting from (2.2) and (2.6) we find on approximation

$$D = c\tau\left(1 + \frac{3\alpha_1}{2}\tau + \cdots\right) \qquad \text{......... (2.8)}$$

so that by (2.5) and (2.8) $q = D$ to a first approximation.

The expression for δ the redshift given by (1.8) is

$$\delta = \frac{R_0}{R(t)} - 1$$

so that on expansion in powers of τ

$$\delta = \alpha_1\tau + \left(\alpha_1^2 - \frac{\alpha_2}{2}\right)\tau^2 + \cdots \qquad \text{...... (2.9)}$$

The distance by apparent size, by putting $q = R_0 r$ in (1.43), has the formula

$$\bar{r} = \frac{R(t)\, q}{R_0(1+q^2/a^2)}$$

so that on expansion we can write to sufficient approximation

$$\bar{r} = c\tau\left(1 - \frac{\alpha_1}{2}\tau + \cdots\right) \qquad \text{......(2.10)}$$

We now turn to the expression $N(r)$ for the number counts of the nebulae. For the models in which there is conservation of galaxies we can write instead of (1.45)

$$N(q) = 4\pi n_0 \int_0^q \frac{q^2\,dq}{(1+q^2/a^2)^3}$$

$$= \frac{4}{3}\pi n_0\left(q^3 - \frac{9q^5}{5a^2} + \cdots\right)$$

so that substituting from (2.6) we have to sufficient approximation

$$N(\tau) = \frac{4}{3}\pi n_0 c^3\left(\tau^3 + \frac{3\alpha_1}{2}\tau^4 + \cdots\right) \qquad \ldots\ldots (2.11)$$

Since $\xi = D = c\tau$ to a first approximation we see that the nebular count can be made locally in any neighbourhood using Euclidean geometry in D viz., for a 'sphere' of radius D a first approximation to N is

$$N(D) \doteq \frac{4}{3}\pi n_0 D^3 \qquad \ldots.(2.12)$$

In the case of the steady state model we have to use equation (1.46). We can evaluate this integral exactly. For an object with coordinate r seen at $r=0$ at time t_0 the light was emitted at time t where, by equation (1.4),

$$c\int_t^{t_0}\frac{dt}{e^{t/T}} = r$$

so that $$cT\left(e^{-t/T} - e^{-t_0/T}\right) = r$$

Hence for fixed time of observation to,

$$dr = -ce^{-t/T}dt$$

so that $$N(t) = 4\pi nc^3T^2\int_t^{t_0} e^{3t/T}\left(e^{-t/T} - e^{-t_0/T}\right)^2 e^{-t/T}dt$$

$$= 4\pi nc^3T^2\int_t^{t_0}\left\{1 - e^{(t-t_0)/T}\right\}^2 dt$$

Thus $$N(\tau) = 4\pi nc^3T^3\left(-\frac{3}{2} + \frac{\tau}{T} + 2e^{-\tau/T} - \frac{1}{2}e^{-2\tau/T}\right) \qquad \ldots. (2.13)$$

where $\tau \equiv t_0 - t$.

To the same approximation as in (2.11) we can therefore write for the steady state model

$$N(\tau) = \frac{4}{3}\pi nc^3\left(\tau^3 - \frac{3}{4T}\tau^4 + \cdots\right) \qquad \ldots. (2.14)$$

By eliminating τ between pairs of these relations we can now find relations between observable quantities which will depend on the model assumed.

(iii) D, δ relation

From (2.9) solving for τ we can write

$$\tau = \frac{\delta}{\alpha_1} + \left(\frac{\alpha_2 - 2\alpha_1^2}{2\alpha_1^3}\right)\delta^2 + \cdots \qquad \text{......(2.15)}$$

which connects redshift with the lapse in cosmological time between emission of the light and its reception. Whence by (2.8) and (2.15)

$$D = \frac{c\delta}{\alpha_1}\left\{1 + \left(\frac{\alpha_1^2 + \alpha_2}{2\alpha_1^2}\right)\delta + \cdots\right\} \qquad \text{......(2.16)}$$

relating luminosity distance to redshift.

(iv) m, δ relation

Equation (1.38) gives the relation between luminosity distance D and the heterochromatic apparent magnitude m_{het}, allowing for the 'correction' term given to order δ. Having pointed out in that place the advantages of using heterochromatic magnitudes directly in the theoretical relations between the observable quantities we shall now drop the suffix to m henceforth, with the understanding that m is the apparent magnitude actually measured by the detecting device, with the appropriate sensitivity function $s(\lambda_0)$ using in the 'correction' terms given by (1.36) and (1.37).

Substituting (2.16) into (1.38) we get a relation between apparent magnitude and redshift, to first order in δ:

$$\log_{10}\left(\frac{c}{\alpha_1}\right) + \log_{10}\delta + \frac{1.086}{5}\left(\frac{\alpha_1^2 + \alpha_2}{\alpha_1^2}\right)\delta = \cdot 2\left\{m - M_0 - 1.086\,(K+L)\,\delta\right\} + 1$$

where we have changed the log to base e and expanded to get the linear term on the left. Rearranging we have

$$m = 5\log_{10}\delta + 1.086\left(1 + \frac{\alpha_2}{\alpha_1^2} + K + L\right)\delta + 5\log_{10}\left(\frac{c}{\alpha_1}\right) + M_0 - 5 \qquad \text{.......(2.17)}$$

The observable data in the form of heterochromatic apparent magnitudes and corresponding redshift can be fitted to this relation and the constant coefficients can be found by least squares. This would yield the constant quantities

$$1 + \frac{\alpha_2}{\alpha_1^2} + K + L \qquad \text{.............} \quad (2.18)$$

and

$$5 \log_{10}\left(\frac{c}{\alpha_1}\right) + M_0 - 5 \qquad \text{.............} \quad (2.19)$$

A knowledge of the average measured (heterochromatic) absolute magnitude M_0 of <u>local</u> nebulae, coupled with the constant of (2.19) given by least squares, would determine the constant $\alpha_1 = \dot{R}_0/R_0$ which is Hubble's constant occurring in the equation (1.24). In fact the redshift law given by (1.24) for the neighbourhood of an observer (δ small) is expressed by (2.17) if we neglect the linear term in δ and use the corresponding approximation in (1.38). This yields, in agreement with (1.24),

$$c\delta \simeq \alpha_1 D \quad (\delta \text{ small}) \qquad \text{.........} \quad (2.20)$$

Substitution of the determined value of α_1 in (2.18) leads to the sign and magnitude of $\alpha_2 = \ddot{R}_0/R_0$, provided that we know $K+L$. We shall show in Section (v) that the quantity $K+L$ may be derived from the number counts, so that if we adopted this value for $K+L$ then the deduced value of $\ddot{R}_0/R_0$ would indicate whether the universe was under accelerating or decelerating expansion at the present time. Alternatively the quantity K defined by (1.36), or for a sufficiently narrow waveband by (1.39), might be determined empirically from studies of local nebulae. The quantity L defined by (1.37), or by

(1.40), depends indirectly on α_1 and α_2 since these quantities serve to characterise the model. In an expanding universe it is reasonable to suppose that $\dot{E}(\lambda,t) \leq 0$ and so $L \leq 0$. If therefore we confined L to a reasonable limit in magnitude we could determine α_2 from the quantity given in (2.18), to within a certain likely error.

The latest observable data on the magnitude - redshift relation have been presented in the paper by Humason, Mayall and Sandage (15) already referred to in Chapter I, Sect.(vi). In this paper they included a brief discussion of the cosmological significance of the results. For this purpose they used an equation, obtained by H. P. Robertson (17), connecting redshift with bolometric magnitude. This necessitated conversion of the measured magnitudes to bolometric ones and the authors included an allowance for evolution on this process. However as Robertson had already allowed for a change in absolute bolometric magnitude with epoch it does not seem clear from their analysis that the correction has been properly made. It would appear to be worth while therefore to apply the analysis of this thesis to the observational results determined by Humason, Mayall and Sandage, as follows.

By the method of least squares Humason, Mayall and Sandage find values for A and B in the relation

$$m = 5 \log_{10}(c\delta) + A\delta + B \qquad \text{......(2.21)}$$

Confining our analysis to the photographic data (suffix P) only, we give their results for the nebulae in the distant clusters, viz.,

$$A_P = -2\cdot 2 \pm \cdot 8 \qquad \text{.........(2.22)}$$

$$B_P = -5\cdot 81 \pm \cdot 09 \qquad \text{.........(2.23)}$$

The magnitudes used in (2.21) are corrected for the latitude obscuration effect of the Galaxy, aperture error, and also for the orthodox redshift correction based on empirical examination of local spectra such as M_{32}. The quantity regarded as the general K correction by Humason, Mayall, and Sandage, as given by their equation B_7, would appear to be the same as our expression (1.33), although the symbols they use are rather loosely defined. This is the same quantity as the minus of our total magnitude correction involved in (1.34) only if evolution is neglected, that is if $E(\lambda, t) = E(\lambda, t_0)$, and for this case what they call K will be equal in our notation to the expression $1\cdot 086\, K\delta$, when approximated to first order in δ. Accordingly by comparing the first and third columns of their table XIII we find that their K correction, assuming no evolution, is effectively in our notation $1\cdot 086\, K_p \delta$ with

$$K_p = 4\cdot 3 \qquad \text{............. (2.24)}$$

Since this orthodox K correction has already been made in the magnitudes used in (2.21) we must accordingly omit K from our expression (2.18). Consequently using (2.22) we must put

$$1 + \frac{\alpha_2}{\alpha_1^2} + L_p = -2 \pm \cdot 8 \qquad \text{.........(2.25)}$$

where L_p is given by (1.37) for photographic wavelengths. This quantity L_p can be written in terms of the rate of change of (heterochromatic) absolute magnitude of nearby nebulae, viz. $\dot{M}_{o_p}$

Thus $$\dot{M}_o = -\frac{d}{dt_o}\left\{2\cdot 5 \log_{10}\int_0^\infty s(\lambda_o)\, E(\lambda_o, t_o)\, d\lambda_o\right\}$$

$$= -2\cdot 5 \log_{10} e \cdot \frac{\int_0^\infty s(\lambda_o)\, \dot{E}(\lambda_o, t_o)\, d\lambda_o}{\int_0^\infty s(\lambda_o)\, E(\lambda_o, t_o)\, d\lambda_o}$$

$$= -2\cdot 5\, \alpha_1 L \log_{10} e .$$

(using (1.37)

Whence $$L_p = -\cdot 92\, \dot{M}_{o_p} / \alpha_1 \qquad \text{.......... (2.26)}$$

Humason, Mayall and Sandage include the expression on the right (multiplied by 1.086), for the case of bolometric magnitude, in their analytic expression for A derived by Robertson. They estimate from the theory of evolution of stars that $\dot{M}_{Bol} < \cdot 3$ mag$/10^9$ yrs. Accordingly we can make the same statement for photographic absolute magnitude, viz.,

$$\dot{M}_{o_p} < \cdot 3 \text{ mag}/10^9 \text{ yrs} \qquad \text{..... (2.27)}$$

Also by investigation of the term B given in our notation by (2.19) they find for the reciprocal of Hubble's constant

$$\frac{1}{\alpha_1} = (5\cdot 4 \pm 1)\, 10^9 \text{ yrs} \qquad \text{.....(2.28)}$$

Hence by (2.26), (2.27), (2.28) we can write

$$L_p > -1\cdot 8 \qquad \text{.....(2.29)}$$

on taking the upper bound for $\frac{1}{\alpha_1}$. Hence by (2.25), (2.29)

$$\frac{\alpha_2}{\alpha_1^2} < -(1\cdot 2 \pm \cdot 8) \qquad \text{.....(2.30)}$$

This result does not agree with that obtained by Humason, Mayall, and Sandage who get $\frac{\alpha_2}{\alpha_1^2} < -(3 \pm \cdot 8)$ from P data.

The discrepancy seems to arise in their evaluation of a term in their expression for A which depends on the evolutionary rate of change of their K correction, viz. $\dot{K}$ It is by no means clear from <u>their</u> definition of K how this term is to be evaluated. Their actual estimate is based on the assumption that the distant spectra of the nebulae will be similar to that of local type I Go supergiants, following Whitford's earlier hypothesis to explain the Stebbins Whitford effect. As has been mentioned in Chapter I Sect. (vi) this effect seems to have been spurious in any case according to Whitford's latest results. The advantage of our analysis is that it does not require the evaluation of such a term.

We can assert therefore that, for the case of an evolutionary universe, our analysis of the observational results of Humason, Mayall and Sandage makes the negative character of $\alpha_2 = \ddot{R}_0/R_0$, deduced from their results, much less certain.

On the other hand it cannot be said that consequently the results are more favourable to the steady state model, for which $\ddot{R}_0/R_0 > 0$. For in the steady state the evolutionary term L is zero and so equation (2.25) would indicate that for that case $\frac{\alpha_2}{\alpha_1^2} = -(3 \pm \cdot 8)$ according to P data. It would follow that if the data of Humason, Mayall, and Sandage are free from systematic errors, such as selective effect, then they indicate that the universe is significantly different from the steady state model.

A systematic effect of selection might arise by favouring galaxies of the distant clusters that are in fact intrinsically brighter than the local galaxies, upsetting the assumption that the term B is a constant depending on the absolute magnitude M_0 of nearby galaxies which are local representatives of the supposedly homogeneous system being analysed. It is clear that this would have the same effect on the A term of (2.21) as an evolutionary effect covered by our L term, but arising in this case from the scatter in absolute magnitude among the nebulae. Consequently such an effect would give a value of α_2/α_1^2 more negative than what would be otherwise deduced. Humason, Mayall, and Sandage attempt to eliminate this effect for the cluster data by letting the order of brightness of the galaxies in the clusters determine the homogeneity of sample.

(v) Relations involving the number counts

For the models in which there is conservation of light sources equation (2.11) gives the number seen at epoch t_0 whose light was emitted at any epoch t such that $0 < t_0 - t \leq \tau$. Taking logarithms we get

$$\log_{10} N(\tau) = \log_{10}\left(\tfrac{4}{3}\pi n_0 c^3\right) + 3\log_{10}\tau + \frac{3\alpha_1}{2}\tau \log_{10} e + \cdots \qquad \ldots\ldots\ (2.31)$$

Using (2.15) we can obtain a relation between the number counts and redshift, to order δ

$$\log_{10} N(\delta) = \log_{10}\left(\frac{4\pi n_0 c^3}{3\alpha_1^3}\right) + 3\log_{10}\delta + \cdot 651\left(\frac{\alpha_2}{\alpha_1^2} - 1\right)\delta + \cdots \qquad \ldots\ldots\ (2.32)$$

From this relation, fitted to the observable data, we may deduce the quantities n_0/α_1^3 and α_2/α_1^2. A knowledge of n_0 which is the average number density (locally) of the sources being counted at epoch t_0 (the present) would yield a check on, the Hubble parameter, obtained otherwise from the magnitude-redshift relation (expression (2.19) where a knowledge of M_0 is required. Vice versa an accurate knowledge of α_1 from the magnitude-redshift data substituted in the value n_0/α_1^3 given from the number counts should give a value of n_0 consistent with observation, and provide a check on the theory of the redshift as arising from a genuine expansion.

The provision of $\alpha_2/\alpha_1^2 \left(= R_0 \ddot{R}_0 / \dot{R}_0^2 \right)$ from the number count theory provides an independent check on the value obtained ~~from~~ for the same quantity from the magnitude-redshift data. It is to be noted that as regards the number counts value there is no uncertainty involved due to the presence of the 'correction' term $K+L$, as in the use of equation (2.17). In fact by means of the number counts we can deduce the value of $K+L$ by substituting the value of α_2/α_1^2 derived from them in the expression (2.18) obtained from the m, δ results. Hence if K is known empirically L can be deduced, and so knowledge would be provided regarding the evolution of the universe.

Another important observable quantity is the gradient of the $\log N$, m relation. This is conveniently obtained in terms of δ by differentiating equations (2.32) and (2.17) with respect to δ. Thus

$$\frac{d(\log_{10} N)}{d\delta} = \log_{10} e \cdot \left\{ \frac{3}{\delta} + \frac{3}{2}\left(\frac{\alpha_2}{\alpha_1^2} - 1\right) + \cdots \right\}$$

$$\frac{dm}{d\delta} = \log_{10} e \cdot \left\{ \frac{5}{\delta} + \frac{5}{2}\left(1 + \frac{\alpha_2}{\alpha_1^2} + K + L\right) + \cdots \right\}$$

so that

$$\frac{d(\log_{10} N)}{dm} = \cdot 6 \left\{ 1 - \left(1 + \frac{K+L}{2}\right)\delta + \cdots \right\}$$

$$\ldots\ldots\ldots\ldots\ldots \quad (2.33)$$

The quantity $K+L$ is provided once again from this relation, this time directly, so that if K is known L is derived. If it could be made technically feasible, of course, the waveband might be chosen so that $K \simeq 0$, as envisaged in connection with equation (1.41), so that the isolation of the evolutionary term would be complete.

Since magnitudes can be measured at ranges where the red-shift cannot it might be convenient to eliminate δ entirely from (2.33) by using the first approximation to (2.17), viz.,

$$\delta \simeq \frac{\alpha_1}{c} 10^{\cdot 2(m - M_0) + 1}$$

so that

$$\frac{d(\log_{10} N)}{dm} = \cdot 6 \left\{ 1 - \left(1 + \frac{K+L}{2}\right) \frac{\alpha_1}{c} 10^{\cdot 2(m - M_0) + 1} + \cdots \right\} \quad \ldots\ldots (2.34)$$

It is believed that equations (2.32), (2.33), (2.34) as presented in this thesis are the first to satisfactorily take account of both the K correction and the evolutionary L correction. In addition the importance of the simple relation (2.32) seems to have been overlooked as regards its ability to provide the fundamental parameter α_2/α_1^2 without being dependent on magnitude measurements and their associated 'corrections', relying as it does only on the number counts and the dependable redshift measurements. The remainder of this discussion on the number count relations is also believed to be presented for the first time; it deals with important applications to radio reception and to the steady state model.

In the form given by (2.34) the relation is suitable for application to radio reception from extra-galactic radio sources. The range of perception by the radio waveband detectors is believed to be considerably greater than that of the present optical instruments, because of the more uniform spectral distribution of energy in the radio waveband with less falling off of energy, in consequence, with the redshift of the spectrum (J.R. Shakeshaft (21)).

As has been mentioned in Chapter I Sect. (vi) the radio reception is generally restricted to narrow wavebands so that the particular form given by (1.39, (1.40) for the K, L terms is highly appropriate.

We see from (2.33) that the local value of the log N, m gradient is equal to .6, that is when $\delta \triangleq 0$. A steeper gradient than this associated with more remote regions of space could arise only if $K+L < -2$. For photographic magnitudes we have inferred that Humason, Mayall, and Sandage adopt $K_p = 4 \cdot 3$ (equation (2.24)). Consequently such a gradient if found would point directly to an evolutionary model with $L_p < -6 \cdot 3$ It will be shown in the next chapter that the non evolutionary steady state model can give only a monotonic decreasing gradient from the local value of .6. This is therefore a criterion for cosmological models.

However as far as the photographic and photovisual regions of the spectrum are concerned a gradient steeper than .6 for remote regions would seem improbable because of the rapid evolutionary trend it demands. We have seen that the estimate of Humason, Mayall, and Sandage based on the theory of stellar evolution has implied, in our notation, $L_p > -1.8$ (equation (2.29)). At radio wavelengths on the other hand a gradient steeper than .6 has been found by M. Ryle and his co-workers for the counts of extra galactic radio 'stars' (22). Ryle has suggested that these radio sources which number so far only in thousands may be colliding galaxies. Whether this gradient phenomenon is ultimately confirmed or not, it is of interest to show that radiation arising from sources generated by a kinetic effect might certainly, in an evolutionary universe, indicate a source number-magnitude gradient higher than .6.

The formula for $N(r)$ given by (1.45) now requires modification for source numbers based on a kinetic effect. If $n(t)$ is the number density of ordinary galaxies defined earlier and a is the cross section for **collisions** of galaxies then a galaxy will make approximately naV collisions per unit time where $V(t)$ is the peculiar velocity of a galaxy relative to the smoothed out substratum of matter at epoch t . Each collision will last a time of order $a^{\frac{1}{2}}/V$ so that the number $\eta(t)$

of collisions occurring per unit volume at any time will be of order $n^2 a^{3/2}$ That is we can put

$$\eta(t) = \beta n^2(t) \quad , \ \beta \ \text{constant}$$

Instead of (1.45) we therefore get

$$N(r) = 4\pi \int_0^r \frac{\eta(t) R^3(t) r^2 dr}{(1 + kr^2/4)^3}$$

$$= 4\pi\beta \int_0^r \frac{n^2(t) R^3(t) r^2 dr}{(1+kr^2/4)^3}$$

Now as before $n(t) = n_0 R_0^3/R^3(t)$ where suffix o indicates the epoch t_0 of observation. Hence

$$N(r) = 4\pi\beta n_0^2 R_0^6 \int_0^r \frac{r^2 dr}{R^3(t)(1+kr^2/4)^3} \quad \text{.......(2.35)}$$

Make now the transformation $q = R_0 r$ used in Sect. (ii) and so

$$N(q) = 4\pi\beta n_0^2 R_0^3 \int_0^q \frac{q^2 dq}{R^3(t)(1+q^2/a^2)^3} \quad \text{.......(2.36)}$$

Using (2.2) and (2.6) we then get on expansion in powers of τ

$$N(\tau) = 4\pi\beta n_0^2 c^3 \left(\frac{\tau^3}{3} + \frac{5}{4}\alpha_1 \tau^4 + \cdots \right)$$

so that by (2.15)

$$\log_{10} N(\delta) = \log_{10}\left(\frac{4\pi\beta n_0^2 c^3}{3\alpha_1^3}\right) + 3\log_{10}\delta + \frac{3}{2}\left(\frac{\alpha_2}{\alpha_1^2} + \frac{1}{2}\right)\delta \log_{10} e + \cdots$$

$$\text{.....(2.37)}$$

The log N, m gradient now becomes, in terms of δ,

$$\frac{d}{dm}\left(\log_{10} N\right) = \frac{\frac{3}{\delta} + \frac{3}{2}\left(\frac{\alpha_2}{\alpha_1^2} + \frac{1}{2}\right) + \cdots}{\frac{5}{\delta} + \frac{5}{2}\left(1 + \frac{\alpha_2}{\alpha_1^2} + K_R + L_R\right) + \cdots}$$

$$= \cdot 6\left\{1 - \left(\frac{1}{4} + \frac{K_R + L_R}{2}\right)\delta + \cdots\right\} \quad \text{.....(2.38)}$$

Here of course K_R is the K term evaluated from equation (1.39) for narrow wavebands of radio emission from a pair of colliding galaxies. The term L_R allows for an evolutionary variation of intensity of the emission arising via the secular change in the peculiar velocities $V(t)$, and is defined by (1.40). We see therefore that the gradient of the $\log N, m$ relation for colliding galaxies would take values greater than .6 if $K_R + L_R < -\frac{1}{2}$ and since for the radio waveband the distribution curve is relatively flat as mentioned earlier, we might expect $K_R \simeq 0$ and so the condition would reduce to $L_R < -\frac{1}{2}$. This implies a much milder demand on evolutionary trends to achieve this than in the case of the galaxy counts at optical wavelengths which demanded $L_P < -1{\cdot}3$. Thus Ryle's phenomenon gains in plausibility, especially when it is realised that evolutionary variation in intensity in this case is much more likely since the random velocities of the galaxies would decrease with expansion.

The appropriate form for the gradient in terms of magnitudes would now be

$$\frac{d}{dm}\left(\log_{10} N\right) = \cdot 6\left\{1 - \left(\frac{1}{4} + \frac{K_R + L_R}{2}\right)\frac{\alpha_1}{c}\, 10^{\cdot 2(m - M_0)+1} + \cdots\right\} \qquad \ldots\ldots(2.39)$$

rather than the equation (2.34) based on a non-kinetic origin. It is understood of course that m and M_0 are the radio magnitudes of the source, apparent and absolute respectively. In this form the relation is applicable to regions where redshift in the optical waveband would be impossible to measure, regions in fact which might be well beyond the limits of the optical instruments. The cosmic epoch of the events recorded would therefore be very early, perhaps earlier than the epoch before which the assumption of conservation of galaxial numbers would be inapplicable, corresponding to regions where galaxies were still being formed or even to regions where no galaxies had yet formed. Indeed as mentioned in Chapter I (Sect.(viii) the graph of log N against log I (Intensity) obtained by Ryle and Scheuer (22) for radio stars might indicate such a phenomenon since it ultimately bends over and flattens out.

As far as the approximate theory is concerned, with which we are now dealing, we shall now consider the results of the number counts expected of the steady state theory. Equation (2.14) is now applicable and to eliminate τ we use (2.15) for the case of the steady state which is

$$\frac{\tau}{T} = \delta - \frac{1}{2}\delta^2 + \cdots \qquad \text{..........} \; (2.40)$$

where T is a constant which is the reciprocal of the Hubble parameter α_1 for the model. Hence (2.14) becomes

$$N(\delta) = \frac{4}{3}\pi n c^3 T^3 \left(\delta^3 - \frac{9}{4}\delta^4 + \cdots \right)$$

so that

$$\log_{10} N(\delta) = \log_{10}\left(\frac{4}{3}\pi n c^3 T^3\right) + 3\log_{10}\delta - \cdot 977\,\delta + \cdots \qquad \text{.....} (2.41)$$

This important result makes definite demands of the observable data for the steady state model. The first term gives the product nT^3 where n is the statistically constant number density of galaxies (as measured locally) and T is the reciprocal of the Hubble parameter. The third term gives a definite coefficient of δ which would have to be realised by the steady state for suitable δ (δ^2 sufficiently small).

On comparing with the corresponding result for the conservation models given by (2.32) we see that a coefficient of δ of the same sign and magnitude in (2.32) would require $\alpha_2 < 0$ whereas for the steady state $\alpha_2 \left(= 1/T^2\right) > 0$. The two possibilities would therefore be distinguished by appealing to the m, δ results, corresponding to the equation (2.17), which would establish the sign of α_2 if we assume that $L \leq 0$, K is known, and use the criterion established from the number counts. Vice versa if the steady state were not ruled out by the results of the m, δ relation (which occurs when, assuming that K is known and $L \leq 0$ as before, the expression (2.18) yields $\frac{\alpha_2}{\alpha_1^2} + L > 0$) it would be distinguished from an evolutionary model by checking the coefficient of δ in the number counts. In the case of an evolutionary model this coefficient would be $> -\cdot 651$ under the supposed condition, since it implies $\alpha_2 > 0$. For the steady state the coefficient is $-\cdot 977$.

By differentiation of (2.41) and (2.17) we get for the gradient of the log N, m relation

$$\frac{d}{dm}\left(\log_{10} N\right) = \frac{\frac{3}{\delta} - \frac{9}{4} + \cdots}{\frac{5}{\delta} + \frac{5}{2}\left(2 + K\right) + \cdots}$$

$$= \cdot 6 \left\{ 1 - \left(\frac{7}{4} + \frac{K}{2}\right)\delta + \cdots \right\} \qquad \ldots\ldots\ (2.42)$$

where we have put $L = 0$ corresponding to the absence of general evolution in the steady state, and put $\alpha_2/\alpha_1^2 = 1$ for that model.

Here again an important difference arises between the steady state and the models for which there is conservation of galaxies - for which the appropriate equation is (2.33). As has been remarked in Sect. (iv) we can expect $L \leq 0$ in an expanding universe, so that whatever be the value of K ($< 4\cdot5$ at most) a sharp difference arises in the coefficients of δ in the two cases. This would be especially noticeable if the nebulae were photographed in red light at wavelengths where $K \doteq 0$ (equation (1.41)) when the ratio of the coefficients would be $\geqslant \frac{7}{4}$ if positive, and might even be negative.

Putting as a first approximation to δ as given by (2.17)

$$\delta \doteq \frac{10^{\cdot2(m-M_0)+1}}{cT}$$

and substituting in (2.42) we have the log N, m gradient in terms of m for the steady state

$$\frac{d}{dm}\left(\log_{10} N\right) = \cdot6\left\{1 - \left(\frac{7}{4} + \frac{K}{2}\right)\frac{10^{\cdot2(m-M_0)+1}}{cT} + \cdots\right\} \qquad \ldots (2.43)$$

It is to be noticed that this formula applies also to the counts of radio sources in the steady state, since for that model there is no mathematical difference between the discussion of a uniform sub-group of the galaxies and that of the galaxies themselves. The difference therefore between the formula (2.43) for radio sources in the steady state and the corresponding formula (2.39) for radio sources in the conservation models provides scope for a decisive part to be played by radio astronomers in the solution of the cosmological problem. For even if the gradient does not exceed .6 in the case of the evolutionary models, as found by Ryle and Scheuer, the coefficient of $\frac{H}{c} \cdot 10^{\cdot 2(m-M_0)+1}$ $(=\delta)$, where H is Hubble's constant, is likely to be significantly different in the two cases. Taking $K_R \doteq 0$ for radio waves and $L_R < 0$ the ratio of the coefficients is at least equal to 7.

We shall end this discussion of the number counts with a reference to systematic error due to selection. This is a grave difficulty in the case of the number counts since at the distances where the counts become important to establish a second order term (such as the term in δ in (2.32)) the less bright galaxies may escape detection. The effect of this would be to lower the adjustable coefficients in the observable relation, which include the second order term of crucial importance.

In the case of the relation (2.32) for the conservation models we see that this leads to a lower estimate of α_2/α_1^2 than the true value. However the distinction between a conservation model and the steady state model seems at the very least a possibility because of the pronounced difference in the number count formulae in the two cases.

A method which we suggest in this thesis to eliminate selective effect in the number counts is as follows.

We have pointed out the immense value of the log N, δ relation given by (2.32), and its counterpart for the steady state (2.41), which might justify the laborious measurements of redshift to the same extent as the counts. According to the m, δ relation (2.17) the redshift observations provide values of A and B in the relation $m = 5 \log_{10} \delta + A\delta + B$. By establishing this particular relation from galaxies of the same order of brightness in each cluster (5th brightest for example) a homogeneous sample of galaxies is obtained which frees this relation from selective error. A uniform sample of galaxies would now also be obtained for the number counts if only those galaxies were counted whose magnitudes never exceeded by more than a suitable fixed amount a magnitude whose dependence on the redshift would be according to the relation given above. This relation would be based on the average standard magnitude M_0 of the 5th brightest galaxy in local clusters, a possible evolutionary change of M_0 being taken account of by the A term (equation (2.17)). The fixed magnitude range would of course be arranged to include nearly all the visible galaxies of the most distant clusters but would exclude many nearer galaxies of low intrinsic luminosity.

(vi) Measurements of angular diameter

Measurements of the angular diameters of the galaxies were made long ago by E. Hubble (25) but no information of value was obtained, partly because the analysis was based on Euclidean geometry in a static universe and also because of the technical difficulty of determining a diameter which did not depend on the time of photographic exposure. Observations have been made again recently by W. A. Baum (20). A suitable empirical definition of the diameter such as is specified by locating its extremities at points where the brightness falls to a certain fraction of that at the centre, coupled with improved instrumental technique may make this an important criterion for the cosmological problem. The measurement of the angular diameters of clusters of galaxies might also lead to results of theoretical importance.

The theoretical formulae established in the paper by McCrea (10), already quoted in Chapter I Sect. (vii), for the measurements of angular diameter were based on the assumption of a constant intrinsic linear diameter of the galaxy independent of epoch, and a constant absolute luminosity. The formulae were in fact suggested as a basis for an observational check on the validity of the metric (1.2) of the expanding universe and in particular of the Doppler effect derived therefrom. Bolometric luminosities were used so that there was no incorporation of the K correction, or of course of the evolutionary L correction, which we shall include in our analysis here. A similar analysis to that of McCrea, but with considerably more elaboration, was included in a paper by E. Hubble and R.C.Tolman (19) later in 1935. This included an analysis of the K term, but otherwise the assumptions were the same as in McCrea's paper. In addition, as already mentioned in Chapter ~~III~~ I Sect. (vi), their formulae incorporate a distance which is neither the luminosity distance nor the distance by apparent size.

There is a need therefore for an analysis which does not make the assumptions mentioned above, but is in fact directed towards providing a basis for detecting the evolution, if any, of the universe. The paper by H. P. Robertson (17) in 1955 contains only a reproduction of the previous work as regards angular diameters, while in his recent book the analysis by G. C. McVittie (16) seems to be based on inferential error and is logically unsatisfactory. The following analysis by the present author is consonant with the other formulae of this thesis.

If d is the proper linear diameter of a galaxy, as it would be measured by an observer on it at the time of emission, then the analysis of Chapter I Sect. (vii) implies that we can write (c.f.(1.42))

$$d = \bar{r}\,\Delta\theta \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2.44)$$

where $\Delta\theta$ is the angular diameter measured by the observer at the spatial origin of the coordinates of the metric (1.2), using the emitted light, while $\bar{r}$ is the distance by apparent size introduced in (1.42). (However this distance would only be arrived at by the observer, using Euclidean geometry calculations, if d was the same as the diameter d_0 of galaxies local to the observer at the origin, i.e. if there was no evolutionary change). The expansion of $\bar{r}$ in powers of τ is given by (2.10), so that eliminating τ by (2.15) we get

$$\bar{r} = \frac{c\delta}{\alpha_1}\left\{1 - \frac{1}{2}\left(3 - \frac{\alpha_2}{\alpha_1^2}\right)\delta + \cdots\right\} \qquad \ldots\ldots(2.45)$$

Whence by (2.44)

$$d = \frac{c\Delta\theta\delta}{\alpha_1}\left\{1 - \frac{1}{2}\left(3 - \frac{\alpha_2}{\alpha_1^2}\right)\delta + \cdots\right\} \qquad \ldots\ldots(2.46)$$

On taking logarithms we have on expansion to order δ,

$$\log_{10}\Delta\theta + \log_{10}\delta - \cdot 217\left(3 - \frac{\alpha_2}{\alpha_1^2}\right)\delta = \log_{10}\left(\frac{\alpha_1 d}{c}\right) \qquad \ldots\ldots(2.47)$$

The quantity α_2/α_1^2 is known from either the m, δ observations based on (2.17) (assuming K and L are known) or directly from the number counts, via say (2.32). The systematic variation of the quantity on the left of (2.47) would therefore indicate the evolutionary change in intrinsic linear diameter of the galaxies. For the case of the steady state in which d must be statistically constant, putting $\alpha_2/\alpha_1^2 = 1$ we have the important relation for observable data

$$\log_{10}\Delta\theta + \log_{10}\delta - \cdot 434\delta = \text{Constant} \qquad \ldots\ldots\ldots(2.48)$$

The quantities measured by Baum seem to be angular diameter $\Delta\theta$ and luminosity (or equivalently apparent magnitude m). To deal with this the only satisfactory procedure seems to be as follows. Equation (1.44) gives the relation between luminosity distance and distance by apparent size. Eliminating τ by (2.44) we get,

$$D = \frac{(1+\delta)^2 d}{\Delta\theta} \qquad \ldots\ldots\ldots\ldots(2.49)$$

so that on taking logs and using (1.38) we derive

$$\cdot 2\left\{m - M_0 - 1{\cdot}086(K+L)\delta\right\} + 1 = \log_{10} d + 2\log_{10}(1+\delta) - \log_{10}\Delta\theta$$

Expanding the binomial log we write to order δ

$$\log_{10}\Delta\theta + \cdot 2\left\{m - M_0 - 1{\cdot}086(K+L+4)\delta\right\} = \log_{10} d - 1 \qquad \ldots\ldots(2.50)$$

We have shown that L or $K+L$ may be derived from the number counts.

Substitution in this equation would therefore permit the detection of a systematic variation in d . Alternatively we can transpose the equation and put all the terms which depend on evolution on the right, thus

$$\log_{10}\Delta\theta + \cdot 2\left\{m - M_0 - 1\cdot 086\,(K+4)\delta\right\} = \log_{10} d + \cdot 217\,L\delta - 1 \qquad \ldots.(2.51)$$

In the absence of evolution then $L=0, d=$ constant, so that the left hand side would be constant. We can eliminate the necessity for redshift measurement by substituting for the first approximation to (2.17), viz. $\delta \doteq \left(\frac{\alpha_1}{c}\right)10^{\cdot 2(m-M_0)+1}$ and derive the result directly applicable to Baum's measurements.

$$\log_{10}\Delta\theta + \cdot 2\left\{m - M_0 - \left(\frac{1\cdot 086\,\alpha_1}{c}\right)(K+4)\,10^{\cdot 2(m-M_0)+1}\right\} = \log_{10} d + \left(\frac{\cdot 217\,L\alpha_1}{c}\right)10^{\cdot 2(m-M_0)+1} - 1 \qquad \ldots\ldots(2.52)$$

so that in the absence of detectable evolutionary effect, which must certainly hold in the case of the steady state theory the following observational relation must be satisfied

$$\log_{10}\Delta\theta + \cdot 2\left\{m - M_0 - \left(\frac{1\cdot 086\,\alpha_1}{c}\right)(K+4)\,10^{\cdot 2(m-M_0)+1}\right\} = \text{constant}. \qquad \ldots.(2.53)$$

α_1 would be known from the Hubble expansion, while K could be derived empirically from local galaxies according to (1.36).

It is possible, though unlikely, that the variations in the two evolutionary terms on the right of (2.52) might cancel out. In this case to distinguish from the steady state we must appeal to (2.47).

The analysis by G. C. McVittie, already referred to, is based on a quantity $\mathcal{J}$ which in our notation might be written $\int_0^\infty S(\lambda_0)\,E\left(\frac{\lambda_0}{1+\delta}, t\right)d\lambda_0$ which he calls the total energy of the source. However this quantity depends on δ even if we ignore the presence of the sensitivity function. He then gives a relation involving angular diameters, luminosities, and the quantity g for two different sources and concludes that if the g's calculated therefrom are different then this gives information as to evolution. This is not so since they would be in general unequal due to the dependence on δ explicitly, regardless of any additional evolutionary effect.

(vii) The measurements of Stebbins and Whitford

The nature of the measurements of the nebular spectra by Stebbins and Whitford (11, 18) has been described in Chapter I Sect. (vi). The theory of their important analysis, which follows logically from the previous theory in this thesis, is as follows.

Suppose, by means of filtering, the radiation from a distant galaxy is confined to a narrow wave band for which the effective wavelength of received energy is λ_0. Let the corresponding heterochromatic magnitude be m, and suppose that heterochromatic absolute magnitude of a local galaxy of the same type, corresponding to the same wave band is M_0. Then equation (1.38) applies with K and L given for narrow wavebands by (1.39), 1.40). Let all these quantities, with dashes, refer to another narrow waveband elsewhere in the received spectrum. Hence we can write

$$\log_{10} D = \cdot 2 \left\{ m - M_0 - 1\cdot 086 (K+L)\delta \right\} + 1$$

$$\log_{10} D = \cdot 2 \left\{ m' - M_0' - 1\cdot 086 (K'+L')\delta \right\} + 1$$

to order δ, so that on subtraction we get

$$m - m' = M_0 - M_0' + 1\cdot086\left\{(K - K') + (L - L')\right\}\delta \qquad \ldots\ldots\ldots \quad (2.54)$$

If C is the colour index for these two wavebands in the case of the distant galaxy, and C_0 is the same index for the local galaxy (or other suitable local standard) then

$$C = C_0 + 1\cdot086\left\{(K - K') + (L - L')\right\}\delta \qquad \ldots\ldots(2.55)$$

Suppose, in the first instance, that $L = L' = 0$ so that by equation (1.40) there is no secular variation in absolute luminosity, then provided $K \neq K'$ we have $C \neq C_0$. The difference in the colour indices for near and distant galaxies of the same type arises in this case from the differential effect of the redshift on the received spectrum. This is a quantity which has to be allowed for in the investigation of a possible evolutionary effect. K and K' will of course be known from an analysis of a local spectrum of the same type of nebula and an empirical knowledge of the sensitivity function for the combined effects of atmosphere and photographic plate, applied to equation (1.39).

An additional effect on the colour index, arising from a differential effect with respect to wavelength of a systematic evolutionary change in the spectrum, appears if $L - L' \neq 0$ Equation (2.55) determines $L - L'$ for different pairs of wavebands. As has been mentioned in Chapter I Sect. (vi), in their original two colour measurements in the blue and yellow Stebbins and Whitford found relatively excess radiation in the yellow for elliptic nebulae. As also stated in Chapter I, Whitford (18), by seven colour photoelectric analysis now finds no significant difference in the different elliptic spectra at varying distances. The original effect has been traced to the sharp dip in the energy curve of the elliptic galaxy M32, which was used as a standard, in the neighbourhood of the violet end of the spectrum.

However we emphasise again here that measurements of this type provide only relative measures of the energy in different colours. Whitford's latest results would therefore apparently rule out relative evolutionary trends in different parts of the spectrum. Until we possess a knowledge of the absolute value of L for one colour of the spectrum however we cannot say that the total energy is not changing with epoch. A knowledge of L can be provided from the number counts, or a combination of the number counts and redshift data, as we have seen earlier in this chapter. For this purpose the counts would have to be made of a particular type of galaxy (elliptic say) and photographed in a definite colour, with the same procedure for the redshift measurements. In the case of the use of the relation (2.32) to obtain α_2/α_1^2 this limitation would of course not be necessary for the number counts. Substitution into the magnitude-redshift relation, for the particular type of galaxy at the particular wavelength, of the value for α_2/α_1^2 derived from the number counts would provide the absolute value of L for that type of galaxy at the particular wavelength.

For the particular case of the steady state we must have on the average $L = L' = 0$ and so the Stebbins Whitford type of measurements may provide an important observable criterion for the solution of the cosmological problem, if they are used in conjunction with the other observable relations.

(viii) Tests on the nature of the redshift

As mentioned in connection with the formula (1.7) for the Doppler redshift, arising in an expanding universe covered by our general metric (1.2), alternative explanations as to the origin of the observed redshift have been given in the past. Denial of the Doppler origin of the redshift logically implies a static universe, for which a mechanism must be found to exhibit a redshift proportional to distance.

In looking for tests to establish the Doppler effect we need not concern ourselves with those static universes, achieved as by E. A. Milne (26) by the artifice of a time transformation on an expanding universe in which the frequencies of the characteristic atomic emissions increase with epoch, since it would be just as legitimate to choose our clock mechanism appropriate to the expanding model with the usual assumption that the frequencies of atomic vibrations are independent of epoch, i.e. we can postulate atomic clocks. There remain those hypotheses such as that by F. Zwicky (27) by which for some unknown reason the energy of a photon becomes degraded with the passage of time between emission and reception (in a manner proportional to distance), and perhaps the type of projective universe recently proposed by E. Holmberg (28).

Tests based on the number counts and apparent diameter of galaxies were proposed by E. Hubble and R.C. Tolman (19) in 1935. For the purposes of comparison they adopted the Einstein universe as the static universe. This is rather unsatisfactory since on the usual interpretation the Einstein universe provides no apparent mechanism for redshift whatever and was in any case proved by Eddington to be unstable. Furthermore the angular diameter test was based on the assumption of constant intrinsic diameter, and the number counts on constant absolute magnitude, which in a static universe is unlikely since it denies the possibility of evolution towards a state of thermodynamic equilibrium.

Lacking static universes with concrete properties to provide a basis for an alternative theory (which of course strengthens the position of the expanding universe theory) perhaps the best type of approach is a suggestion by W.H.McCrea (10) that any periodic phenomenon such as the luminosity of the Cepheids must exhibit the Doppler effect if it is genuine. It would be hard to imagine an alternative theory of the redshift which could affect such a different phenomenon in the same way as light itself.

Unfortunately the Cepheids cannot be detected in regions of significant redshift. A most promising alternative has been suggested and analysed recently by S.N.Milford (29). He suggests the study of the very bright type I supernova for which the consistent nature of the light curve has been established. Milford however makes assumptions in his derivation which are incorrect, arising from the fact that in his analysis he effectively works in terms of bolometric magnitudes and does not allow for the selective effects arising from redshift. Although Milford makes reference to the K correction which would actually have to be applied we shall show that it is not in fact the usual K correction that is involved. Consequently Milford's result is incorrect.

According to equation (1.29) we can express the measured (heterochromatic) luminosity of a distant supernova in the form

$$L^*_{het} = \left\{ \frac{1}{1+\delta} \int_0^\infty s(\lambda_0)\, E\left(\frac{\lambda_0}{1+\delta}, t\right) d\lambda_0 \right\} \Big/ 4\pi D^2 \qquad \text{....(2.56)}$$

where all the quantities have the interpretations obvious from the analysis in connection with (1.29); in particular t is the cosmological epoch of emission, that is as would be registered on the clock of an observer in the neighbourhood. By definition of magnitude the apparent (heterochromatic) magnitude registered is

$$m = -2.5 \log_{10} L^*_{het} \quad + \quad \text{constant} \qquad \text{........(2.57)}$$

Now $\int_0^\infty s(\lambda_0)\, E\left(\frac{\lambda_0}{1+\delta}, t\right) d\lambda_0$

$$= \int_0^\infty s(\lambda_0) \left\{ E(\lambda_0, t) - \delta \lambda_0\, E'(\lambda_0, t) + \cdots \right\} d\lambda_0$$

$$= \int_0^\infty s(\lambda_0)\, E(\lambda_0, t)\, d\lambda_0 - \delta \int_0^\infty s(\lambda_0)\, \lambda_0\, E'(\lambda_0, t)\, d\lambda_0 + \cdots$$

where $E'(\lambda_0, t) = \frac{\partial}{\partial \lambda_0} \left\{ E(\lambda_0, t) \right\}_{t \text{ constant}}$

Hence $m = -2.5 \log_{10} \left\{ \int_0^\infty s(\lambda_0)\, E(\lambda_0, t)\, d\lambda_0 - \delta \int_0^\infty s(\lambda_0)\, \lambda_0\, E'(\lambda_0, t)\, d\lambda_0 + \cdots \right\}$

$$+ 2.5 \log_{10} \left\{ (1+\delta) D^2 \right\} + \text{constant}. \qquad \text{.......(2.58)}$$

We now differentiate m with respect to the time of observation t_0, noting that for a supernova explosion the variation of magnitude takes place in a period during which the change in δ or D is utterly negligible. Hence to order δ we can write

$$\frac{dm}{dt_0} = -2.5 \log_{10} e \cdot \frac{dt}{dt_0} \left\{ \frac{\int_0^\infty s(\lambda_0)\, \dot{E}(\lambda_0,t)\, d\lambda_0 - \delta \int_0^\infty s(\lambda_0)\, \lambda_0\, \dot{E}'(\lambda_0,t)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, E(\lambda_0,t)\, d\lambda_0 - \delta \int_0^\infty s(\lambda_0)\, \lambda_0\, E'(\lambda_0,t)\, d\lambda_0} \right\}$$

where
$$\dot{E}(\lambda_0,t) = \frac{\partial}{\partial t}\left\{E(\lambda_0,t)\right\}$$
$$\dot{E}'(\lambda_0,t) = \frac{\partial}{\partial t}\left\{E'(\lambda_0,t)\right\}$$

By (1.5), (1.8) we have $\frac{dt}{dt_0} = \frac{1}{1+\delta}$. Hence to order δ

$$\frac{dm}{dt_0} = -2.5 \log_{10} e \cdot \frac{f(\delta,t)}{1+\delta} \cdot \frac{\int_0^\infty s(\lambda_0)\, \dot{E}(\lambda_0,t)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, E(\lambda_0,t)\, d\lambda_0} \qquad \text{........(2.59)}$$

where
$$f(\delta,t) = \frac{1 - \delta \left\{ \dfrac{\int_0^\infty s(\lambda_0)\, \lambda_0\, \dot{E}'(\lambda_0,t)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, \dot{E}(\lambda_0,t)\, d\lambda_0} \right\}}{1 - \delta \left\{ \dfrac{\int_0^\infty s(\lambda_0)\, \lambda_0\, E'(\lambda_0,t)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, E(\lambda_0,t)\, d\lambda_0} \right\}} \qquad \text{........(2.60)}$$

We now consider a supernova near the observer for which δ is effectively zero.

Let suffix star $*$ apply to this case. Hence

$$\frac{dm}{dt_0} = \frac{dm_*}{dt_0} \cdot \frac{f(\delta,t)}{1+\delta} \cdot \left\{ \frac{\dfrac{\int_0^\infty s(\lambda_0)\, \dot{E}(\lambda_0,t)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, E(\lambda_0,t)\, d\lambda_0}}{\dfrac{\int_0^\infty s(\lambda_0)\, \dot{E}_*(\lambda_0,t_*)\, d\lambda_0}{\int_0^\infty s(\lambda_0)\, E_*(\lambda_0,t_*)\, d\lambda_0}} \right\} \qquad \text{......(2.61)}$$

Since Milford does not include the sensitivity function $s(\lambda_0)$ in his definition of absolute luminosity, so that he effectively takes $s(\lambda_0)=1$, he is able to put the expression in curly brackets equal to unity by identifying the phase of the local light curve corresponding to t_* with that of the distant supernova at phase corresponding to t, without necessarily assuming that the absolute luminosities at corresponding phases are equal. We can only do this clearly if $s(\lambda_0)=1$ which is certainly not the case in practice. However if we examine the supernova in a sufficiently narrow wave band (a particular colour filter) then we need only consider the expressions in curly brackets in the small range $d\lambda_0$ at the particular wavelength λ_0 so that the total factor approximates to unity as Milford assumes.

We note however that even on this assumption the expression given by (2.60) does not become unity. In fact our final result is, for sufficiently narrow waveband,

$$\frac{dm}{dt_0} = \frac{f(\delta,t)}{1+\delta} \cdot \frac{dm_*}{dt_0} \qquad \text{..........(2.62)}$$

Thus the times between corresponding phases are not exactly in the Doppler ratio $1 : 1+\delta$ as found by Milford. However a systematic analysis of the supernova light curves could no doubt yield limits for $f(\delta,t)$ in the neighbourhood of unity so that this test may indeed provide a satisfactory verification of the Doppler effect.

In conclusion we remark that the concept of an expanding universe would receive its best support if the observable relations, interpreted in terms of the metric (1.2), can establish the parameters of a model of that metric consistent with these relations.

CHAPTER III: EXACT OBSERVABLE RELATIONS FOR THE STEADY STATE MODEL ASSUMING $K = 0$.

(i) Introduction

As we have already pointed out, because of the statistical constancy of the absolute magnitude of the galaxies in the steady state, the evolutionary L 'correction' given by equation (1.37) vanishes for that model. We have shown also in connection with equation (1.41) that, by a suitable choice of waveband (in the red), the K correction may be taken to be zero as well. We shall make use of this possibility to obtain more insight into the observable features to be expected of the steady state, by deriving exact relations where in some cases it would not otherwise be feasible. The work is original unless otherwise stated.

(ii) Luminosity distance, redshift and apparent magnitude

The luminosity distance for the steady state is got by putting $R(t) = e^{t/T}$ and $k = 0$ in equation (1.23). Thus

$$D = r\, e^{(2t_0 - t)/T}$$

or putting $t_0 - t \equiv \tau$

$$D = r\, e^{(t_0 + \tau)/T} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

The relation of r to t is got for the steady state from (1.4) in the form

$$c\int_t^{t_0} e^{-t/T}\, dt = r$$

yielding $\quad CT\left(e^{-t/T} - e^{-t_0/T}\right) = r \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2)$

Whence

$$D = cT\left(e^{2\tau/T} - e^{\tau/T}\right) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

This connects luminosity distance to cosmological time lapse τ between emission and reception, although neither of these quantities is directly observed.

For the redshift δ equation (1.8) yields for the steady state

$$\delta = e^{\tau/T} - 1 \qquad \text{.........................(3.4)}$$

By (3.3) and 3.4) we have the D, δ relation

$$D = cT\delta(1+\delta) \qquad \text{.........................(3.5)}$$

We shall formally define the velocity of recession in terms of the redshift according to the classical Doppler relation, viz.,

$$V = c\delta \qquad \text{.........................(3.6)}$$

Since by (3.5) when δ is small $D \doteq cT\delta$ we have

$$V \doteq \frac{D}{T} \qquad (\delta \text{ small}) \quad \text{.............(3.7)}$$

For small δ, V will be the usual velocity and, as we have shown in Chapter I, D is then the usual distance by measuring rod or trigonometric parallax, so that (3.7) represents the Hubble law for the steady state in the neighbourhood of the observer. Thus $1/T$ is the Hubble parameter in agreement with the general result given by equation (1.24).

The results published by Humason, Mayall and Sandage (<u>15</u>) indicate that we must take at present for the reciprocal of the Hubble parameter

$$T = 5\cdot 4 \times 10^9 \text{ yrs} \qquad \text{......................... (3.8)}$$

with a possible error of 20%. An error of this order can arise even from the assumption of linearity in the Hubble law out to the distance of the most remote clusters so far observed. For by (3.5), (3.6) the general relation between V and D is

$$V = \frac{D}{T(1+V/c)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.9)$$

The most distant spectra measured for redshift so far correspond to $\delta \simeq \cdot 2$, which by (3.6) means a velocity of recession equal to $\cdot 2c$ or one fifth of the velocity of light. By (3.9) this means the Hubble 'constant' is 17% less than the local value. By (3.5) we note that for this redshift the distance obtained by astronomers would be $1{\cdot}3 \times 10^9$ lt yrs.

The relation of luminosity distance to apparent magnitude is especially simple for the steady state, assuming that the waveband is chosen so that $K = 0$. By (1.38) this relation is

$$\log_{10} D = \cdot 2(m - M_0) + 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.10)$$

where m and M_0 are the apparent and absolute heterochromatic magnitudes for this waveband of sensitivity. It follows from (3.10) and (3.5) that the exact magnitude-redshift relation is

$$m = 5\log_{10}\left\{\delta(1+\delta)\right\} + 5\log_{10}(cT) + M_0 - 5 \quad \ldots\ldots\ldots\ldots (3.11)$$

which would have to be satisfied by the steady state when $K = 0$.

A unique observable feature of the steady state model is that over a sufficiently long period new galaxies can be seen to form in the spaces between the dispersing galaxies, due to continual creation. Such creation events in other cosmological models of general relativity are confined to the precincts of so called observational horizons which exist in some of those models, e.g. the Einstein - de Sitter universe. In this latter universe the observational horizon of every observer advances through matter with a speed which is locally that of light, bringing into view (provided it is luminous at that epoch) matter newly created in the experience of the observer. In the steady state these events can occur anywhere in intergalaxial space. There are however observational limits in the steady state and these we shall now investigate.

Equation (3.2) connects the time of emission t of light from a particle of fixed coordinate r with the time t_0 of its reception at the space origin. We can write this equation in the form

$$\ell_t = cT\left(1 - e^{(t-t_0)/T}\right) \qquad \text{.................} \quad (3.12)$$

where $\ell_t = r e^{t/T}$ is by (1.10) the integrated proper distance from the origin at the time of emission. We see that when increases the lapse of time $\tau \equiv t_0 - t$ between emission and reception steadily increases until it is infinite when $\ell_t = cT$. Thus nothing beyond this distance, <u>at the time of emission,</u> can ever be seen by the observer. The horizon in this sense which exists in the steady state is what has been defined as an 'event horizon' by W. Rindler (<u>30</u>) in a recent investigation of visual horizons in world models. Since this horizon is at constant proper distance from the observer in the steady state model the matter of the universe steadily passes beyond this horizon in the expansion of the model. However it is important to notice that the observer never witnesses this - no particle ever disappears from view in his finite experience; this we see as follows.

Suppose a particle 'r' is seen at finite time t_0 then the time of emission t must be given by (3.2), and $re^{t/T}$ will be $< cT$. As t_0 increases to $+\infty$ there is always a finite solution for t from (3.2) for constant r, and in the limit as $t_0 \rightarrow +\infty, re^{t/T} \rightarrow cT$ Thus the particle once seen never disappears. Again from (3.2) we see that for a finite time of observation t_0 when $t \rightarrow -\infty$ then $r \rightarrow +\infty$. Thus the whole range of particles which can be seen at time t_0 is covered by the infinite range of r This observed population is in fact infinite as we shall show later. Thus there is no anomaly arising from the fact that although no galaxy disappears from view and galaxies are continually being created, there is nevertheless a steady state, since the visible population is in fact infinite.

On the other hand there is a practical limit to the number of galaxies which can be observed in the steady state depending on the sensitivity of the instruments of detection. For by (3.4) when the lapse of time $\tau \rightarrow \infty$ we see that the Doppler redshift also becomes infinite. The astronomical distance D also tends to infinity theoretically but is correspondingly limited in practice.

(iii) Number counts

The exact value of the number of visible galaxies corresponding to lapse of time $\leq \tau$ has been given by equation (2.13), viz.

$$N(\tau) = 4\pi n c^3 T^3 \left(-\frac{3}{2} + \frac{\tau}{T} + 2e^{-\tau/T} - \frac{1}{2}e^{-2\tau/T}\right) \quad \ldots\ldots\ldots\ldots\ldots(3.13)$$

where n is the statistically constant number of galaxies, counted locally, per unit volume. For τ small this reduces to

$$N(\tau) \doteq \tfrac{4}{3}\pi n c^3 \tau^3$$

which, since $D \doteq c\tau$ by (3.3) when τ is small, means that

$$N(D) \doteq \tfrac{4}{3}\pi n D^3 \qquad \text{..................(3.14)}$$

Thus the count can be made locally at any point using Euclidean geometry in D. However if D is not small this is not the case and such an assumption for remote regions would not reveal a constant number of nebulae per unit volume by measure D, which might be supposed to indicate the steady state as distinct from the expanding conservation models in which there is no creation of new galaxies. However since it is most practical to use D measure let us find what number density can be expected at any particular range in D.

At range D the Euclidean volume of a sphere radius D is

$$v(D) = \tfrac{4}{3}\pi D^3 \qquad \text{.............(3.15)}$$

It will be easier to use τ as parameter related to D by (3.3). Thus

$$v(\tau) = \tfrac{4}{3}\pi c^3 T^3 \left(e^{6\tau/T} - 3e^{5\tau/T} + 3e^{4\tau/T} - e^{3\tau/T}\right) \qquad \text{....(3.16)}$$

The number density at range D in terms of τ is therefore

$$\frac{dN}{dv} = \frac{n\left(1 - 2e^{-\tau/T} + e^{-2\tau/T}\right)}{2e^{6\tau/T} - 5e^{5\tau/T} + 4e^{4\tau/T} - e^{3\tau/T}} \qquad \text{..........(3.17)}$$

We have seen in Sect. (ii) that the theoretical range of visibility corresponds to infinite τ and infinite D, both these quantities tending to infinity monotonically over the entire range of the coordinate r. From (3.15) it follows that the theoretical range of observation corresponds to infinite 'Euclidean' volume in measure D and from (3.17) we see that the number density tends monotonically to zero from the local value n. At the range corresponding to $\delta \doteq \cdot 2$ for which we found in Sect. (ii) that $D \doteq 1{\cdot}3 \times 10^9$ lt yrs we obtain from (3.4), $e^{\tau/T} \doteq 1{\cdot}2$ so that

$$\left(\frac{dN}{dv}\right)_{\delta=\cdot 2} = \cdot 29\, n \qquad \ldots\ldots(3.18)$$

Thus the apparent number density at this range (now accessible) would be less than $3/10$ of the local value, which affords another observable requirement to be satisfied by the steady state. These results are of course independent of whether we take K to be zero or not, although the measure of D via the magnitudes depends on K.

We shall now find the exact expressions for $N(\delta)$ and **the** gradient of $\log_{10} N, m$, found approximately in Chapter II. By elimination of τ between (3.4) and (3.13) we find

$$N(\delta) = 4\pi n c^3 T^3 \left\{ -\frac{3}{2} + \log_e(1+\delta) + \frac{2}{1+\delta} - \frac{1}{2(1+\delta)^2} \right\} \qquad \ldots\ldots(3.19)$$

This formula and also equation (3.11) were given without proof in the original paper by H. Bondi and T. Gold (<u>13</u>) presenting the steady state theory.

We note that the total visible population, which we have seen corresponds to δ infinite, is infinite.

By (3.11) and (3.19) we find that the gradient of the $\log_{10} N, m$ relation is exactly, for $K=0$, on reduction

$$\frac{d}{dm}\left(\log_{10} N\right) = \frac{2\,\delta^3}{5(1+2\delta)\left\{2(1+\delta)^2 \log_e(1+\delta) - \delta(2+3\delta)\right\}} \qquad \ldots\ldots(3.20)$$

It is easily verified that this gradient is a monotonic expression, decreasing from the local value of .6 in the neighbourhood of the observer $(\delta \simeq 0)$ to zero at the limit of observation $(\delta \to +\infty)$. This justifies the assertion made in Chapter II Sect. (v) that the steady state model cannot yield gradients higher than .6, either for the galaxies themselves or for a uniform sub-group such as say colliding galaxies, so that this model could not support Ryle's phenomenon found for radio stars.

Considering once again the range for which $\delta \simeq \cdot 2$ we find

$$\left[\frac{d}{dm}\left(\log_{10} N\right)\right]_{\delta = \cdot 2} = \cdot 46 \qquad .....(3.21)$$

Another criterion for the steady state model, which is exact, for the case $K = 0$ and the assumed value of T.

(iv) Angular diameter

For the distance $\bar{r}$ by apparent size the formula (1.43) yields for the steady state

$$\bar{r} = r e^{t/T} \qquad ....(3.22)$$

Eliminating r by (3.2) we find

$$\bar{r} = cT\left(1 - e^{(t-t_0)/T}\right) \qquad ...(3.23)$$

Comparing (3.23) with (3.12) we see that for the steady state the distance by apparent size is equal to the integrated proper distance of the source at the time of emission. Using (3.4) we have the relation between distance by apparent size and redshift:

$$\bar{r} = \frac{cT\delta}{1+\delta} \qquad .....(3.24)$$

Thus we see that while the luminosity distance D of sources is theoretically unbounded the distance by apparent size has the upper bound cT, which of course is the maximum proper distance of observable sources at the time of emission, already dealt with in Sect. (ii).

If we now use the relation $d = \bar{r}\,\Delta\theta$ implied by (1.42), connecting linear diameter and angular diameter, and eliminate $\bar{r}$ by (3.24) we obtain the important relation, which must be satisfied exactly by the steady state model, between angular diameter and redshift:

$$\Delta\theta = \frac{d}{cT}\left(1 + 1/\delta\right) \qquad \text{.... (3.25)}$$

The linear diameter d is statistically constant for the steady state so that we find that $\Delta\theta$ decreases monotonically with redshift to a lower limit d/cT for the most distant sources $(\delta \rightarrow +\infty)$. This might be regarded as a remarkable observable aspect of a model for which the observable population of galaxies is theoretically infinite. It is of course associated with the fact that the distance by apparent size, in this case equal to the proper distance at the time of emission, has the upper bound cT. Such a result cannot be deduced for the conservation models because of the uncertainty regarding the evolutionary effect on the linear diameter d. It is nevertheless an important observable characteristic of the steady state which can be applied as a test to the actual universe, to individual galaxies or to a cluster.

The limiting value of $\Delta\theta$ is very sensitive to the average value of d assumed and this would first have to be determined from local observations. If we take $d = 20,000$ lt. yrs., and T the reciprocal of the Hubble parameter equal to 5.4×10^9 yrs., we find that the limiting angular diameter is .77 seconds of arc. For the distance corresponding to $\delta \simeq .2$, at present under observation, the steady state would give, on these assumptions, $\Delta\theta \simeq 4.6$ seconds of arc.

The previous exact results of this section do not depend on the K term. Assuming $K=0$ we shall now relate angular diameter to apparent magnitude m. The general relation between luminosity distance D and distance by apparent size $\bar{r}$ is given by (1.44), so that putting $d = \bar{r}\,\Delta\theta$ we find

$$D = \frac{(1+\delta)^2 d}{\Delta\theta} \qquad \ldots\ldots(3.26)$$

Whence by (3.10) for $K = 0$

$$10^{.2(m-M_0)+1} = \frac{(1+\delta)^2 d}{\Delta\theta} \qquad \ldots\ldots(3.27)$$

Eliminating δ by (3.25) we obtain for the m, $\Delta\theta$ relation

$$10^{.2(m-M_0)+1} = \frac{d\,\Delta\theta}{(\Delta\theta - d/cT)^2} \qquad \ldots\ldots(3.28)$$

which would have to be satisfied by the steady state when $K=0$.

(v) Concluding remarks

The exact results of this chapter (some depending on the possibility of making $K=0$) are to be regarded as ancillary to those approximate results obtained incidentally for the steady state in the general analysis of Chapter II.

In general the value of these results for the steady state is of a negative character. If any of the relations is significantly contradicted by observation then the steady state is ruled out, so that the only alternative consistent with present theory is an evolutionary model. On the other hand if a particular relation is satisfied, within error, by actual observations then we can conclude that the universe does not differ significantly from the steady state in this respect at least, but that in general it would be conceivable that an evolutionary model might be a better fit.

However there are particular relations involving the number counts which, as has been carefully pointed out in Chapter II, are significantly different for the steady state and irreconcilable with any other model. These alone, if satisfied by the observational data, would serve to confirm the steady state, provided the range of likely error can be sufficiently restricted.

CHAPTER IV: APPLICATION OF THE FIELD EQUATIONS OF GENERAL RELATIVITY TO OBTAIN THE CURVATURE

(i) Introduction

In the observable relations developed in this thesis we have given the expansions in series to order δ only in the redshift or equivalent variable, neglecting terms in δ^2 and of higher order. We have demonstrated that to this approximation these relations applied to the observational data could provide the Hubble parameter α_1 determining the present rate of expansion of the universe, the parameter α_2 determining the sign and magnitude of the acceleration of the expansion, and finally the quantity L which gives the first approximation to the rate of evolution of the emission spectrum at arbitrary wavelength. Ample distinctions between the steady state model and an evolutionary model have been provided by this approximation.

It can be shown that terms in δ^2, or the equivalent in alternative variable, may involve the third derivative of the function $R(t)$, the curvature of the 3-space $t =$ constant given by $k/R^2(t)$ including its sign indicated by k, together with terms involving second order 'corrections' resulting from the expansion of the exact correction term involved in equation (1.34) to order δ^2. Examples of relations involving such terms of order δ^2 are to be found in the publications, already referred to, by H. P. Robertson (17) and G. C. McVittie (16).

Although it is of fundamental theoretical importance to obtain the sign of the curvature, there are too many associated unknowns in the coefficients of δ^2 in these relations for even the sign of the curvature to be reliably determined in this way at the present time. Such knowledge required to evaluate it would involve refinements of detail regarding the spectra of the nebulae including their possible evolutionary change much in excess of that possessed at the present time. In addition these terms are unlikely to be larger than the observational errors at the present range of the instruments ($\delta \simeq \cdot 2$).

(ii) Analysis of the field equations

We therefore turn to the field equations of general relativity, in an attempt to obtain the curvature and other information relevant to the solution of the cosmological problem. For the metric (1.2) the field equations, with the cosmological constant put equal to zero, yield (see for example H. Bondi (31))

$$\frac{3\dot{R}^2}{R^2} + \frac{3kc^2}{R^2} = 8\pi G\rho \qquad \text{......(4.1)}$$

$$\frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2} = -\frac{8\pi G}{c^2}p \qquad \text{......(4.2)}$$

where $R \equiv R(t)$, ρ is the average cosmological energy density in mass units, p is the average pressure, and G is the Newtonian constant of gravitation.

According to the substitutions made in Chapter II Sect.(ii) we can write these equations, for the present cosmological epoch $t = t_0$, in terms of the quantities involved in the observable relations derived in this thesis.

$$3\alpha_1^2 + \frac{12c^2}{a^2} = 8\pi G\rho \qquad \text{.....(4.3)}$$

$$2\alpha_2 + \alpha_1^2 + \frac{4c^2}{a^2} = -\frac{8\pi G}{c^2}p \qquad \text{.....(4.4)}$$

where as in Chapter II the sign of the curvature has been absorbed into a^2 according to the relation $\frac{k}{R_0^2} = \frac{4}{a^2}$.

The formulae relating the observable quantities as far as order δ we have shown are in principle adequate to determine both α_1 and α_2. Therefore, provided that observational error can be reduced to a minimum and selective bias eliminated, the range of the present optical and other instruments is sufficient to ultimately provide reliable values for these quantities. By combining (4.3), (4.4) we get

$$\alpha_2 = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \qquad \ldots\ldots(4.5)$$

Thus if $\alpha_2 > 0$ then $\rho + \frac{3p}{c^2} < 0$ ......(4.6)

and if $\alpha_2 < 0$ then $\rho + \frac{3p}{c^2} > 0$ ......(4.7)

In the case of (4.6) we cannot escape the conclusion that one, or both, of ρ and p is negative. In either event this would involve a complete revision of our concept of energy and stress, particularly as to their zero points. In the case of the steady state model, treated according to the equations of general relativity, this conclusion has already been arrived at, by W. H. McCrea (32). In the case of the 'point'source models however the application of the condition (4.6) might also be necessary on the ground that if $\alpha_2 < 0$ then the present age of the universe must be less than the reciprocal of the Hubble parameter, which according to Humason, Mayall and Sandage (equation (3.8)) is $5{\cdot}4 \times 10^9$ years. This figure is very close to the estimated age of some stars.

Nevertheless $\alpha_2 < 0$ is the result found in Chapter II Sect. (iv) on applying the analysis of this thesis to the observational data obtained by Humason, Mayall and Sandage. However the result is uncertain, as was pointed out in Chapter II, not least because of the lack of knowledge regarding a possible evolutionary effect. It has however been pointed out in Chapter II how the number counts can either independently or in association with the redshift data , provide the necessary knowledge of the magnitude of the evolutionary effect.

If we assume for the moment that $\alpha_2 < 0$ then equation (4.7) permits the orthodox interpretation of ρ and p. Accordingly observation would indicate that p, associated with the pressure of radiation, cosmic rays, and the random motion of the galaxies, is very small compared with ρc^2 , so that we can take $p \doteq 0$ and ρ to be approximately the average density of matter in the universe. Equation (4.5) would therefore provide this value of ρ on substituting the value of α_2 obtained from the observational data. Observation of the star population in nearby galaxies provides a check on this value of α_2 in the form of an upper limit, arising from a lower limit for ρ . The curvature is then obtained directly from (4.3) on substituting this value of ρ and the value of the Hubble parameter α_1 .

For the conventional interpretation of ρ and $p\ (\geq 0)$ which we have associated with the condition for a decelerating expansion ($\alpha_2 < 0$), combined with our assumption that the cosmological constant is zero, the general behaviour for the models of general relativity theory corresponding to the different signs of the curvature is well known (see for example R. C. Tolman (12)). If the curvature is positive then space is closed and the model can only oscillate between zero radius and a radius of finite value, while if the curvature is zero or negative then space is open and the function $R(t)$ can only increase from zero monotonically to infinity so that the contents of the universe are ultimately infinitely diffused.

If once again we take the reciprocal of the Hubble parameter to be $5{\cdot}4 \times 10^9$ yrs then we find from (4.1) that

$k \gtreqless 0$ according as $\rho \gtreqless 6{\cdot}2 \times 10^{-29}$ gm/c.c. .....(4.8)

This result holds whatever be the sign of α_2 or whether we take p to be zero or not, and is the criterion for curvature in terms of the total energy density ρ in mass units, at the present epoch. The critical value of ρ is lower than the generally estimated value of total density (33) so that on prima facie evidence space would appear to be closed.

The previous analysis depends on the general assumption that Λ the cosmological constant is zero. Consequently such universes of positive curvature which depend on a non zero Λ, such as the ever expanding Eddington - Lemaitre universe for which $\Lambda > 0$, are excluded from consideration. Taking Λ to be zero is justified in the first place from the point of view of logical economy, since the introduction of non static models permits a non zero density (and positive pressure) without

Λ. In the second place, although originally introduced by Einstein to permit the full incorporation of Mach's principle in general relativity, the presence of Λ is in fact inconsistent with Mach's principle since it prevents the complete dependence of inertia on matter alone, giving rise as it does to the partial dependence of inertia on a dematerialised constant.

Returning again to our equations (4.3), (4.4) let us now consider the possibility that observation indicates that $\alpha_2 > 0$. Then according to (4.6), as has been said, revision is necessary of our ideas of energy and stress, since there is no <u>visible</u> evidence of negative energy or negative stress on the cosmological scale. In this case the field equations could not provide directly a knowledge of the curvature until the nature of ρ and p was fully established, and cosmology would have taken on a radically different significance. (see Part III)

<u>(iii) The steady state model</u>

As has been stated, in the case of the steady state model for which $\alpha_2 > 0$, McCrea has shown how the admission of a negative pressure permits the creation process in that model to be consistent with the field equations of general relativity. The newly created matter is regarded as the mass equivalent of the work done by the negative pressure in the expansion of the model.

The metric in this case is described by $R(t) = e^{t/T}$, $k=0$ so that substitution in (4.3), (4.4) gives

$$8\pi G\rho = \frac{3}{T^2} \qquad \text{......(4.9)}$$

$$\frac{8\pi G}{c^2} p = -\frac{3}{T^2} \qquad \text{.......(4.10)}$$

Since $k = 0$, ρ has the value of the critical density given in (4.8) if we assume that $T = 5.4 \times 10^9$ yrs. This would in principle provide another observable criterion for the steady state. However it is doubtful if a zero point stress of a negative character can be admitted without allowance for the possibility of ρ incorporating a zero point energy density. In fact it would appear to be this possibility that would alone permit the acceptance of models in which the curvature is zero or negative, because of the very low density associated with them, according to equation (4.8).

If we do not accept that the field equations of general relativity are the proper field equations to describe the steady state, as might be the point of view of the originators of the theory, then we must rely on the observable relations dealt with in Chapter II and Chapter III, since the steady state theory is at present without an alternative field theory. Fortunately, as we have demonstrated in Chapter II, the unique character of these relations for the steady state, especially those involving the number counts, and the absence of evolutionary effects, make this alternative a practicable one.

REFERENCES

(1) A. Einstein, Ann. d. Phys. 49 769 (1916)

(2) H. P. Robertson, Proc. Nat. Acad. Sci. 15 822 (1929)

(3) H. P. Robertson, Ap. J. 82 284 (1935)

(4) A. G. Walker, Proc. Lond. Math. Soc. (2) 42 90 (1936)

(5) R. C. Tolman, Relativity, Thermodynamics and Cosmology; Oxford: Clarendon Press 1934 p. 364.

(6) G. Lemaitre, Ann. Soc. Sci. Bruxelles 47 A 49 (1927)

(7) E. Hubble, Proc. Nat. Acad. Sci. 15 168 (1929)

(8) E. T. Whittaker, Proc. Roy. Soc. 133 A 93 (1931)

(9) R. C. Tolman, Proc. Nat. Acad. Sci. 16 511 (1930)

(10) W. H. McCrea, Zs. f. Ap. 9 290 (1934-35)

(11) J. Stebbins and A. E. Whitford Ap. J. 108 413 (1948)

(12) R. C. Tolman, Op. cit. (5) p. 394

(13) H. Bondi and T. Gold, Mon. Not. R.A.S. 108 252 (1948)

(14) F. Hoyle, Mon. Not. R.A.S. 108 372 (1948)

(15) M. L. Humason, N. U. Mayall and A. R. Sandage, Astron. J. 61 97 (1956)

(16) G. C. McVittie, General Relativity and Cosmology; London: Chapman and Hall, 1956 p.154.

(17) H. P. Robertson, Pub. Astr. Soc. Pac. 67 82 (1955)

(18) A. E. Whitford, Reported in Sky and Telescope 16 222 (1957)

(19) E. Hubble and R. C. Tolman, Ap. J. 82 302 (1935)

(20) W. A. Baum, Astron. J. 58 211 (1953)

(21) J. R. Shakeshaft, Phil. Mag. 45 1136 (1954)

(22) M. Ryle and P. Scheuer, Proc. Roy. Soc. A. 230 448 (1955)

(23) G. C. McVittie, Proc. Phys. Soc. 51 529 (1939)

(24) O. Heckmann, Theorien der Kosmologie; Berlin: Springer (1942)

(25) E. Hubble, Ap. J. 64 321 (1926)

(26) E. A. Milne, Kinematic Relativity; Oxford: Clarendon Press 1948 p.24

(27) F. Zwicky, Proc. Nat. Acad. Sci. 15 773 1929

(28) E. R. R. Holmberg, Mon. Not. R.A.S. 116 691 (1956)

(29) S. N. Milford, Ap. J. 122 13 (1955)

(30) W. Rindler, Mon. Not. R.A.S. 116 662 (1956)

(31) H. Bondi, Cosmology; Cambridge: C.U.P. 1952 p.103

(32) W. H. McCrea, Proc. Roy. Soc. A 206 562 1951

(33) H. Bondi, Op. cit. 31 p. 45

(34) B. Y. Mills, Scientific American 195 No. 6, 8, (1956)

PART II

GENERAL RELATIVITY AND MACH'S PRINCIPLE

CHAPTER I: THE PRINCIPLE OF MACH

(i) Newtonian Mechanics

In Newtonian mechanics space and time were regarded as absolute, forming a background to physical events but remaining independent of and unaffected by the matter featuring in these events. Relative to absolute space matter moved according to Newton's three laws of inertia. The geometry was assumed to be Euclidean in absolute space.

A logical shortcoming of the theory is that it gave no means of identification of this space relative to matter, except for the statement in the first law that matter under no force would be without acceleration in this space. The absence of force however is not detectable in any other way. In practice it is found that relative to the inertial plane of a swinging pendulum (Foucault pendulum) the earth rotates with the same angular velocity as it does relative to the distant stars. Thus the distant stars would seem to be at rest in absolute space.

The forces taken account of in the theory included gravitational force, which according to Newton's universal law of gravitation was proportional to the mass of the body upon which it acted. In the absence of a gravitational field in absolute space the equation of motion of a particle would be in accordance with the inertial laws, under inertial forces only (non-gravitational). This is found to apply in practice in any frame of reference in which the local gravitational field vanishes and in which the plane of a swinging pendulum remains without rotation. These are the so-called inertial frames, 'freely falling' and without rotation relative to the distant stars. Thus the frame of reference which effectively constituted absolute space for Newton, that is in which the sun and distant stars are at rest, is inertial if we neglect the local gravitational field of the solar system.

However if a reference frame is chosen rotating and accelerating relatively to an inertial frame then the Newtonian laws hold for the motion of a particle only by the introduction of the so-called fictitious forces - centrifugal force and force of Coriolis, together with an artificial homogeneous gravitational field. Thus the inertial frames appear as preferred frames of reference in which it happens that the distant stars are not revolving. The theory supplied no causal connection between these two facts.

The philosopher and scientist E. Mach (1) regarded this as unsatisfactory and put forward the view that the notion of absolute space was a 'forlorn' concept 'bereft of all scientific significance' (op.cit. p.229). Thus Mach observes (p.229): 'When we say that a body K alters its direction and velocity, solely through the influence of another body K', we have asserted a conception that it is impossible to come to, unless other bodies A, B, C ...... are present with reference to which the motion of the body K has been estimated. In reality therefore we are simply cognisant of a relation of the body K to A, B, C ....'.

Instead of absolute space Mach proposed to substitute the hypothesis, generally known as Mach's Principle, that inertia and the inertial frames are determined by a kind of gravitational effect of all the matter in the universe. Thus he says (p.231): 'When we reflect that we cannot abolish the isolated bodies A, B, C .... that is cannot determine by experiment whether the part they play is fundamental or collateral, that hitherto they have been the sole and only competent means of the orientation of motions and of the description of mechanical facts it will be found expedient provisionally to regard all motions as determined by these bodies'.

It was envisaged by Mach that the inertial frames are determined by the fact that in them the average motion of universal matter, measured in some definite way, is zero. Thus he asserts (p.234): 'Instead of saying, the direction and velocity of a mass μ in space remain constant, we may also employ the expression, the mean acceleration of the mass μ with respect to the masses m, m', m''....... at the distances r, r', r'' is zero,.......' In view of this remark, associated with the quotation to follow, it is clear that Mach looked upon inertia itself as arising from acceleration relative to the universe as a whole: 'It is known from recent hydrodynamical investigations that a rigid body experiences resistance in a frictionless fluid only when its velocity <u>changes</u>. True, this result is derived theoretically from the notion of inertia; but it might, conversely, be regarded as the primitive fact from which we have to start'.

By Mach's Principle the chain of cause and effect in the physical phenomena would accordingly be closed, as Mach desired, and the a priori concept of absolute space would be eliminated. However the limitations, at that period, of the notions of distance and time prevented Mach from providing an analytic theory of his general ideas which would satisfactorily provide for the convergence of the world effects of matter on matter. If his views, which have appealed to many scientists, are correct, then both the experiment of Foucault's pendulum and the practical constancy of inertial mass indicate that the effect of distant matter in the universe must heavily predominate over that of local matter, such as the rotating earth or the revolving planets, in determining inertia and the inertial frames. To the accuracy of the present means of detection no effect of local matter, in the sense of Mach's Principle, has yet been observed.

Nevertheless, a physical theory that described physical phenomena at least as well as Newtonian mechanics, and which in addition was independent of the concept of absolute space, but incorporated a basis for the effect of matter on inertia, at least in principle, would be heuristically preferable to Newtonian mechanics.

(ii) Relativity Mechanics and the General Theory of Relativity

Another fact for which the Newtonian theory offered no fundamental explanation was that according to the gravitational law the gravitational force on a body was proportional to its mass so that all bodies fell with the same acceleration in a gravitational field, independently of their masses. For all other forces the acceleration produced was inversely proportional to the mass. Accordingly the masses of two bodies could be compared either by the ratio of their attractions on a standard body, or by the inverse ratio of the accelerations produced in them by the same force. The equality of the 'gravitational' mass and the 'inertial' mass, confirmed to great accuracy by the experiments of R. Eotvos (2), was without physical explanation in the Newtonian theory.

As is well known this equality was made a logical consequence of the Principle of Equivalence, put forward by A. Einstein to form the basis of the general theory of relativity. Previous to this Newtonian mechanics in <u>inertial</u> frames had given place to the mechanics of special relativity. According to the Principle of Special Relativity the same laws of electrodynamics and mechanics were to be valid in all inertial frames i.e. frames in which there was no gravitational field, real or fictitious. Thus, as had already been found for Newtonian mechanics, all inertial frames moving relatively to one another with uniform velocity were to be physically equivalent, thus eliminating the notion of absolute rest. This was emphasised by the particular postulate that the velocity of light in a vacuum was to be the same in all inertial frames. Thus special relativity, which was developed kinematically from this postulate, was in agreement with the fact that no velocity of the earth could be obtained relative to the supposed aether (Michelson-Morley experiment). In addition it was in agreement with the observed fact that the electrodynamic phenomena seemed to depend only on the <u>relative</u> motion of conductors, magnets etc., and not on the concept of absolute space.

An important feature in the development of special relativity was that it led to the celebrated mass-energy relationship $E = mc^2$, and to the realisation that energy had inertia as well as matter in accordance with this relation, which has been amply confirmed by experiment. In addition it was realised that the estimated length of a rod would depend on the observer's relative velocity to the rod; neither would observers moving relatively to one another with uniform velocity agree as to the interval of time between two events. Thus the notions of absolute space and absolute time had no place in the analysis of the physical phenomena in inertial frames. Nevertheless special relativity had nothing to say about the comparison of observations between relatively accelerating observers.

By the Principle of Equivalence Einstein maintained that a reference frame having acceleration g relative to an inertial frame was physically equivalent to one that was stationary in a homogeneous gravitational field of intensity-g. According to the first way of looking at it, therefore, a given mass could be measured by its inertia, while according to the second by its weight. Thus the identity of the gravitational and inertial masses followed as a matter of course.

The confirmation by the experiments of Eotvos of the truth of this logical inference from the Principle of Equivalence demonstrates the validity of this fundamental principle of general relativity. It also makes possible the origin of inertia according to Mach's Principle so that general relativity may be an expression of that principle. For consider the following observations (by the present writer).

(a) If a particle is at rest in an inertial frame then the bringing up of a large gravitating body will render the frame non-inertial. On the other hand a reference frame co-moving with the particle, now accelerating (without rotation) relatively to the old frame, will have become inertial. Thus the inertial frames are certainly affected by the distribution of matter in the universe. Since an inertial frame, to continue inertial, must, if allowed, always move so as to nullify any suggestion of a gravitational field on a particle at rest relative to it, it seems inevitable to conclude that in fact in so doing it cancels the field, measured in a suitable way, of the whole universe. We can justify and illustrate this conclusion as follows.

(b) If, according to the Principle of Equivalence, we cannot distinguish between the apparent gravitational field generated by an inertial force and a 'real' gravitational field, it follows that the inertial mass of a particle accelerating relatively to an inertial frame must be interpreted as proportional to its weight in the field of the rest of the universe. If, as we have seen Mach suggested, universal matter is on the average at rest in any inertial frame then the universe as a whole is accelerating relative to the particle. It is clear therefore that the measure of the effect of the whole universe on inertia must take account of its motion in addition to its Newtonian attraction. We shall see later that general relativity permits such a measure in detail.

(c) Let us now remove the inertial force, but allow the particle to maintain the same acceleration (at the instant) as before, relatively to the old inertial frame, by placing a large gravitating body in a suitable position. The rest frame of the particle will now be inertial, assuming that it does not rotate relatively to the previous inertial frame. It follows that this body has cancelled the field of the rest of the universe at the particle. We see therefore how universal matter can determine inertia and the inertial frames in accordance with Mach's Principle, if we invoke the Principle of Equivalence.

As soon as we accept that the difference between inertial and non-inertial frames is equivalent to the difference between frames in which there is or is not a gravitational field, then both types of frames 'may with equal right be looked upon as "stationary", that is to say they have an equal title as systems of reference for the physical description of phenomena'. (A. Einstein (3)). This led Einstein to make the general postulate of relativity that the laws of physics must be covariant with respect to arbitrary changes of the reference systems by means of coordinate transformations, the coordinates themselves to be of arbitrary physical definition. Einstein further justified this requirement of general covariance of the physical laws by the observation that, since the description of the universe depended fundamentally on the coincidence of space-time events with pointer readings on measuring instruments, then there is no *a priori* reason for preferring the physical description of one observer to any other.

It is now well known how, in the theory of general relativity, this covariance of the physical laws was achieved (3). By an arbitrary transformation of coordinates the invariant

$$ds^2 = c^2\,dT^2 - dx^2 - dy^2 - dz^2 \qquad \text{......(1.1)}$$

characterising the space-time measurements of special relativity, applicable in any locally inertial frame, was generalised to the form

$$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu \qquad \text{......(1.2)}$$

applicable in a reference frame moving arbitrarily relatively to the locally inertial frame. The relation of the invariant interval ds to the measurements, by metric stick and clock, of freely falling observers in the neighbourhood of the events connected by (1.2) has been described in Part I, Chapter I of this thesis. Since the $g_{\mu\nu}$ are functions of the x^{μ} it is evident that the spatial geometry in a general reference frame will not be Euclidean; nor will clocks, located at different spatial points of these generally non-inertial frames, keep time with one another as in the case of special relativity.

The differential equations of motion of a free particle in special relativity, viz. $\frac{d^2X}{ds^2} = \frac{d^2Y}{ds^2} = \frac{d^2Z}{ds^2} = \frac{d^2T}{ds^2} = 0$, applicable in a locally inertial frame, become in general relativity the covariant differential equations of the time-like geodesic in the space-time of an arbitrary reference frame, viz.

$$\frac{d^2x^{\mu}}{ds^2} + \left\{ {\mu \atop \alpha\beta} \right\} \frac{dx^{\alpha}}{ds} \frac{dx^{\beta}}{ds} = 0 , \qquad \ldots\ldots(1.3)$$

where the $\left\{ {\mu \atop \alpha\beta} \right\}$ depend on the $g_{\mu\nu}$ and their first derivatives.

As explained in Part I, Chapter I the general metric (1.2) can only locally in an infinitesimal region of space-time be reduced to the metric (1.1) of special relativity. Accordingly the general space-time variability of the $g_{\mu\nu}$ in (1.2) must be associated with the presence of the gravitational field in this general reference frame. Indeed by supposing that the $g_{\mu\nu}$ are independent of the time coordinate (static field), and differ only infinitesimally from the 'Galilean' values of special relativity, the equation (1.3) for the motion of a free particle <u>in space</u> can be written to a first approximation (as is easily shown),

$$\frac{d^2x^k}{d(x^4)^2} = -\frac{\partial}{\partial x^k}\left(\frac{1}{2}g_{44}\right), \quad \left(k = 1,2,3\right)$$

Accordingly the potential of the gravitational field, in the Newtonian sense, has to be identified with $\frac{1}{2}g_{44}$, or at least its derivatives with those of $\frac{1}{2}g_{44}$.

For the equations of the gravitational field Einstein therefore looked for covariant relations between the $g_{\mu\nu}$ and the source of the gravitational field, that is the distribution and motion of matter in the reference frame. This latter was characterised by the covariant stress-energy-momentum tensor $T_{\mu\nu}$, whose components in a locally inertial system were to have the usual interpretation of special relativity. Since the ordinary divergence of $T^{\mu\nu}$ vanished in special relativity, Einstein postulated the vanishing of a generally covariant tensor divergence of $T^{\mu\nu}$, already known from the work of Ricci and Levi-Civita. Einstein therefore looked for a tensor of the second rank, involving ~~(linearly)~~ derivatives of the $g_{\mu\nu}$ up to the second order only (in analogy with Poisson's equation), and whose tensor divergence vanished, which he could equate to the $T_{\mu\nu}$. He thus finally obtained the generally covariant equations of the field

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -\kappa T_{\mu\nu} \qquad \text{..........(1.4)}$$

where $R_{\mu\nu}$ is a tensor involving derivatives of the $g_{\mu\nu}$ up to the second order and R is the invariant obtained by the contraction of this tensor. κ is a constant which is identified as $8\pi G/c^2$, where G is the Newtonian gravitational constant, by obtaining the analogue of Poisson's equation from (1.4) for a weak static field. The energy density of matter for this purpose is taken in mass units.

By these tensor equations, which are second order differential equations for the $g_{\mu\nu}$ in terms of the components of the matter tensor $T_{\mu\nu}$, Einstein hoped that Mach's Principle had been fully incorporated in general relativity. In Chapter II we shall examine how far this hope would appear to have been realised.

(iii) Sciama's theory of Mach's Principle

An independent theory of Mach's Principle has been put forward tentatively by D. W. Sciama (4). Sciama takes the view that general relativity has failed to account satisfactorily for the inertial properties of matter. His views in this connection and those of others who hold the same opinion will be presented in Chapter II. Suffice it to say here that he rested on the conclusion that the field equations of general relativity imply that a single gravitating body in an otherwise empty universe produces inertial effects equal to that produced by a full universe. Accordingly Sciama desired to find a theory of gravitation that would imply that matter has inertia only in the presence of other matter. The theory with this property which appear to him to be the simplest is outlined below.

For the influence of matter on matter Sciama defines a scalar and vector potential field analogous to that in classical electromagnetic theory, with associated 'gravo-electric' and 'gravo-magnetic' field intensities. He makes the fundamental postulate that 'in the rest frame of any body the gravitational field of the universe as a whole cancels the gravitational field of local matter, so that in this frame the body is "free". Thus in this theory inertial effects arise from the gravitational field of a moving universe'. Assuming flat space-time Sciama therefore postulates the following equation to hold at the particle in its rest frame

$$\underline{E} = -\operatorname{grad}\Phi - \frac{1}{c}\frac{\partial \underline{A}}{\partial t} \qquad \text{............ (1.5)}$$

where $\Phi = -\int_{r=0}^{r=c\tau} \frac{\sigma\, dV}{r}$, $\underline{A} = -\int_{r=0}^{r=c\tau} \frac{\sigma \underline{u}\, dV}{c r}$ .... (1.6)

Here σ is the cosmic gravitational mass density, but it is clear that Sciama does not distinguish between this and the cosmic inertial mass density. The limits of integration correspond to a cosmic volume of radius $c\tau$, where τ is the reciprocal of the Hubble parameter, to allow in a plausible manner for relativistic cut off in an expanding universe. The particle is in the position $r = 0$.

By application of his theory to the motion of a free particle relative to a local mass M at rest in the universe, which is regarded as otherwise smoothed out, Sciama derives the equation

$$\frac{M}{r^2} = -\left(\frac{\Phi + \phi}{c^2}\right)\frac{dv}{dt} \qquad \text{..........(1.7)}$$

where $\phi = -M/r$, and v is the speed of the particle relative to the mass M.

Making contact with the corresponding Newtonian equation Sciama concludes that the inertial mass of the particle has to be interpreted as proportional to

$$-G(\Phi + \phi)/c^2 \qquad \text{....................(1.8)}$$

where the constant of proportionality is an invariant for the particle independently of position. This requires

$$\left.\begin{array}{l} G\Phi \triangleq -c^2 \\ \text{or } G\sigma\tau^2 \triangleq 1 \end{array}\right\} \qquad \text{....................(1.9)}$$

Thus a knowledge of G and τ by experiment and observation leads to the value of σ the average gravitational mass density of the universe, and therefore provides knowledge of the effective mass of the universe.

Sciama claims therefore that his theory incorporates a combination of Newton's laws of motion and of gravitation, with inertia and the inertial frames determined in accordance with Mach's Principle. This claim, as far as his tentative theory goes, appears to be justified. However it will be shown in Chapter II that general relativity is fully consistent with this interpretation of Mach's Principle by Sciama, and will in fact appear to be the general tensor theory which Sciama desired to find, to which he recognized his theory could only be a first approximation. Indeed we have already shown in Section (ii) of this Chapter how, on the basis of the Principle of Equivalence, inertia can be regarded as arising from the gravitational influence of the whole universe, with the inertial frames determined by the fact that the gravitation of local matter is cancelled at a test-particle by the gravitation arising from a moving universe. This is precisely the qualitative basis of Sciama's theory. In Chapter II we shall show further that, in the first approximation, general relativity supports this interpretation by equations of the same mathematical form as in Sciama's theory. That is, the equations postulated by Sciama are already contained in general relativity.

CHAPTER II: GENERAL RELATIVITY AS AN EXPRESSION OF MACH'S PRINCIPLE (1)

(i) The motion of a free particle when space-time is quasi-Galilean

The deviations of the $g_{\mu\nu}$ from the 'Galilean' values which obtain in special relativity (c.f. equation (1.1)) are a measure of the departure of the reference frame from an inertial system. The $g_{\mu\nu}$ also determine the geodesic world line of a free particle projected in the system. To this extent therefore the field equations (1.4) are consistent with Mach's Principle since, as we have pointed out in Chapter I, they indicate that the $g_{\mu\nu}$ depend on the distribution and motion of mass in the reference frame. We shall now examine in detail how far the inertia of a free particle appears to depend on the $g_{\mu\nu}$, and to what extent inertia, the inertial frames, and the $g_{\mu\nu}$ are determined in accordance with Mach's Principle.

We restrict our considerations in the first instance to a quasi-Galilean system, that is whose departure from an inertial system is slight and confined to the neighbourhood of the spatial origin so that the $g_{\mu\nu}$ take on the Galilean values at sufficient distance. For this case of a weak gravitational field, in which the spatial velocity of a free particle is small compared with the velocity of light, the motion of such a particle can be described by a Maxwell type pondermotive equation. This idea is not new and has in fact, with limited application, been presented by Einstein (5). But since

Einstein's derivation of the result appears to contain errors of detail we give our own derivation here, before investigating its significance for Mach's Principle.

Here, and henceforth in this thesis, let Latin letters refer to the numbers 1,2,3 associated with spatial coordinates; Greek letters will, as previously, cover all space-time coordinates. It is well known that the equations of the geodesic followed by a free particle, given by (1.3), can be written in the equivalent form

$$\frac{d}{ds}\left(g_{\mu\alpha}\frac{dx^\alpha}{ds}\right) = \frac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} \qquad \ldots\ldots\ldots(2.1)$$

where by (1.2)

$$ds^2 = g_{44}(dx^4)^2 + 2g_{4h}\,dx^4 dx^h + g_{hq}\,dx^h dx^q \qquad \ldots\ldots\ldots(2.2)$$

For $\mu = i$ equations (2.1) may be written

$$\left(\frac{ds}{dx^4}\right)^2\frac{d}{ds}\left(g_{ih}\frac{dx^h}{ds}\right) = \frac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \left(\frac{ds}{dx^4}\right)^2\frac{d}{ds}\left(g_{i4}\frac{dx^4}{ds}\right) + \frac{\partial g_{h4}}{\partial x^i}\frac{dx^h}{dx^4} + \frac{1}{2}\frac{\partial g_{hq}}{\partial x^i}\frac{dx^h}{dx^4}\frac{dx^q}{dx^4}$$

Write now $x^4 \equiv t$, $dx^h/dt \equiv v^h$ and neglect squares and products of the spatial coordinate velocities v^h , getting

$$\frac{ds}{dt}\frac{d}{dt}\left(g_{ih}\frac{dt}{ds}v^h\right) = \frac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{ds}{dt}\frac{d}{dt}\left(g_{i4}\frac{dt}{ds}\right) + \frac{\partial g_{h4}}{\partial x^i}v^h \qquad \ldots\ldots(2.3)$$

The general metric (2.2) may be expressed in the form

$$ds^2 = (1+\gamma_{44})dt^2 + 2\gamma_{4h}\,dx^h dt - (1-\gamma_{11})(dx^1)^2 - (1-\gamma_{22})(dx^2)^2 - (1-\gamma_{33})(dx^3)^2 + \gamma_{hq}\,dx^h dx^q \qquad (h \neq q).$$

$$\ldots\ldots(2.4)$$

where we have taken c the velocity of light as unity.

It is evident that the $\gamma_{\mu\nu}$ are the deviations of the $g_{\mu\nu}$ from the Galilean values. They are the $\gamma_{\mu\nu}$ of Einstein's analysis except for the modification due to his employment of imaginary x^4.

We now make the assumption that the squares and products of the $\gamma_{\mu\nu}$ and those of their derivatives can be neglected. To this approximation therefore, retaining only first order terms, we can reduce equation (2.3) to

$$\frac{d}{dt}\left(\frac{g_{ii}v^i}{\sqrt{g_{44}}}\right) = \frac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{d}{dt}(g_{i4}) + \frac{\partial g_{h4}}{\partial x^i}v^h$$

on using (2.2) to find the appropriate approximation for ds/dt in each term. On rearranging we can write

$$\frac{d}{dt}\left(\frac{-g_{ii}v^i}{\sqrt{g_{44}}}\right) = -\frac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{\partial}{\partial t}(g_{4i}) + \left\{\frac{\partial}{\partial x^h}(g_{4i}) - \frac{\partial}{\partial x^i}(g_{4h})\right\}v^h \quad \ldots(2.5)$$

so that in terms of the metric (2.4) the equations can be expressed, to the required approximation

$$\frac{d}{dt}\left\{\left(1-\gamma_{ii}-\tfrac{1}{2}\gamma_{44}\right)v^i\right\} = -\frac{1}{2}\frac{\partial\gamma_{44}}{\partial x^i} - \frac{\partial}{\partial t}(\gamma_{4i}) + \left\{\frac{\partial}{\partial x^h}(\gamma_{4i}) - \frac{\partial}{\partial x^i}(\gamma_{4h})\right\}v^h \quad \ldots(2.6)$$

In solving the field equations to this approximation Einstein showed (6) that the $\gamma_{\mu\nu}$ were the solutions to the equations

$$\left\{\left(\frac{\partial}{\partial x^1}\right)^2 + \left(\frac{\partial}{\partial x^2}\right)^2 + \left(\frac{\partial}{\partial x^3}\right)^2 - \left(\frac{\partial}{\partial t}\right)^2\right\}\gamma_{\mu\nu} = 2K\left(T_{\mu\nu} - \tfrac{1}{2}\delta_{\mu\nu}T\right) \quad \ldots(2.7)$$

provided $\dfrac{\partial^2}{\partial x^\nu\partial x^\alpha}\left(\gamma^\alpha_\mu - \tfrac{1}{2}\delta^\alpha_\mu\gamma^\beta_\beta\right) + \dfrac{\partial^2}{\partial x^\mu\partial x^\alpha}\left(\gamma^\alpha_\nu - \tfrac{1}{2}\delta^\alpha_\nu\gamma^\beta_\beta\right) = 0 \quad \ldots(2.8)$

to the order of the approximation. Here κ is as defined with equations (1.4), and $\gamma^{\alpha}_{\mu} = \delta^{\alpha\beta}\gamma_{\mu\beta}$ where $\delta^{\alpha\beta}$ are the Galilean values of the $g^{\alpha\beta}$. Assuming the contribution of stress to the stress-energy-momentum tensor to be vanishingly small compared with densities of mass and momentum for the case he was considering, Einstein obtained the solution:

$$\left.\begin{aligned} &\gamma_{11} = \gamma_{22} = \gamma_{33} = \gamma_{44} = -\frac{\kappa}{4\pi}\int\frac{[\rho]\,dV}{r}, \quad \gamma_{4p} = \frac{\kappa}{2\pi}\int\frac{[\rho u^p]\,dV}{r} \\ &\text{and } \gamma_{pq} = 0, \quad p \neq q. \end{aligned}\right\} \quad \ldots(2.9)$$

In this solution ρ is the mass density, u^p the space coordinate velocity, of the element of mass in the coordinate volume $dV \; (= dx^1 dx^2 dx^3)$ at distance r from the point where the $\gamma_{\mu\nu}$ are evaluated. Square brackets indicate retarded values corresponding to the propagation of the field with the unit velocity.

The integrals, supposed convergent, are over all matter producing the field. Such a solution of the wave equation of Lorentz (equation (2.7)) is well known to be valid only if the quantities solved for (the $\gamma_{\mu\nu}$) tend to zero in a suitable way. This is associated with the fact that the solution implies that the densities of mass and momentum must vanish at 'infinity'. Thus it is clear that the solution considered by Einstein involves a metric which is Galilean at sufficient distance, associated with mass concentrations only in the neighbourhood of the space origin, with empty space at infinity.

The solution has to be consistent with the condition (2.8) which will be satisfied if the expressions $\frac{\partial}{\partial x^\alpha}\left(\gamma^\alpha_\mu - \frac{1}{2}\delta^\alpha_\mu \gamma^\beta_\beta\right)$ vanish, for all μ , to the first order in the $\gamma_{\mu\nu}$. Using (2.9) and the fundamental conservation equations $T^{\mu\nu}_{;\nu} = 0$, it is easily seen that for integrals over a finite region of mass these expressions are indeed second order quantities.

If we now put

$$\phi = -G\int\frac{[\rho]}{r}dV , \qquad A^h = -4G\int\frac{[\rho u^h]}{r}dV \qquad .....(2.10)$$

so that $\gamma_{11} = \gamma_{22} = \gamma_{33} = \gamma_{44} = 2\phi , \ \gamma_{4h} = -A^h$ .....(2.11)

then equation (2.6) can be written in vector form, covering $i = 1,2,3$

$$\frac{d}{dt}\left\{(1-3\phi)\underline{v}\right\} = -\mathrm{grad}\,\phi - \frac{\partial \underline{A}}{\partial t} + \underline{v}\wedge \mathrm{curl}\,\underline{A} \qquad .....(2.12)$$

This is the required Maxwell type pondermotive equation of the field. The three assumptions made during its derivation are:

(a) The particle velocity $\underline{v}$ in the reference frame is assumed small such that v^2/c^2 is negligible compared with v/c .

(b) The deviations of the $g_{\mu\nu}$ from the Galilean values are small such that their squares and products, and those of their derivatives, can be neglected.

(c) The deviations $\gamma_{\mu\nu}$ vanish at 'infinity' so that the quantities $\underline{A}$, ϕ are defined in terms of convergent integrals.

If <u>in addition</u> we now further assume that

(d) The source velocities of the field are also small in the reference frame so that the same remark as in (a) applies for them then equation (2.12) reduces to

$$\frac{d\underline{v}}{dt} = -\mathrm{grad}\,\phi - \frac{\partial \underline{A}}{\partial t} + \underline{v}\wedge \mathrm{curl}\,\underline{A} \qquad .....(2.13)$$

The equation obtained by Einstein was (our notation)

$$\frac{d}{dt}\left\{(1-\phi)\underline{v}\right\} = -\operatorname{grad}\phi - \frac{\partial \underline{A}}{\partial t} + \underline{v}\wedge \operatorname{curl}\underline{A}$$

Since he assumed condition (d) as well as (a), (b) and (c) his result is incorrect to the order he was considering and misleading. In obtaining this result he put $[h4, i] = \frac{1}{2}\left(\frac{\partial g_{4i}}{\partial x^h} - \frac{\partial g_{4h}}{\partial x^i}\right)$ thereby omitting or neglecting the term $\partial g_{ih}/\partial x^4$ which when $h = i$ contributes to our result in equation (2.6) as the term $\frac{d}{dt}\left(-\gamma_{ii}\right)$ in the coefficient of v^i in the left hand side. The neglect of this term is of course consistent with condition (d), but on the other hand the retention of the term $\frac{d}{dt}(-\phi)$ in the coefficient of $\underline{v}$, arising in our approximation from the term $\frac{d}{dt}\left(-\frac{1}{2}\gamma_{44}\right)$ in the coefficient, is not consistent with Einstein's assumptions.

(ii) <u>Interpretation of the pondermotive equation for quasi-Galilean fields</u>

As Einstein pointed out equation (2.12) indicates that general relativity goes far towards incorporating Mach's Principle. It may be compared with the Newtonian equation, viz.:

$$\frac{d\underline{v}}{dt} = -\operatorname{grad}\phi$$

The additional terms are small in the quasi-Galilean frame considered by Einstein and, as he said, beyond physical measurement. Nevertheless they show in the sense of Mach's Principle how concentrated matter affects the inertial mass of a freely moving particle, and the acceleration of its co-moving local inertial frame relative to the given frame, in the following respects:

(a) The inertial mass is apparently proportional to $1-3\phi$.

(b) The locally inertial rest frame of the particle is accelerated by means of the following effects:

(i) Gravitational attraction towards the local mass concentrations indicated by the term $-\operatorname{grad}\phi$.

(ii) An inductive effect of local accelerating matter in the same sense as the acceleration, indicated by the term $-\partial \underline{A}/\partial t$.

(iii) An inductive effect of matter which is rotating relative to the compass of inertia (to use Godel's phrase) at 'infinity' in the sense of the rotation, as indicated by the term $\underline{v} \wedge \operatorname{curl} \underline{A}$. This is of the same type as the 'fictitious' Coriolis force familiar in Newtonian dynamics, with its associated centrifugal force, when a reference frame is used which is rotating relative to the compass of inertia.

It is clear therefore that general relativity certainly incorporates in detailed manner the aspects of Mach's Principle indicated above. For a satisfactory theory of Mach's Principle however Einstein realised the necessity of showing how inertia depended on the entire cosmic distribution of matter. The position of general relativity with regard to this fundamental issue will now be discussed.

(iii) The influence of the cosmological distribution of mass on inertia

Since the field equations (1.4) of general relativity are differential equations (second order) for the $g_{\mu\nu}$ in terms of the components of the stress-energy-momentum tensor, the boundary conditions regarding the cosmological distribution of mass are necessary for their complete solution. The crucial question vis-à-vis Mach's Principle is whether the boundary conditions, imposed to determine the $g_{\mu\nu}$ fully, are consistent with the identification of the whole of the inertia of a particle, and of the existence of inertial frames, with an integrated influence of cosmic mass.

In his early attempts at a cosmological solution Einstein regarded the full determination of the $g_{\mu\nu}$ in terms of boundary conditions at spatial infinity. He argued that according to Mach's Principle the inertia of a particle sufficiently far from other matter should fall to zero. By our equation (2.5) which holds quite generally subject only to conditions (a) and (b) of Section (i) we see that the inertia of a particle, in a space-time in which $g_{11} = g_{22} = g_{33}$, is proportional to $g_{11}/\sqrt{g_{44}}$. It follows that, as Einstein has pointed out (7), if the density of cosmic mass tends to zero at spatial infinity then Mach's Principle would require that $g_{11} \to 0$, or $g_{44} \to \infty$, at spatial infinity. Yet it is clear that, as we have demonstrated in Section (i), Einstein's solution of the field equations assuming empty space at infinity, even allowing for the presence of matter in finite space, leads directly to a metric which is Galilean at infinity. The exact Schwarzschild solution for a single body in empty space leads to the same result. According to the metrics of these solutions, given for the approximate solution by (2.9), the matter present in finite space accounts for only that part of the $g_{\mu\nu}$ which represents the deviations from the values at spatial infinity.

This failure by Einstein to obtain boundary conditions at spatial infinity consistent with Mach's Principle led him to consider the possibility of a cosmological solution in which the universe was spatially finite (re-entrant), so that the difficulty of boundary conditions at spatial infinity would be eliminated. He found that the gravitational equations admitted a static, homogeneous and isotropic, universe in which space was spherical of a radius R related to cosmic mass density ρ. The solution demanded however a negative pressure p equal in magnitude to one third of the energy density ρc^2. Accordingly Einstein modified the field equations by introducing a logically permissible cosmological constant Λ which allowed the Einstein universe to correspond to positive or zero pressure (7).

However, although logically permissible on other grounds, it cannot be said that Einstein satisfactorily accounted for inertia in terms of Mach's Principle by the device of the cosmological constant. For now inertia (the $g_{\mu\nu}$) depended, via the radius of curvature, on a dematerialised constant as well as on cosmic mass density.

On the other hand, if we accept the possibility of a negative zero point stress, then the cosmological constant can be regarded as never necessary for Einstein's purpose. By the same token the cosmic space-time discovered by de Sitter (8), which has usually been regarded as empty thereby eschewing Mach's Principle in general relativity (see for example H. Bondi (9)), need not be so regarded. For the field equations, whether Λ be present or not, demand only of the de Sitter metric that $\rho c^2 + p = 0$.

The de Sitter metric has been found to be appropriate to the space-time of the steady state model proposed by H. Bondi and T. Gold (10), and by F. Hoyle (11). Indeed the negative pressure associated with the de Sitter metric has been invoked by W. H. McCrea (12) to render the creation of matter in the steady state consistent with general relativity. It will in fact be shown later (Chapter IV) that the steady state model interpreted according to general relativity, would seem to provide an origin for inertia completely consistent with Mach's Principle.

The failure by Einstein to find boundary conditions at spatial infinity to allow inertia to be fully determined by matter has remained an essential reason for the view held by many that general relativity does not fully incorporate Mach's Principle (4), (9). However in the view of the writer of this thesis it is not legitimate to postulate empty space at infinity as a cosmic solution to the field equations. For empty space can only be described by a reference frame which is purely conceptual, defined without reference to matter or radiation, and restoring to space an objective quality, independent of matter, which the field equations of general relativity would patently deny. It is unsatisfactory to suppose that matter can affect inertia, as has been shown in detail in Sections (i), (ii), and yet not wholly cause it; but this situation arises in Section (i) from the unrealistic postulate of empty space at infinity.

In addition it should be pointed out that the metric of general relativity arises from a general transformation of the metric of special relativity, which was based on the real existence of a quasi-Galilean reference system extending at least over the whole solar system. If this quasi-Galilean frame owes its existence in fact to world gravitation, as the experiment of the Foucault pendulum, etc. would imply, then it is not surprising that we do not obtain consistency if we postulate the absence of world gravitation. We can of course arrive at the notion of empty space as a limiting process by letting the cosmic density tend to zero. We shall see in Chapter IV that, in the case of the steady state model at least, this would be compensated by the fact that the radius of space would tend to infinity in such a way that inertia can always be accounted for.

In the next Chapter we shall put forward an analysis to show that general relativity permits a theory of Mach's Principle which, in the first approximation, is entirely consistent in essence with that advanced by Sciama.

CHAPTER III: GENERAL RELATIVITY AS AN EXPRESSION OF MACH'S PRINCIPLE (2)

(i) The inductive effect of the universe in general relativity

If we reject as inadmissible in general relativity the consideration of the hypothetical situation of empty space at infinity, we have to look upon Einstein's quasi-Galilean solution, presented in Chapter II, Section (i), as applicable to a reference frame which, over an extensive region, would be inertial in the field of the distant matter of the universe but is disturbed from this state in the neighbourhood of the spatial origin by concentrated matter there. Thus the integrals giving the $\gamma_{\mu\nu}$ in (2.9) would be taken over this concentrated matter only. At much greater distances the reference frame would cease to be quasi-Galilean and deviations $\gamma_{\mu\nu}$, not necessarily remaining small, would begin to appear depending now on the <u>cosmic</u> distribution of mass and its motion. On such a scale of course <u>local</u> disturbances of space-time would cease to be important.

Accordingly, if local concentrated matter has the effect on the $g_{\mu\nu}$ indicated by equations (2.9), and on the motion of a free particle as indicated in the pondermotive equation (2.12) (whose interpretation in accordance with Mach's Principle is given in Chapter II, Section (ii)), we now enquire to what extent and in what manner do the $g_{\mu\nu}$ and inertia depend on the cosmic distribution of mass. This is a question that Einstein did not answer, beyond showing that in the Einstein universe the $g_{\mu\nu}$ depended on the mass density of that universe, via the radius of closed space which entered explicitly into the $g_{\mu\nu}$.

It is natural to look for an analytic theory of these cosmic effects that will give expression to the qualitative aspects of world gravitation which we have shown, in Chapter I, Section (ii), to be consistent with the Principle of Equivalence. Furthermore we look to this theory to be formally consistent with the inertial effects, demonstrated by Einstein to be inherent in general relativity, of concentrated matter in a quasi-Galilean region, but now in relation to the distribution of mass in the universe as a whole, and taking account of its motion relative to the particular reference frame adopted.

Accordingly we investigate the extent to which we may generalise the circumstances when the motion of a free particle may be described by a Maxwell type pondermotive equation. For this purpose we make the assumptions less restrictive than in the quasi-Galilean case considered by Einstein, as follows:

(a) The particle velocity $\underline{v}$ in the reference frame is assumed small such that v^2/c^2 is negligible compared with v/c.

(b) The velocities of the sources of the field in the region of space time in the neighbourhood of the particle event are also small of the same order, so that the same remark applies as in (a). This does not restrict the velocities of sources outside the region of interest, and such sources will of course influence the field in the region of interest.

(c) The deviations of the $g_{\mu\nu}$ from the Galilean values are small in the region of interest, such that their squares and products and those of their derivatives can be neglected. We do not however assume that the deviations vanish at 'infinity', nor that they even remain small outside the specified range.

It is clear from equation (2.5) that the equation of motion of a free particle can in these circumstances be written

$$\frac{dv^i}{dt} = -\frac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{\partial}{\partial t}(-g_{4i}) + \left\{\frac{\partial}{\partial x^h}(g_{4i}) - \frac{\partial}{\partial x^i}(g_{4h})\right\}v^h \quad \ldots.(3.1)$$

for $i = 1,2,3$. Comparing with (2.5) it is to be noted that we have omitted the coefficient of v^i on the left hand side since its time derivative will be of the same order as the velocity of the particle, and so the corresponding term in the expansion of the left hand side is neglected by conditions (a), (b). On the other hand the coefficient of v^h on the right hand side is retained since it is of the dimensions of an <u>angular</u> velocity of the field sources, so that the whole term is an acceleration which in general is of the same order as the left hand side.

The equations (3.1) are generally covariant in the sense that, in all reference frames and regions of space-time which do not violate the assumptions above, they describe the space motion of a free particle in terms of the derivatives of the $g_{\mu\nu}$ involved. We now generalise the quantities $\underline{A}$, ϕ occurring in the quasi-Galilean analysis by defining

$$\left(\underline{A}, \Phi\right) \equiv \left(-g_{41}, -g_{42}, -g_{43}, \tfrac{1}{2}g_{44}\right) \quad \ldots\ldots.(3.2)$$

The three equations in (3.1) may then be written concisely

$$\frac{d\underline{v}}{dt} = -\operatorname{grad}\Phi - \frac{\partial \underline{A}}{\partial t} + \underline{v} \wedge \operatorname{curl}\underline{A} \qquad \ldots\ldots (3.3)$$

The vector notation implies the vector character of the terms for purely spatial transformations. For space-time transformations however the quantities ($\underline{A}$, Φ) do not transform as a 4-vector but as components of the tensor $g_{\mu\nu}$. This is because unlike the corresponding electromagnetic pondermotive equation, the permitted transformations are not necessarily between inertial frames and therefore not in general linear.

It is to be noted that here we have not as in the quasi-Galilean case identified $\underline{A}$, Φ with the deviations from their Galilean values of the $g_{\mu\nu}$ involved, but, consistent with our endeavour to account for the whole of inertia according to Mach's Principle, in terms of the total $g_{\mu\nu}$. The covariance of (3.3) is secured by the tensor character of the total $g_{\mu\nu}$ involved; the deviations do not transform as tensors for general transformations. Indeed according to the field equations (1.4) it is the total $g_{\mu\nu}$ field that is related inseparably to the distribution of mass in the whole universe.

Bearing in mind therefore the physical interpretation of the quantities $\underline{A}$, ϕ in the quasi-Galilean case we should expect analogous interpretation of $\underline{A}$, Φ in (3.3) which would, if Mach's Principle is to be satisfied, take account of the distribution and motion of matter in the whole universe, relative to the particular reference frame being used. It would be an immediate consequence of such an interpretation that the 'fictitious' forces of Newtonian mechanics in accelerating or rotating reference frames would become directly attributable to the inductive effect of a moving universe.

In particular in a reference frame in which a freely moving particle was permanently at rest equation (3.3) would reduce to

$$-\operatorname{grad}\Phi - \frac{\partial \underline{A}}{\partial t} = 0, \qquad \ldots\ldots(3.4)$$

holding at the particle. This is the equation postulated by Sciama. To use Sciama's expression the 'gravo-electric' field of the whole universe would be zero at the particle and it would be gravitationally 'free' in its own rest frame. Thus equation (3.4) demonstrates that general relativity permits the same interpretation of Mach's Principle as is givem by Sciama - that the Newtonian attraction of local matter is cancelled in the rest frame of a free particle by the inductive effect of a moving universe. Indeed (3.4) gives analytic expression to the qualitative interpretation of the Principle of Equivalence, consistent with Mach's Principle, which was given In Chapter I, Section (ii).

The choice of reference frames co-moving with a free particle is infinite. However general relativity allows the particular choice of a geodesic local coordinate system, with spatial origin permanently located at the particle, in which for an infinitesimal region of space-time containing the particle event the $g_{\mu\nu}$ take the Galilean values of special relativity. Thus in this region the metric is

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \qquad \ldots\ldots\ldots(3.5)$$

In this system the coordinates represent locally the natural measures of space and time by an observer moving with the particle, and we note that the terms in (3.4) now vanish individually at the particle, on account of the stationary values of the $g_{\mu\nu}$ at the particle event.

In this thesis we attach importance for Mach's Principle to the total $g_{\mu\nu}$ and not just their derivatives. Justification of this has already been given in connection with equation (3.2) and will be given further in the physical interpretation of ($\underline{A}$, Φ) dealt with in Chapter IV. It follows from (3.2) that the static potential Φ in the field of the whole universe, at the spatial origin of a frame whose metric in the neighbourhood is given by (3.5), would be

$$\left.\begin{aligned} \Phi_0 &= \tfrac{1}{2} g_{44}(0) = \tfrac{1}{2} \\ &= \tfrac{c^2}{2} \text{ in general units} \\ \text{while} \qquad \underline{A}_0 &= 0 \end{aligned}\right\} \ldots(3.6)$$

The dimensions of Φ and the significance we are trying to associate with it would require Φ_0 to be of order $-GM/R$, where M is the effective gravitational mass of the universe and R its effective radius. Sciama's approach is to <u>define</u> what corresponds to our Φ_0 by the expression $-\int_{r=0}^{r=cT} \frac{\sigma\, dV}{r}$ as in equation (1.6), where σ is the gravitational mass density in the universe. As in (1.9) he gets $G\Phi \doteq -c^2$. Both results are seen to be of the same order, noting that Sciama omits G from his definition of Φ. The discrepancy in sign will occupy us later (Chapter IV) when we investigate to what extent general relativity justifies this tentative physical interpretation of $\underline{A}, \Phi$. We first give some applications of the foregoing analysis.

(ii) Applications of the inductive theory in general relativity

(a) Sciama considers the case of a free particle in rectilinear motion in the gravitational field of a mass M which is at rest relative to the 'smoothed out' universe. His analysis and results were briefly outlined in Chapter I, Section (iii). We shall now consider the same problem according to general relativity.

If we choose a reference frame at rest relative to the smoothed out universe, assumed homogeneous and isotropic, in the form of the well known cosmological reference systems of co-moving coordinate mesh, then the metric will be of the general form (13).

$$ds^2 = dt^2 - R^2(t)\frac{\{(dx')^2 + (dy')^2 + (dz')^2\}}{(1 + k(r')^2/4)^2} \qquad \ldots\ldots(3.7)$$

Here $(r')^2 = (x')^2 + (y')^2 + (z')^2$, k is a constant taking the values $1, 0,$ or -1 according to the geometrical nature of the universe, and $R(t)$ is a function of cosmic time which also depends on the nature of the model. As has been shown in Part I, Chapter I of this thesis, where the properties of the metric (3.7) have been examined in detail, this metric can be identified for the neighbourhood of any 'fundamental' particle of the model (having constant x', y', z') with the Galilean (inertial) metric given by (3.5). That is t is the naturally measured time on a fundamental particle, and the naturally measured spatial distances at time t corresponding to coordinate increments dx', dy', dz' will be $dx = R(t)dx'/(1+kr'^2/4)$ etc.

Accordingly, any fundamental observer at rest relative to the smoothed out universe can describe local events in terms of the metric (3.5), just as in the case of an observer on any free particle, even although on the cosmological scale there may be stress present. As pointed out in Part I, Chapter I this is because such stress must necessarily be uniform in the smoothed out models of metric (3.7), and the world lines of fundamental particles are consequently geodesics - an aspect of Mach's Principle, in view of our analysis of the geodesic world line.

Let us now imagine a concentration of mass M at the spatial origin of a fundamental observer's local frame at cosmic epoch t. In a sufficiently small neighbourhood we can neglect the deviations from the Galilean values of the $g_{\mu\nu}$, as far as they arise from the universe as a whole, and include only the deviations due to the mass M. Thus to this approximation the metric can be taken as

$$ds^2 = \left(1 - \frac{2GM}{r}\right)dt^2 - \left(1 + \frac{2GM}{r}\right)\left(dx^2 + dy^2 + dz^2\right) \qquad .....(3.8)$$

This follows from Einstein's approximate solution for a metric that is Galilean at 'infinity' given by (2.9), or directly from Schwarzschild's solution based on the same assumption, and taken to the same order.

It is to be noted that, according to the ideas presented in this thesis, the contribution to the $g_{\mu\nu}$ 'potentials' from the universe as a whole is present in the Galilean terms of the $g_{\mu\nu}$. Since the universe is at rest in this frame we have by (3.2)

$$\left.\begin{aligned} \underline{A} &= 0 \\ \text{while} \quad \Phi &= \tfrac{1}{2}\left(1 - \frac{2GM}{r}\right) \end{aligned}\right\} \qquad \text{..................(3.9)}$$

Suppose a particle is moving freely towards the mass M along the x axis. If its space coordinates are $(x_1, 0, 0)$ at coordinate time t then its coordinate speed is $dx_1/dt = -v$, where $v>0$. Make now the transformation to a suitable rest frame for the particle, by means of the relations

$$x = X + x_1 \,,\ y = Y \,,\ z = Z \,,\ t = T \qquad \text{.......(3.10)}$$

yielding $dx = dX - v\,dT$, $dy = dY$, $dz = dZ$, $dt = dT$. We get therefore to sufficient order for the covariance of equation (3.3)

$$ds^2 = \left(1 - v^2 - \frac{2GM}{r}\right)dT^2 + 2v\,dX\,dT - \left(1 + \frac{2GM}{r}\right)\left(dX^2 + dY^2 + dZ^2\right) \quad \text{..(3.11)}$$

Thus in the particle's rest frame

$$\left.\begin{aligned} \underline{A} &= (-v, 0, 0) \\ \Phi &= \tfrac{1}{2}\left(1 - v^2 - \frac{2GM}{r}\right) \end{aligned}\right\} \qquad \text{..................(3.12)}$$

Apply now equation (3.4) in the particle's rest frame, yielding

$$-\frac{\partial}{\partial X}\left\{\frac{1}{2}\left(1-v^2-\frac{2GM}{r}\right)\right\} - \frac{\partial}{\partial T}\left(-v\right) = 0 \qquad \ldots\ldots\ldots\ldots\ldots(3.13)$$

leading to

$$-\frac{GM}{r^2} + \frac{dv}{dt} = 0 \qquad \ldots\ldots\ldots\ldots\ldots(3.14)$$

on substituting the original coordinates. This is the Newtonian equation of motion of the particle and is also the equation which would follow from the general pondermotive equation (3.3), applied to the original frame, by (3.9).

On examining (3.12) and (3.13) we see that the origin of the inertial term $-\frac{\partial(-v)}{\partial T}$ in (3.13) lies in the relative motion of the universe, yielding $\underline{A} = (-v, 0, 0)$ in the particle's rest frame, and thus creating at the particle an inductive field which balances the local attraction due to the mass M, thus connecting with Sciama's ideas.

We note also that the $\underline{A}$, Φ in (3.12) arise by transformation of the <u>whole</u> $g_{\mu\nu}$ and not just their deviations from the Galilean values, in accordance with our tentative interpretation of the Galilean values as the static 'potentials' of the whole universe.

(b) The other case considered by Sciama is that of a particle moving with uniform motion in a circle under the attraction of a mass M at the centre, this mass being again at rest relative to the smoothed out universe.

Transform therefore from the metric (3.8) to a suitable rest frame for the particle according to the relations

$$\left.\begin{aligned} x &= X\cos\omega T - Y\sin\omega T \\ y &= Y\cos\omega T + X\sin\omega T \\ z &= Z \\ t &= T \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots(3.15)$$

so that to sufficient order, putting $X^2+Y^2=R^2$,

$$ds^2 = \left(1-\frac{2GM}{R}-\omega^2R^2\right)dT^2 - 2\omega\left(-Y\,dX\,dT + X\,dY\,dT\right) - \left(1+\frac{2GM}{R}\right)\left(dX^2+dY^2+dZ^2\right) \qquad \ldots\ldots\ldots\ldots\ldots(3.16)$$

Thus in this frame

$$\left.\begin{aligned} \underline{A} &= \left(-\omega Y, \omega X, 0\right) \\ \Phi &= \tfrac{1}{2}\left(1-\frac{2GM}{R}-\omega^2R^2\right) \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots(3.17)$$

The equation (3.4) then yields

$$-\frac{GM}{R^2} + \omega^2 R = 0 \qquad \ldots\ldots\ldots\ldots\ldots(3.18)$$

which is the Newtonian equation of motion, and incidentally, putting $R=r$, what would be given by (3.3) when applied to the original frame, using (3.9).

Connecting with Sciama's ideas we say that the gravitational attraction by M is balanced by the gravitational field induced by a rotating universe, giving rise to the 'vector potential' indicated by $\underline{A}$ in (3.17).

(c) As a final example we shall show how, by means of the covariance of (3.3), the Newtonian 'fictitious' forces may be attributed to the inductive effect of a moving universe, in the most general Newtonian motion of the reference frame relative to a locally inertial frame.

Consider a free particle in a reference frame which is at rest relative to the smoothed out universe. Then by a suitable transformation to local natural coordinates we can as before describe <u>neighbouring</u> events in terms of the metric (3.5), assuming the absence of local concentrated matter. Let $\underline{r}$ be the position vector of the particle relative to this locally inertial frame. Then by (3.5), (3.2), (3.3) we have

$$\ddot{\underline{r}} = 0 \qquad \text{.........................(3.19)}$$

Transform to a second frame whose space origin has variable velocity $\underline{V}$ and which has variable spin $\underline{\omega}$ relative to the first frame. If the position vector of the particle in this frame is $\underline{R}$ then a well known kinematic result of Newtonian motion gives

$$\dot{\underline{r}} = \underline{V} + \dot{\underline{R}} + \underline{\omega} \wedge \underline{R} \qquad \text{.........................(3.20)}$$

$$\ddot{\underline{r}} = \dot{\underline{V}} + \underline{\omega} \wedge \underline{V} + 2\underline{\omega} \wedge \dot{\underline{R}} + \dot{\underline{\omega}} \wedge \underline{R} + \underline{\omega} \wedge (\underline{\omega} \wedge \underline{R}) + \ddot{\underline{R}} \qquad \text{......(3.21)}$$

differentiation being with respect to the common Newtonian time of either frame. Thus for the particle in the second frame

$$\ddot{\underline{R}} = -\left[\dot{\underline{V}} + \underline{\omega} \wedge \underline{V} + 2\underline{\omega} \wedge \dot{\underline{R}} + \dot{\underline{\omega}} \wedge \underline{R} + \omega \wedge (\underline{\omega} \wedge \underline{R})\right] \qquad \text{.....(3.22)}$$

on using (3.19). It is to be noted that in this equation $\underline{V}, \dot{\underline{V}}, \underline{\omega}, \dot{\underline{\omega}}$ are as measured relative to the non-inertial frame.

The transformation connecting the two frames is by (3.20) in differential form

$$\left.\begin{aligned} d\underline{r} &= (\underline{V} + \underline{\omega} \wedge \underline{R})\,dT + d\underline{R} \\ dt &= dT \end{aligned}\right\} \qquad \text{..................(3.23)}$$

Hence

$$ds^2 = dt^2 - d\underline{r}^2$$

$$= \left[1 - \underline{V}^2 - 2\underline{V}\cdot(\underline{\omega}\wedge\underline{R}) - (\underline{\omega}\wedge\underline{R})^2\right] dT^2 - 2\left(\underline{V} + \underline{\omega}\wedge\underline{R}\right)\cdot d\underline{R}\, dT - d\underline{R}^2 \quad \ldots (3.24)$$

Thus in the non-inertial frame

$$\left.\begin{aligned} \underline{A} &= \underline{V} + \underline{\omega}\wedge\underline{R} \\ \Phi &= \tfrac{1}{2}\left[1 - \underline{V}^2 - 2\underline{V}\cdot(\underline{\omega}\wedge\underline{R}) - (\underline{\omega}\wedge\underline{R})^2\right] \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (3.25)$$

Now

$$\operatorname{grad}\Phi = \underline{\omega}\wedge\underline{V} + \underline{\omega}\wedge(\underline{\omega}\wedge\underline{R})$$

$$\frac{\partial \underline{A}}{\partial T} = \frac{\partial \underline{V}}{\partial T} + \frac{\partial \underline{\omega}}{\partial T}\wedge\underline{R}$$

$$= \dot{\underline{V}} + \dot{\underline{\omega}}\wedge\underline{R}$$

while $\operatorname{curl}\underline{A} = 2\underline{\omega}$

Hence by (3.3)

$$\ddot{\underline{R}} = -\left[\dot{\underline{V}} + \underline{\omega}\wedge\underline{V} + 2\underline{\omega}\wedge\dot{\underline{R}} + \dot{\underline{\omega}}\wedge\underline{R} + \underline{\omega}\wedge(\underline{\omega}\wedge\underline{R})\right] \quad \ldots\ldots (3.26)$$

giving complete agreement with (3.22).

Thus our theory gives an exact treatment of the fictitious forces as the inductive effect of a moving universe.

CHAPTER IV: THE STEADY STATE UNIVERSE AS A SOLUTION OF THE FIELD EQUATIONS INCORPORATING MACH'S PRINCIPLE

(i) The cosmic distribution of matter

According to our analytic theory, developed in the last Chapter, inertia and the inertial frames appear to be determined in general relativity in accordance with Mach's Principle, provided that $\underline{A}, \Phi$ occurring in equations (3.3), (3.4) can be physically interpreted in terms of the distribution and motion of mass in the whole universe.

It would be natural to look for physical identifications of $\underline{A}, \Phi$, defined in terms of the total $g_{\mu\nu}$ by (3.2), which are similar to those found by Einstein in terms of local matter for the $\underline{A}, \phi$ arising in the quasi-Galilean case, where they were defined in terms of the $\gamma_{\mu\nu}$. This of course implies, as was indicated in Chapter III, that in some way the Galilean terms themselves are related to world gravitation. Indeed, for the particular case when the $\underline{A}, \Phi$ are evaluated at the space origin of a reference frame of locally Galilean metric, we found the $\underline{A}, \Phi$ dependent on the values of the Galilean $g_{\mu\nu}$ only, viz. $\underline{A}_0 = 0$ and $\Phi_0 = c^2/2$ equation 3.6)). We assume that these special values are related to the fact that, for an observer permanently located at the space origin, the coordinates represent the naturally measured proper distance and proper time which were discussed in Chapter I, Part I of this thesis.

As has been pointed out in detail in Part I of this thesis the simplest possible assumption consistent with observation is that the universe in the large consists of galaxies (of stars) whose distribution has a statistically homogeneous and isotropic character; furthermore this system is in the process of homogeneous expansion. The metric of such universes, assuming the contents smoothed out to an ideally homogeneous state has already been given (equation (3.7)), and was discussed in detail in Part I.

As was pointed out in connection with (3.7) this metric can be identified in the neighbourhood of any spatial point (x', y', z') with a Galilean metric, the coordinate t being in fact the naturally measured (proper) time of all fundamental observers. Accordingly we identify the values of $\underline{A}_0$ and Φ_0, found appropriate in (3.6) to the space origin of any locally Galilean frame, with the values appropriate to the space origin of the world metric (3.7). Furthermore it will be our purpose to interpret these values in terms of the distribution of mass, and its motion, in that particular universe among those given by (3.7) which seems the most satisfactory from the point of view of Mach's Principle. In doing so we shall adopt what seems the most appropriate and physically meaningful extension of local proper coordinates, namely cosmic time t, and integrated proper distance measured by the aggregate of fundamental observers in the simultaneity of cosmic epoch t.

A knowledge of the undetermined parameters in the models of metric (3.7), namely the expansion factor $R(t)$, whose reciprocal squared represents the magnitude of the curvature of the subspace $t =$ constant, and k which gives the sign of the curvature, is in effect a knowledge of the boundary conditions of the universe. For these quantities determine the rate of expansion of space, whether this is accelerating or retarding, whether space is open or closed, and according to the field equations they together determine the density and pressure of the smoothed out universe.

It seems possible therefore that the satisfactory incorporation of Mach's Principle in general relativity may be bound up with the determination of these parameters for the actual universe. Indeed it may well be that by a satisfactory theory of Mach's Principle these boundary conditions may be inferred. It is of course conceivable that by an automatic adjustment of the radius of space to cosmic density _any_ cosmological distribution of matter is consistent with Mach's Principle, in that inertia can be satisfactorily accounted for in terms of this distribution, and its motion.

In this connection, however, it must be remembered that the constants occurring in the field equations, namely G and c, are derived from local experimental evidence for the present cosmic epoch only, but are nevertheless assumed to be independent of epoch when cosmological solutions involving changing universes are considered. It seems very likely that many such dimensional 'constants' would in fact be a function of epoch in a changing universe. In this case the solution of the cosmological problem might well become impossible. For these reasons it would appear simplest and logically most appropriate to regard the steady state model, first proposed by H. Bondi and T. Gold (10), as providing the unique cosmological boundary conditions in general relativity in its present form.

Bondi and Gold introduced the steady state model for the reasons of simplicity and logical economy mentioned above, but at the time did not regard general relativity as a relevant field theory because of the apparent non conservation of energy, due to their postulated spontaneous creation of matter to maintain the steady state. The steady state model was proposed independently by F. Hoyle (11) and in his theory the rate of creation of matter, and the rate of expansion, were made dependent on an additional term introduced into the field equations of general relativity. In Hoyle's model the cosmic pressure was taken to be zero. By introducing the concept of a negative cosmic stress W. H. McCrea (12) was able to show that this stress performed the same function mathematically in the conventional field equations as Hoyle's modifying term. In addition the creation of matter could be regarded as the materialisation of the work performed by this stress in the expansion of the universe.

We therefore propose to examine the steady state model according to McCrea's interpretation, with a view to finding a physical justification of our analytic theory of Mach's Principle in general relativity.

(ii) Physical identification of A, Φ in the steady state universe

The metric of the steady state model is that of the de Sitter universe which in the Robertson-Lemaitre form (14) is

$$ds^2 = c^2 dt^2 - e^{2t/T}\left\{(dx')^2 + (dy')^2 + (dz')^2\right\} \qquad \text{.............}(4.1)$$

or in polars

$$ds^2 = c^2 dt^2 - e^{2t/T}\left\{(dr')^2 + (r')^2 d\theta^2 + (r')^2 \sin^2\theta \, d\phi^2\right\} \qquad \text{........}(4.2)$$

Here c is the velocity of light, returning to general units, and T is a constant which will be identified below in terms of observable characteristics of the model. The model is clearly that one among the general models of metric (3.7) got by putting $R(t) = e^{t/T}$, and $k = 0$. The geometry of the subspace $t = t_0$ (constant) is thus Euclidean.

According to the expression giving the Hubble parameter for the expanding models of general metric (3.7), derived in Part I (equation (1.14)), we see that in the case of the steady state model the reciprocal of the Hubble parameter is the constant T. For the sake of clarity in the subsequent work of this Chapter we reproduce here the analysis for the steady state.

The integrated proper distance from the space origin of the frame of metric (4.2) to the fundamental particle at (r', θ, ϕ), as measured in the simultaneity of cosmic time t by the aggregate of fundamental observers on the radial line joining these two points, is

$$\ell = e^{t/T}\int_0^{r'} dr' = r' e^{t/T} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(4.3)$$

The rate of increase of ℓ for this particle is

$$\dot{\ell} = r' e^{t/T}/T = \ell/T \qquad \ldots\ldots\ldots\ldots\ldots\ldots(4.4)$$

From this we see that $1/T$ is the Hubble parameter of the model as derived locally by any fundamental observer since, as was explained in Part I, Chapter I, Section (v), for near objects the distance ℓ can be identified with distance as measured by astronomers.

From (4.4) we note that when

$$\dot{\ell} = c \qquad \ldots\ldots\ldots\ldots\ldots\ldots(4.5)$$

then $\ell = cT = \bar{R}$ (say)

This distance is in fact the integrated proper distance to the horizon of the model at all epochs. An object beyond this distance at the time of emission of its light can never be seen by the fundamental observer at the space origin. This property of the model was carefully examined in Part I, Chapter III (in connection with equation (3.12)). There it was pointed out that when the distance at the time of emission approached the distance $\bar{R}$ then, for a finite time of observation t_0, the time of emission tended to $-\infty$. In addition it was noted that the total population of objects seen at the space origin at time t_0 corresponded in fact to the total range of the coordinate r' of the metric (4.2).

The importance for Mach's Principle is that the horizon distance $\bar{R}$ is related to σ the density of gravitational mass, in the model. In general relativity gravitational mass density in a continuous distribution of mass is defined so as to lead to the Gauss flux theorem for small regions of space (see for example J. L. Synge (15)). For the isotropic cosmological models $\sigma = \rho + 3p/c^2$, where ρ is the proper inertial mass density as measured by fundamental observers and p is the proper pressure. The gravitational 'force' on unit mass which is involved in the flux theorem has to be identified, in the case of these world models, with the relative proper acceleration $\ddot{\ell}$.

For the general model of metric (3.7) the field equations of general relativity, taking the cosmological constant to be zero, yield for ρ and p (16),

$$8\pi G\rho = \frac{3\dot{R}^2}{R^2} + \frac{3kc^2}{R^2} \qquad \text{.............(4.6)}$$

$$-\frac{8\pi Gp}{c^2} = \frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2} \qquad \text{.............(4.7)}$$

Where $R \equiv R(t)$ and dots indicate differentiation. Thus in the steady state for which $R(t) = e^{t/T}$ and $k = 0$ we obtain

$$\rho = -\frac{p}{c^2} = \frac{3}{8\pi GT^2} \qquad \text{.............(4.8)}$$

so that both inertial mass density and the pressure are constant in the steady state. The negative pressure is the zero-point stress introduced by McCrea. It follows from (4.8) that the density of gravitational mass in the model is negative, viz.,

$$\sigma = \rho + \frac{3p}{c^2} = \frac{-3}{4\pi C T^2} \qquad \text{..........(4.9)}$$

By differentiating (4.4) we obtain

$$\ddot{\ell} = \frac{\ell}{T^2} \qquad \text{..........(4.10)}$$

and hence using (4.9)

$$\ddot{\ell} = -\tfrac{4}{3}\pi G \sigma \ell \qquad \text{..........(4.11)}$$

This equation is identical in form with the Newtonian gravitational field equation in a sphere of constant density σ.

Comparing (4.11) with the pondermotive equation (3.3) we see that, for a fundamental observer whose radial space coordinate is the proper distance ℓ and whose coordinate time is the cosmic time t, we obtain

$$-\frac{d\Phi}{d\ell} = -\tfrac{4}{3}\pi G \sigma \ell \qquad \text{..........(4.12)}$$

We note however that while in the case of the pondermotive equation (3.3), when $\underline{A}$ and Φ are defined in terms of the $g_{\mu\nu}$ by (3.2), the validity of the equation is restricted to a region of space-time in which the deviations of the $g_{\mu\nu}$ from the Galilean values are small, the equation (4.11) on the other hand is true for all ℓ. Consequently so is (4.12). We assume that this difference can be attributed to the physical significance of proper distance and cosmic time employed in (4.12). It is clear however that any significance physically which we can attach to Φ in (4.12) must be attached also to Φ defined by (3.2), since when ℓ is small it becomes the radial space coordinate of the fundamental observer's locally Galilean metric.

Since, as we have shown, the proper distance $\bar{R} = cT$ constitutes an effective radius of the model as regards optical **effects** we shall assume that this distance constitutes also an effective limit of gravitational influence in the model. Consider therefore, in analogy with Newtonian gravitational theory, the 'work' done by the gravitational field of the steady state universe in moving, under its expansion, a particle of unit mass from a position at proper distance ℓ from the fundamental observer to his optical and gravitational horizon. This will be

$$\Phi_\ell = -\tfrac{4}{3}\pi G\sigma \int_\ell^{\bar{R}} \ell \, d\ell$$

$$= -\tfrac{2}{3}\pi G\sigma \bar{R}^2\left(1 - \frac{\ell^2}{\bar{R}^2}\right)$$

which by (4.9) is $\quad \dfrac{c^2}{2}\left(1 - \dfrac{\ell^2}{\bar{R}^2}\right)$ ....(4.13)

This quantity may be regarded as the analogue of the Newtonian potential at distance ℓ from the centre of a sphere of constant density σ and radius $\bar{R}$, isolated in otherwise empty space. Putting $\ell = 0$ we obtain

$$\Phi_0 = \frac{c^2}{2} \qquad \text{..................(4.14)}$$

which therefore provides, in a natural way, a physical identification of the Φ of our analytic theory in Chapter III when evaluated at the space origin of a Galilean frame (equation (3.6)).

Unlike the rest of the general models of metric (3.7) the steady state model allows a static metric to be used which is Galilean at the space origin without further change of coordinates. Thus the motion of a particle in the neighbourhood of the space origin can be described in terms of the static coordinates in accordance with the pondermotive equation (3.3). The metric is the well known form of the de Sitter metric

$$ds^2 = c^2\left(1-\frac{\ell^2}{\bar{R}^2}\right)d\bar{t}^2 - \frac{d\ell^2}{1-\frac{\ell^2}{\bar{R}^2}} - \ell^2 d\theta^2 - \ell^2 \sin^2\theta\, d\phi^2 \qquad \text{.......(4.15)}$$

connected to (4.2) by the Robertson-Lemaitre transformation. Here $\bar{R} = cT$ as in (4.5), and the coordinate ℓ of the event $(\ell, \bar{t})$ is in fact the integrated proper distance of our previous analysis between the fundamental particle at the space origin and the particle at the event $(\ell, \bar{t})$, evaluated at the cosmic time t of the event.

In terms of this metric the Φ of our analysis in Chapter III, defined by (3.2), will be

$$\Phi = \frac{c^2}{2}\left(1-\frac{\ell^2}{\bar{R}^2}\right) \qquad \text{......................(4.16)}$$

which agrees with the Φ_ℓ of (4.13) and therefore allows the physical interpretation associated with (4.13). Since also $\underline{A} = 0$ for this metric the pondermotive equation (3.3) yields for the region of space in which ℓ is small

$$\frac{d^2\ell}{d\bar{t}^2} = \frac{\ell}{T^2} \qquad \text{.......................(4.17)}$$

This agrees with (4.10) for small ℓ since then t becomes equal to $\bar{t}$.

We shall now show that the value of $\Phi_0 = c^2/2$ given in (3.6) and identified by (4.14) with the potential energy of unit mass in the field, relative to a fundamental observer, of the whole universe, can also be identified with a Newtonian type potential integral over the universe of matter within the horizon of such an observer. This was envisaged in Chapter III, in connection with (3.6), as a natural extension of the result found for ϕ in terms of matter in a quasi-Galilean frame (equation (2.10)).

For the steady state σ and $\bar{R}$ are constants. We may therefore write for the potential energy in the field of unit mass at the space origin, of the cosmic mass in the volume element dV at coordinate ℓ

$$d\Phi_0 = -G\int_{\ell}^{\bar{R}} \frac{\sigma\, dV}{\ell^2}\, d\ell \qquad \text{...............(4.18)}$$

Hence $$\Phi_0 = -G\sigma\int_{\ell=0}^{\ell=\bar{R}}\left(\frac{1}{\ell} - \frac{1}{\bar{R}}\right)dV \qquad \text{...............(4.19)}$$

Here we have assumed a Newtonian inverse square law which is justified by the Newtonian type field equation (4.11), and also more generally by W. H. McCrea's Newtonian cosmology, which is based on proper distance ℓ and cosmic time t, and leads to the same field equations (4.6), and (4.7), as are given by general relativity (<u>12</u>). Equation (4.19) may be regarded as the extension of formula (2.10) to a cosmic range of integration in the steady state universe. On performing the integration we derive

$$\Phi_0 = -G\sigma\int_0^{\bar{R}}\left(\frac{1}{\ell}-\frac{1}{\bar{R}}\right)4\pi\ell^2\,d\ell$$

$$= \frac{c^2}{2}, \text{ by } (4.9). \qquad \text{.........................(4.20)}$$

This result is in agreement with (4.14) and justifies our interpretation of Φ_0, given by (3.6), as the gravitational potential of unit mass at the space origin of a fundamental observer's locally Galilean frame, in the field of all the mass in the universe which has gravitational influence on, or is apparent to, such an observer.

The quantity $\underline{A}$ of our analytic theory in Chapter III, defined by (3.2), is zero for the metrics of the steady state that we have discussed. Since however, in the region where ℓ is small, the metric (4.15) is approximately Galilean we can, on changing to Cartesian space coordinates, apply the transformation employed in the first example of Section (ii), Chapter III. In that example we showed that a non-zero $\underline{A}$ arises, by transformation of the $g_{\mu\nu}$, in the rest frame of a particle moving with speed v along the common x, X axes towards the origin of the Galilean frame. In fact in deriving $\underline{A}$ in the particle's rest frame we had to multiply the Galilean g_{11} in the original frame, viz. $g_{11}(0) = -1$, by v in the course of the transformation of the $g_{\mu\nu}$. This led to, restoring general units,

$$\underline{A} = (-cv, 0, 0) \qquad \text{......................(4.21)}$$

in the particle's rest frame, which was interpreted as the 'vector' potential arising from the relative motion with speed v of the smoothed out universe.

According to the subsequent application of the pondermotive equation in this example the quantity $-g_{11}(0)$ turns out to be proportional to the inertial mass of the particle in its motion relative to the Galilean frame. From the point of view of the particle's rest frame the inertial term in the equation of motion was interpreted as due to the inductive effect of the moving universe. Can we in fact justify this physical interpretation of $\underline{A}$ in (4.21) as we have done for Φ ?

In this connection we recall the result for $\underline{A}$ in the quasi-Galilean case given by (2.10). Associating $\underline{A}$ with inertial mass rather than gravitational mass, we tentatively adopt the following analogue of (2.10) for the $\underline{A}$ arising from the relative motion, with velocity $\underline{v}$, of the smoothed out universe in the steady state

$$\underline{A} = -\frac{4G}{c}\int_{l=0}^{l=\bar{R}} \rho\left(\frac{1}{l} - \frac{1}{\bar{R}}\right)\underline{v}\, dV \qquad \text{...........(4.22)}$$

$$= -\frac{4G\rho\underline{v}}{c}\int_0^{\bar{R}}\left(\frac{1}{l} - \frac{1}{\bar{R}}\right)4\pi l^2\, dl$$

$$= -c\,\underline{v} \qquad \text{by (4.8),................(4.23)}$$

which agrees with (4.21).

We conclude therefore that the analytic theory of Mach's Principle in general relativity, presented in Chapter III, can in fact be substantiated in terms of the distribution and motion of mass in the steady state universe.

(iii) Comparison with Sciama's theory

In his paper, already referred to, Sciama claimed that his theory of the origin of inertia differed from general relativity principally in the following respects:

(a) It enabled the amount of matter in the universe to be estimated from a knowledge of the gravitational constant, whereas no such estimate was afforded by general relativity.

(b) The Principle of Equivalence was a consequence of his theory and not an initial axiom.

(c) It implied that gravitation must be attractive, whereas in general relativity the sign of the field was not determined.

We shall discuss these claims in turn:

(a) We have indicated in Chapter I, Section (iii) that in Sciama's theory the possibility of estimating the amount of matter in the universe arises from his equation $G\sigma\tau^2 \simeq 1$ (equation (1.9)), by which a knowledge of G, and of Hubble's constant τ, would lead to a knowledge of σ the average gravitational mass density in the universe. However, as we pointed out in Chapter I, Sciama tacitly assumes that cosmic gravitational mass density is equal to inertial mass density, that is, in our notation, he assumes $\sigma = \rho$. Consequently σ occurs not only in his definition of Φ, but also in his definition of $\underline{A}$ (equation (1.6)). This allows him to write for the inductive force due to the relative motion, with velocity $\underline{v}$, of the smoothed out universe,

$$-\frac{1}{c}\frac{\partial \underline{A}}{\partial t} = -\frac{\Phi}{c^2}\frac{d\underline{v}}{dt}$$

Accordingly Sciama has to identify the inertial mass of a particle with the negative of its gravitational potential energy, i.e. $G\Phi \simeq -c^2$ or $G\sigma\tau^2 \simeq 1$.

On the other hand if Sciama had employed σ in his definition of Φ, and ρ in his definition of $\underline{A}$, then Φ would not have come in to his inductive term and he would have obtained $G\rho\tau^2 \simeq 1$.

In the analytic theory of Mach's Principle provided by general relativity, as presented in Chapter III, the gravitational potential energy of a particle of unit mass, at rest in the neighbourhood of a fundamental observer, was viewed by him to be $\Phi_0 = c^2/2$ (equation (3.6)). In the present Chapter we have identified this expression with the integral (4.19), defined in terms of the gravitational mass distribution in the steady state model. For that model it follows from (4.9) that

$$G\sigma T^2 = -\frac{3}{4\pi} \qquad \text{.....................(4.24)}$$

Thus σ is negative (due to the contribution of negative stress), which can be interpreted via equation (4.11) as being consistent with a cosmic gravitational repulsion.

For the case of the relative motion, with velocity $\underline{v}$, of the smoothed out universe our analytic theory of Chapter III led to $\underline{A} = -c\,\underline{v}$ (equation (4.21)) whose time derivative accounted for the inertial term in the equation of motion. This value of $\underline{A}$ was in turn identified, in a manner analogous to the quasi-Galilean case, with the integral (4.22), leading to (4.23). This integral was defined in terms of the <u>inertial</u> mass distribution in the steady state model, for which (4.8) yields

$$G\rho T^2 = \frac{3}{8\pi} . \qquad \text{.....................(4.25)}$$

From (4.24) and (4.25) it follows that general relativity does indeed provide for the origin of inertia in accordance with Mach's Principle, and implies too the same consistency relation between the gravitational constant, Hubble's constant and the cosmic mass density (apart from sign) as does Sciama's theory. Like Sciama we can infer (in our case from (4.22)) that 99% of the contribution to local inertia is due to matter at a distance exceeding 10^8 light years.

(b) As was mentioned in Chapter I, Section (iii) Sciama takes the view that in general relativity one gravitating mass in an otherwise empty universe produces the same inertial effects as in his theory, and since there is no universe in this case to give rise to an inductive field at a test particle 'it is difficult to see why the Principle of Equivalence should be true'.

This argument is of course based on a quasi-Galilean solution of the field equations. In Chapter II, Section (iii) we have given detailed reasons why we reject as inadmissible in general relativity a solution based on empty space at infinity. The logical function of general relativity, according to the field equations, is to explain the existence of quasi-inertial frames, which extend over regions as large as the solar system, in terms of the matter in a full universe. It is this matter too that serves to characterise and physically define the coordinates of the world reference frame. The metric of special relativity, approximately valid in these quasi-inertial frames, forms the basis to which the metric of an arbitrary reference frame can be reduced locally. It cannot be taken for granted that this would be physically valid in a hypothetically empty universe (or empty save for a single massive body). Thus the only cosmic boundary conditions logically valid for general relativity are those of the actual universe.

In these views the present author is supported by Mach, for he says (op. cit. (1) p.229): 'If we now suddenly neglect A, B, C......., the actual massive bodies of the universe, and attempt to speak of the deportment of the body K in absolute space we implicate ourselves in a two-fold error. In the first place we cannot know how K would act in the absence of A,B,C....; and in the second place every means would be wanting of forming a judgment of the behaviour of K and of putting to the test what we had predicated - which latter therefore would be bereft of all scientific significance'. Again on p.235 he says: 'When we reflect that the time-factor that enters into the acceleration is nothing more than a quantity that is the measure of the distances (or angles of rotation) of the bodies of the universe we see that even in the simplest case, in which apparently we deal with the mutual action of only two masses the neglecting of the rest of the world is impossible. Nature does not begin with elements as we are obliged to begin with them.'

For reasons stated in Section (i) of the present Chapter we have regarded the steady state model as the simplest, and logically most appropriate to the equations of general relativity. In terms of that model general relativity can indeed account for inertia by means of the analytic theory of Mach's Principle presented in this thesis.

We see by equations (4.24) and (4.25) that however near the concept of an empty universe we approach as a limit (σ and $\rho \rightarrow 0$), then this is compensated in the steady state by the fact that the effective radius of space $\bar{R} = cT$ must tend correspondingly to infinity, so that inertia is always accounted for.

(c) Sciama's claim that his theory implies that gravitation is attractive (for concentrated matter) is of doubtful significance. It is based on his equation of motion for a test particle in the field of a concentrated mass M at rest in the smoothed out universe. This equation is given by (1.7), and in the form in which Sciama directly derives it from his postulated basic equation (1.5) it may be written

$$\underline{E} = \left\{ -\frac{M}{r^2} - \left(\frac{\Phi + \phi}{c^2} \right) \frac{dv}{dt} \right\} \hat{\underline{r}} = 0 \quad ,$$

where $\hat{\underline{r}}$ is a unit vector drawn from M towards the test particle, and v is the velocity of the universe relative to the particle in the direction of $\hat{\underline{r}}$. The equation expresses that the total 'gravoelectric' field is zero at the particle.

Sciama concludes that for $\underline{E}$ to vanish then $\frac{dv}{dt}$ must be positive, so that the particle must accelerate <u>towards</u> the mass at rest in the smoothed out universe. This assertion is however based on the assumption that his Φ as defined by (1.6) should be negative, which means that σ the gravitational mass density of the smoothed out universe must be positive. There are no grounds in his theory for this assumption; nor would there be grounds in his theory for taking ρ , the smoothed out cosmic inertial mass density, to be positive, supposing that he did in fact distinguish between ρ and σ , and it was granted that the inertia of concentrated matter was positive.

In general relativity the sign of the coefficient of $T_{\mu\nu}$ in the field equations (1.4) is chosen negative, so that the gravitation of local concentrated matter should be attractive. However this leads to cosmic gravitational mass density σ being interpreted as negative in any model, such as the steady state model, in which the expansion is accelerating (c.f.equation (4.11). Thus in the steady state model, if local concentrated matter is regarded as attractive, then $\sigma < 0$ and $\rho > 0$ by (4.8) and (4.9). However if the coefficient of $T_{\mu\nu}$ in the field equations is chosen positive, then concentrated matter is repulsive, and the signs of σ and ρ are reversed in the case of the steady state.

The physical interpretation given in this Chapter of the analytic theory of Mach's Principle presented in Chapter III is valid in either case, with appropriate choice of signs throughout.

Thus the sign of the field would appear to be of no intrinsic importance in any theory that allows for its dependence on the state of the gravitating mass.

Conclusion

Our broad conclusion is that general relativity is entirely consistent in essence with the theory of Mach's Principle put forward by D. W. Sciama and, on the basis indicated in this thesis, general relativity would appear to fully incorporate Mach's Principle in the case of a steady state cosmology.

R E F E R E N C E S

(1) E. Mach. The Science of Mechanics; 2nd Ed. Chicago 1893

(2) R. Eötvös. Math. and Naturw. Ber. aus Ungarn 8 65 (1890)

(3) A. Einstein. Ann. d. Phys. 49 769 (1916)

(4) D.W. Sciama. Mon. Not. R.A.S. 113 34 (1953)

(5) A. Einstein. The Meaning of Relativity; 5th Ed. London: Methuen, 1951 p. 97.

(6) A. Einstein. Op. cit. p. 83

(7) A. Einstein. S.B. d. Preuss. Akad. d. Wiss 142 (1917)

(8) W. de Sitter. Proc. Akad. Wetensch. Amsterdam 19 1217 (1917)

(9) H. Bondi. Cosmology; Cambridge: C.U.P. 1952 p. 98

(10) H. Bondi and T. Gold. Mon. Not. R. A.S. 108 252 (1948)

(11) F. Hoyle. Mon. Not. R.A.S. 108 372 (1948)

(12) W.H. McCrea. Proc. Roy. Soc. A 206 562 (1951)

(13) H. P. Robertson. Proc. Nat. Acad. Sci. 15 822 (1929)

(14) J. Lemaitre. J. Math. and Phys. (M.I.T.) 4 188 (1925)
H.P. Robertson. Phil. Mag. 5 835 (1928)

(15) J.L. Synge. Proc. Edin. Math. Soc. 2nd Series 5 93 (1937)

(16) e.g. H. Bondi. Op. cit (9) p. 103

* (17) W. Davidson. Mon. Not. R.A.S. 117 212 (1957)

----

* Publication by the author, referred to in the Abstract; a copy is submitted with this thesis as supporting material.

PART III

THE MECHANISM OF STEADY STATE COSMOLOGY ACCORDING TO GENERAL RELATIVITY

INTRODUCTION

As we have had occasion to state elsewhere in this thesis, it has been shown by W.H.McCrea (1) that the continuous creation of matter required by steady state cosmology, first proposed by H. Bondi and T. Gold (2), and by F. Hoyle (3), can be based on the existing equations of general relativity, without introducing, ad hoc, extra terms into these equations. In the theory of continuous creation advanced by Hoyle an extra term is added to the field equations of general relativity, distinct from the energy tensor, so that energy is no longer conserved and matter is in fact created out of nothing at a rate determined by the extra term.

McCrea showed that by absorbing the extra term proposed by Hoyle into the energy tensor it could be interpreted as a cosmological stress term, which in Hoyle's model was taken to be zero. The metric of space-time in the steady state model is established, on the basis of Bondi and Gold's 'perfect cosmological principle', as that of the de Sitter universe. Accordingly, the application by McCrea of the field equations of general relativity to this metric showed that the stress must be uniform, independent of cosmic epoch, and negative, being the negative of the constant energy density in the model. The matter created in the model, at the constant rate necessary to maintain the steady state, was then interpreted by McCrea as the mass equivalent of the work done by the negative stress in the expansion of the model.

McCrea did not propose any physical basis for this negative stress, except to point out that a zero-point of stress, other than what was usually assumed arbitrarily on the basis of classical mechanics, was consistent with the field equations of general relativity, and with the virtual zero-point properties accorded to space by the quantum theory of fields. Such a stress could be important cosmologically and yet be beyond detection in the laboratory because of its minute order and uniformity on that scale.

In the event of the steady state being established as the true cosmology then the principle of conservation of energy, although hitherto a dependable feature of the physical laws, would be seriously jeopardised by the theories of Bondi and Gold, and of Hoyle. It would be satisfactorily redeemed by McCrea's interpretation only if a physical interpretation of the zero-point stress can be found. This physical interpretation must be, as far as possible, consistent with the accepted physical phenomena. It would also be desirable that it should be suggested by existing physical theory, or at least involving only natural generalisation or extension of such theory.

It is our purpose in the following pages to present such a physical basis for zero-point stress, and associated zero point energy. The work is entirely original, although, as will be pointed out, there is a qualitative similarity of part of the theory in Chapter III to certain ideas advanced by F. Pirani (4), these ideas having been arrived at independently by the author before he learned of Pirani's paper.

CHAPTER I: MASS CONCENTRATION IN THE STEADY STATE UNIVERSE ACCORDING TO GENERAL RELATIVITY - EXTERIOR SOLUTION

(i) The steady state metric interpreted according to general relativity.

The metric of the general expanding cosmological models, which are spatially isotropic at every point, has the form

$$ds^2 = dt^2 - R^2(t)\,\frac{(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2)}{(1 + kr^2/4)^2} \qquad \text{....(1.1)}$$

where we have taken c the velocity of light as unity. This metric has been examined in detail in Part I of this thesis. It is to be considered as applicable to a universe whose local irregularities due to mass concentrations have been ideally smoothed out.

The field equations of general relativity applied to this metric yield for the uniform, and in general time dependent, energy density ρ and pressure p (Part I, Chapter IV),

$$K\rho = \frac{3\dot{R}^2}{R^2} + \frac{3k}{R^2} \qquad \text{....(1.2)}$$

$$-Kp = \frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{k}{R^2} \qquad \text{....(1.3)}$$

where $\dot{R} \equiv d\{R(t)\}/dt$, and $K = 8\pi$ if we adopt $c = 1$ and $G = 1$, G being the Newtonian gravitational constant. The cosmological constant has no place in the following theory of the steady state and has been put equal to zero.

The steady state model has the metric corresponding to $R(t) = e^{t/R_s}$ (R_s constant), and $k = 0$ $(\underline{2})$, viz.

$$ds^2 = dt^2 - e^{2t/R_s}(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2) \qquad \text{....(1.4)}$$

For this metric equations (1.2), (1.3) yield

$$\rho_s = -p_s = \frac{3}{KR_s^2} \qquad \text{....(1.5)}$$

It should be noted that ρ_s and p_s are by definition the density and pressure as measured locally by an observer moving along with the fluid of the model.

The metric (1.4) is the line element of the de Sitter universe in the form obtained by G. Lemaitre, and independently by H.P. Robertson, (5). This was related to one of the earlier forms, viz.

$$ds^2 = \left(1 - \frac{\ell^2}{R_s^2}\right)dT^2 - \frac{d\ell^2}{1 - \frac{\ell^2}{R_s^2}} - \ell^2 d\theta^2 - \ell^2 \sin^2\theta \, d\phi^2 \qquad \text{....(1.6)}$$

according to the transformation

$$\tau = \frac{\ell e^{-T/R_s}}{\sqrt{1 - \ell^2/R_s^2}} \;, \quad t = T + \tfrac{1}{2} R_s \log\left(1 - \ell^2/R_s^2\right) \qquad \text{....(1.7)}$$

The maintenance of constant energy density and pressure in the steady state will be examined in accordance with general relativity in Chapter III.

(ii) <u>Mass concentration superimposed on the steady state - the metric of the exterior solution and its properties.</u>

Consider the metric

$$ds^2 = \left(1 - \frac{2m}{\ell} - \frac{\ell^2}{R_s^2}\right)d\tau^2 - \frac{d\ell^2}{1 - \frac{2m}{\ell} - \frac{\ell^2}{R_s^2}} - \ell^2 d\theta^2 - \ell^2 \sin^2\theta \, d\phi^2 \qquad \text{....(1.8)}$$

where m is a constant. For ℓ sufficiently small (ℓ^2/R_s^2 negligible compared with $2m/\ell$) this reduces to the Schwarzschild metric for the exterior field due to a mass m located at the space origin $\ell = 0$. For ℓ sufficiently large the metric goes over to that of the steady state given by (1.6). We therefore infer that (1.8) represents the stationary space-time due to a spherically symmetric mass concentration m

superimposed on the smoothed out steady state universe. We shall examine this metric according to general relativity.

A stationary metric exhibiting spherical symmetry about $\ell = 0$ has the general form

$$ds^2 = e^{\nu} d\tau^2 - e^{\lambda} d\ell^2 - \ell^2 d\theta^2 - \ell^2 \sin^2\theta \, d\phi^2 \qquad \text{....(1.9)}$$

where λ, ν are functions of ℓ only. For the non-zero components of the stress-energy-momentum tensor T_μ^ν the field equations give for this metric (6),

$$\left.\begin{aligned} KT_1^1 &= -e^{-\lambda}\left(\frac{\nu'}{\ell} + \frac{1}{\ell^2}\right) + \frac{1}{\ell^2} \\ KT_2^2 = KT_3^3 &= -e^{-\lambda}\left(\frac{\nu''}{2} - \frac{\lambda'\nu'}{4} + \frac{\nu'^2}{4} + \frac{\nu'-\lambda'}{2\ell}\right) \\ KT_4^4 &= e^{-\lambda}\left(\frac{\lambda'}{\ell} - \frac{1}{\ell^2}\right) + \frac{1}{\ell^2} \end{aligned}\right\} \qquad \text{....(1.10)}$$

where dashes denote derivatives. Here the indices 1,2,3,4 are associated with the coordinates ℓ, θ, ϕ, τ respectively. Substituting from (1.8) we easily obtain

$$T_1^1 = T_2^2 = T_3^3 = T_4^4 = \frac{3}{KR_s^2} \qquad \text{....(1.11)}$$

We now assume that the stress in the fluid, as measured locally by co-moving observers, is still everywhere isotropic despite the presence of the mass concentration i.e. the medium constitutes a perfect fluid characterised by a proper energy density ρ and pressure p. The T_μ^ν are therefore related to the 4-vector of velocity of the fluid by the following relation, holding at an arbitrary point outside the mass concentration:

$$T_\mu^\nu = g_{\mu\alpha}(\rho + p)\frac{dx^\alpha}{ds}\frac{dx^\nu}{ds} - g_\mu^\nu p \qquad \text{....(1.12)}$$

Accordingly, for the non-vanishing T_μ^ν corresponding to metric (1.8) we derive

$$\left.\begin{aligned} T_1^1 &= g_{11}(\rho+p)\left(\frac{dl}{ds}\right)^2 - p \\ T_2^2 &= g_{22}(\rho+p)\left(\frac{d\theta}{ds}\right)^2 - p \\ T_3^3 &= g_{33}(\rho+p)\left(\frac{d\phi}{ds}\right)^2 - p \\ T_4^4 &= g_{44}(\rho+p)\left(\frac{d\tau}{ds}\right)^2 - p \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots(1.13)$$

By spherical symmetry we can put $d\theta/ds = d\phi/ds = 0$ whence $T_2^2 = T_3^3 = -p$, so that by (1.11)

$$p = -\frac{3}{KR_s^2} \qquad \ldots.(1.14)$$

Furthermore we can substitute $T_4^4 = -p$ in the last equation of (1.13) so that we find

$$g_{44}(\rho+p)\left(\frac{d\tau}{ds}\right)^2 = 0,$$

and since $g_{44} \neq 0$, $d\tau/ds \neq 0$, we must have $\rho + p = 0$, whence

$$\rho = \frac{3}{KR_s^2} \qquad \ldots\ldots\ldots\ldots\ldots(1.15)$$

It transpires, therefore, that at all points for which the metric (1.8) is valid, i.e. outside the mass m, the density and pressure are those of the smoothed out steady state. We interpret this as requiring that the natural equilibrium state of internebular gas and radiation in the steady state universe should be characterised by a total density ρ_s and pressure p_s as given by (1.5). Although outside the mass m we have ignored the other mass concentrations in the actual universe (galaxies) this is probably justified because of their great distance apart, so that their aggregate effect is that of the smoothed out universe we are considering. It is also justified by our interpretation of the consequence of this assumption - that the internebular density is likely to be comparable with the smoothed out density. In what follows the mass m will have several interpretations. It may be identified with a galaxy, or a cluster of galaxies where the average density is only a few times the average cosmological density, or a mass of internebular fluid whose density may be slightly higher, or slightly lower ($m < 0$), than the equilibrium density. The region of space where the metric (1.8) is not valid will define the interior of the mass concentration (or rarefaction) and there ρ and p must differ from ρ_s and p_s, since otherwise m would vanish.

Among the components of T_μ^ν which vanish identically for the metric (1.9) is T_4^1. For a general orthogonal metric describing the space-time occupied by a perfect fluid we have

$$T_4^1 = g_{44}(\rho + p) \frac{dl}{ds} \frac{d\tau}{ds} \qquad \text{....(1.16)}$$

Since this component vanishes identically one would normally conclude that $dl/ds = 0$, and this would mean that the fluid was at rest relative to the reference frame, that is the coordinate system would be co-moving. But for our case when $\rho + p = 0$ this is not the conclusion, but rather the vanishing of $\rho + p$ indicates that the <u>net</u> algebraic rate of flow of positive and negative energy (which will include the rate of work by the negative pressure p) is zero. The 4-velocity of the medium of net positive density but negative pressure will not however vanish.

It is a well known result in general relativity that if the pressure in a fluid (isotropic) is everywhere constant then the world lines of the fluid elements are geodesics in space-time. This theorem normally follows from the fact that the covariant divergence of $T^{\mu\nu}$ vanishes. But in our case when $\rho+p$ actually vanishes the proof breaks down. However we shall assume for two reasons that the result nevertheless holds in this case too; firstly by reason of continuity since the proof holds however near $\rho+p$ is to zero without actually vanishing, and secondly because the constituent of the medium that is ordinary matter must, in a locally inertial frame, follow the law of inertia for a medium of uniform pressure. Because of the homogeneous character of the model all constituents of the fluid will have the same average motion.

Consider therefore a radial geodesic in the space-time of metric (1.8), corresponding to $\theta = \text{const.}$, $\phi = \text{const.}$. For the more general metric (1.9) the differential equations are easily shown to be

$$\left.\begin{aligned} &\frac{d^2 l}{ds^2} + \frac{1}{2}\lambda'\left(\frac{dl}{ds}\right)^2 + \frac{1}{2}e^{\nu-\lambda}\nu'\left(\frac{d\tau}{ds}\right)^2 = 0 \\ &\frac{d^2\tau}{ds^2} + \nu'\frac{dl}{ds}\frac{d\tau}{ds} = 0 \end{aligned}\right\} \qquad \ldots.(1.17)$$

Integration of the second equation yields

$$\frac{dT}{ds} = \alpha e^{-\nu} \qquad (\alpha \text{ constant}) \qquad \text{....(1.18)}$$

The first equation then becomes

$$\frac{d^2 l}{ds^2} + \tfrac{1}{2}\lambda' \left(\frac{dl}{ds}\right)^2 + \tfrac{1}{2}\alpha^2 \nu' e^{-\lambda-\nu} = 0 \qquad \text{....(1.19)}$$

This has a first integral provided by the metric (1.9) itself, which with (1.18) leads to

$$\left(\frac{dl}{ds}\right)^2 = e^{-\lambda}\left(\alpha^2 e^{-\nu} - 1\right) \qquad \text{....(1.20)}$$

For the metric of interest (1.8) we therefore get

$$\left(\frac{dl}{ds}\right)^2 = \alpha^2 - 1 + \frac{2m}{l} + \frac{l^2}{R_s^2} \qquad \text{....(1.21)}$$

while (1.19) becomes

$$\frac{d^2 l}{ds^2} = -\frac{m}{l^2} + \frac{l}{R_s^2} \qquad \text{...........(1.22)}$$

We see that by (1.21), whatever the parameter α of the motion of an individual particle, the velocity for sufficiently large l is approximately $dl/ds = l/R_s$. In fact the difference between the actual velocity and the value l/R_s clearly tends to zero for all α. It follows that there is a natural motion for the fluid substratum which can be characterised in the undisturbed state ($m=0$) by taking $\alpha = 1$. There are now three cases to consider in the disturbed state.

(a) Suppose first that m is sufficiently large so as to provide a solution for l_0 in the equation

$$-\frac{m}{l_0^2} + \frac{l_0}{R_s^2} = 0,$$

$$\text{or } l_0^3 = m R_s^2, \qquad \text{.................(1.23)}$$

which is greater than ℓ_1 , the radial coordinate of the spherical boundary to the mass m . In this case the surface $\ell = \ell_0$ forms a neutral surface where the attraction of the mass m is balanced by the cosmical repulsion, which is characterised in the steady state by the second term in (1.22). A material particle shot from the mass m with sufficient velocity to reach the neutral surface would escape for ever into outer space. Its actual behaviour would depend on the parameter α of its motion. It is easy to see that if the velocity just vanishes at $\ell = \ell_0$ then

$$\alpha = \left[1 - 3\left(\frac{m}{R_s}\right)^{2/3}\right]^{\frac{1}{2}} \qquad \text{..............(1.24)}$$

This is very near unity even if m represents a cluster of galaxies, since R_s is equal to the radius of the observable universe (Part II, Chapter IV).

It follows that the substratum itself will be at rest at the neutral surface. In fact if the fluid had an inward velocity at the neutral surface then it would have an even higher velocity inwards at more remote points to overcome cosmical repulsion. This is contrary to our previous conclusion regarding the natural motion of the fluid as one of expansion. Nor could there be an outward velocity of the medium at the neutral surface, for this would imply that the medium, including its material constituent, was steadily leaving the mass m with a velocity greater than the escape velocity. This would be such an unnatural phenomenon that it must immediately be rejected.

Accordingly, we have to regard the fluid medium as at rest on the neutral surface $l = l_0$, accelerating towards the mass m for $l < l_0$, and accelerating outwards for $l > l_0$. It should be noticed however that there is no net accretion of mass by the mass concentration, in agreement with the assumed steady state and constancy of m, since the condition $\rho_s + p_s = 0$, which holds in the exterior field, means that there is no net flux of energy through any spherical surface, concentric with the space origin and of fixed radius, which lies entirely in the exterior fluid outside the mass concentration. We can assume the whole mass inside this surface to be contracting or expanding in gravitational pulsations, if we please, without upsetting the exterior field, provided spherical symmetry is maintained. This is the analogue of Birkhoff's theorem for the steady state universe (see Section (iv)). If the mass was luminous, and actually radiating energy, this steady state condition would of course be violated, and in the exterior fluid there would be a net flow of energy radially outwards.

The assumption of a steady state, however, serves to illustrate the effect of mass concentrations on the surrounding substratum (inter-galactic space for instance) in which the density and pressure maintain the values ρ_s and p_s respectively. We quote here results which will be proved in Chapter III (Sect.(vi) and which show that the rate of creation of matter in the steady state universe, treated in accordance with general relativity, is not a statistically uniform process throughout space but in fact strongly dependent upon the local gravitational field.

Inside the neutral surface the rate of creation of matter is so reduced in relation to that obtaining in the undisturbed steady state, due to the gravitational influence of the mass concentration, that in fact annihilation prevails for $\ell < \left\{2/(1+\sqrt{5})\right\} \ell_0 \ (\simeq .62\,\ell_0)$. Outside the neutral surface the rate of creation increases steadily to reach the undisturbed steady state value at great distance.

(b) If m is still positive but not sufficiently large to provide a solution for $\ell_0 > \ell_1$, then the neutral surface does not exist. In this case the substratum everywhere expands but is retarded in the neighbourhood of the mass m, when compared with the undisturbed steady state. The rate of creation of matter near m will be less than the average rate in space.

(c) If m is negative, so that there is a 'hole' in the substratum, then in the neighbourhood of m cosmical repulsion is assisted by the local gravitational repulsion of m itself. In this neighbourhood the rate of creation of matter will exceed the average rate in space.

These conclusions are reminiscent of certain very general arguments that have been put forward in the past by R.O. Kapp (7), as the present author realised when he had formulated this analysis. Kapp advanced the notion of annihilation of matter proceeding simultaneously with its creation, and suggested that whether annihilation or creation predominated in a region must be related to the degree of local concentration of matter. The analysis therefore serves to substantiate Kapp's ideas to this extent.

The cases of (b) and (c) have particular relevance to the theory of the maintenance of the steady state, in a manner consistent with the conservation of energy, which will be put forward in Chapter III.

<u>(iii) Alternative form of the metric of the exterior solution</u>

It is of interest to consider in relation to the metric (1.8) a metric which is a special case of a more general line element obtained by G.C.McVittie (8). McVittie obtained the metric of space-time corresponding to a massive particle superimposed on the general smoothed out cosmological model whose line element is given by (1.1), and in which the reference frame was designed to be co-moving with the fluid. This metric took the general form, in our notation,

$$ds^2 = \left\{\frac{1 - \frac{\phi(t)}{2r}\left(1+kr^2/4\right)^{\frac{1}{2}}}{1 + \frac{\phi(t)}{2r}\left(1+kr^2/4\right)^{\frac{1}{2}}}\right\}^2 dt^2 - \left\{\frac{\left\{1 + \frac{\phi(t)}{2r}\left(1+kr^2/4\right)^{\frac{1}{2}}\right\}^4}{\left(1+kr^2/4\right)^2}\right\} R^2(t)\left(dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right) \qquad \text{..........(1.25)}$$

where $$\frac{\dot{\phi}}{\phi} = -\frac{\dot{R}}{R} \qquad \text{..............................(1.26)}$$

It follows that, for the steady state background of metric (1.4), equation (1.26) yields

$$\phi = m\, e^{-t/R_s} \qquad \text{..................(1.27)}$$

where m is a constant which we identify with the constant m in the metric (1.8). Accordingly, since $k=0$ for the steady state, (1.25) becomes

$$ds^2 = \left(\frac{1 - \frac{me^{-t/R_s}}{2r}}{1 + \frac{me^{-t/R_s}}{2r}}\right)^2 dt^2 - \left(1 + \frac{me^{-t/R_s}}{2r}\right)^4 e^{2t/R_s}\left(dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right) \quad \ldots\ldots(1.28)$$

When r is large this metric goes over to that of the steady state given in the form (1.4). Suppose now r is small and put $r' = re^{t/R_s}$ so that r' is also small for a finite value of t, then $dr' \simeq e^{t/R_s}\,dr$ if we neglect r'/R_s. In these circumstances (1.28) becomes approximately

$$ds^2 = \left(\frac{1 - \frac{m}{2r'}}{1 + \frac{m}{2r'}}\right)^2 dt^2 - \left(1 + \frac{m}{2r'}\right)^4 \left\{(dr')^2 + (r')^2 d\theta^2 + (r')^2\sin^2\theta\, d\phi^2\right\} \quad \ldots\ldots(1.29)$$

which is the well known isotropic form of the Schwarzschild metric for a mass m in otherwise empty space. That is, by our approximation in (1.29), we have neglected the effect of the cosmological background to the same extent as is done in deriving the Schwarzschild metric from the field equations.

To verify that the metric (1.28) does in fact represent the same model as does (1.8) we shall now find the proper density and pressure corresponding to this metric. For the general orthogonal metric of the form

$$ds^2 = e^{\nu}\,dt^2 - e^{\mu}\left(dr^2 + r^2 d\theta^2 + r^2\sin^2\theta\, d\phi^2\right) \quad \ldots\ldots(1.30)$$

where ν, μ are now functions of both r and t, the non vanishing components of the T_μ^ν are found to be (9),

$$\left.\begin{aligned}
KT_1^1 &= -e^{-\mu}\left(\frac{\mu'^2}{4}+\frac{\mu'\nu'}{2}+\frac{\mu'+\nu'}{r}\right)+e^{-\nu}\left(\ddot{\mu}+\frac{3\dot{\mu}^2}{4}-\frac{\dot{\mu}\dot{\nu}}{2}\right)\\
KT_2^2 = KT_3^3 &= -e^{-\mu}\left(\frac{\mu''}{2}+\frac{\nu''}{2}+\frac{\nu'^2}{4}+\frac{\mu'+\nu'}{2r}\right)+e^{-\nu}\left(\ddot{\mu}+\frac{3\dot{\mu}^2}{4}-\frac{\dot{\mu}\dot{\nu}}{2}\right)\\
KT_4^4 &= -e^{-\mu}\left(\mu''+\frac{\mu'^2}{4}+\frac{2\mu'}{r}\right)+\frac{3}{4}e^{-\nu}\dot{\mu}^2\\
KT_4^1 &= e^{-\mu}\left(\dot{\mu}'-\frac{\dot{\mu}\nu'}{2}\right)\\
KT_1^4 &= -e^{-\nu}\left(\dot{\mu}'-\frac{\dot{\mu}\nu'}{2}\right)
\end{aligned}\right\}\quad\ldots.(1.31)$$

where a dash indicates a partial derivative with respect to r, and a dot with respect to t. After straightforward calculations, which will not be given here, it is found that for the metric (1.28)

$$\left.\begin{aligned}
T_1^1 = T_2^2 = T_3^3 = T_4^4 &= \frac{3}{KR_S^2}\\
T_4^1 = T_1^4 &= 0
\end{aligned}\right\}\quad\ldots.(1.32)$$

Having regard to the expression for the T_μ^ν given by (1.12), assuming a perfect fluid as before, and following exactly the same arguments as for the metric (1.8), we find for the proper density and pressure

$$\rho = -p = \frac{3}{KR_S^2}\qquad\ldots.(1.33)$$

Thus we have confirmed that the metric (1.28) represents the same physical model as that described by metric (1.8). Accordingly, there must exist a transformation of coordinates connecting the two metrics. By comparing the coefficients of $d\theta^2$ in the two metrics it is easy to see that one equation of the transformation is

$$\ell = r\left(1 + \frac{me^{-t/R_s}}{2r}\right)^2 e^{t/R_s} \qquad \text{....(1.34)}$$

It may also be deduced by comparison of the metrics that the differential equations for T are

$$\frac{\partial T}{\partial r} = \frac{\frac{r}{R_s}\left(1 + \frac{me^{-t/R_s}}{2r}\right)^4 e^{2t/R_s}}{\left(\frac{1 - \frac{me^{-t/R_s}}{2r}}{1 + \frac{me^{-t/R_s}}{2r}}\right)^2 - \frac{r^2}{R_s^2}\left(1 + \frac{me^{-t/R_s}}{2r}\right)^4 e^{2t/R_s}} \qquad \text{....(1.35)}$$

$$\frac{\partial T}{\partial t} = \frac{\left(\frac{1 - \frac{me^{-t/R_s}}{2r}}{1 + \frac{me^{-t/R_s}}{2r}}\right)^2}{\left(\frac{1 - \frac{me^{-t/R_s}}{2r}}{1 + \frac{me^{-t/R_s}}{2r}}\right)^2 - \frac{r^2}{R_s^2}\left(1 + \frac{me^{-t/R_s}}{2r}\right)^4 e^{2t/R_s}} \qquad \text{.....(1.36)}$$

The author has not however succeeded in finding integrals of these relations.

As expressed in equations (1.32) we find $T_4^1 = 0$ for the metric (1.28), as we found also for the metric (1.8).

McVittie postulated the condition $T^1_4 = 0$ in all cases included by the general metric (1.25), in order to make the coordinate system co-moving. This is achieved in all cases except the special case of our metric (1.28) for which $\rho + p = 0$, so that the condition

$$T^1_4 = g_{44}(\rho + p)\frac{dr}{ds}\frac{dt}{ds} = 0 \qquad \text{..............(1.37)}$$

does not necessarily require $dr/ds = 0$ for the fluid elements. But if as before we require that the world lines of the fluid elements should be geodesics, corresponding to uniform pressure (which does not hold generally for McVittie's models where the coordinates are co-moving, since the pressure is non-uniform), then the world lines can be derived. We shall examine their general character.

It is clear that for sufficiently large r the equations of a radial geodesic of the space-time (1.28) will be those derived for the limiting form of the metric at great distance i.e. the metric (1.4). Consequently for large r the differential equations of the radial geodesic can be written approximately

$$\frac{d^2r}{ds^2} + \frac{2}{R_s}\frac{dr}{ds}\frac{dt}{ds} = 0 \qquad \text{......................(1.38)}$$

$$\frac{d^2t}{ds^2} + \frac{e^{2t/R_s}}{R_s}\left(\frac{dr}{ds}\right)^2 = 0 \qquad \text{.....................(1.39)}$$

Equation (1.38) has the first integral

$$\frac{dr}{ds} = \beta e^{-2t/R_s} \qquad (\beta \text{ constant}) \quad ....(1.40)$$

Accordingly (1.39) has the first integral obtainable from (1.40) and the line element itself (at large r),

$$\left(\frac{dt}{ds}\right)^2 = 1 + \beta^2 e^{-2t/R_s} \quad ....(1.41)$$

It follows, therefore, that ultimately any particle moving on a radial geodesic has the world line $r =$ const., $dt/ds = 1$. Once again this motion has to be interpreted, consequently, as the natural state of motion of the fluid, which it will therefore possess at sufficient distance from the mass m.

The reference points $r =$ const. steadily recede from the mass m as is seen by examining its motion relative to the static reference system associated with metric (1.8). Putting $t = \text{const.}$ in the relation (1.34) we find that ℓ is a monotonic function of r provided that $|m| < 2re^{t/R_s}$, which can be assumed to hold outside the mass (c.f. (1.28)). Further, on differentiating (1.34), we obtain

$$\frac{d\ell}{dt} = e^{t/R_s}\left(1 - \frac{m^2 e^{-2t/R_s}}{4r^2}\right)\left(\frac{dr}{dt} + \frac{r}{R_s}\right) \quad ....(1.42)$$

Hence if $dr/dt = 0$ we find

$$\frac{d\ell}{dt} = \frac{r}{R_s} e^{t/R_s}\left(1 - \frac{m^2 e^{-2t/R_s}}{4r^2}\right) \quad ....(1.43)$$

This expression is always positive in the range in which ℓ is a monotonic increasing function of r, and tends to the value given by the relation applicable to the undisturbed steady state, viz.

$$\frac{d\ell}{dt} = \frac{\ell}{R_s} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1.44)$$

since $\ell \to r e^{t/R_s}$ by (1.34). It follows that the reference points $r = \text{const.}$ move through the static frame (1.8), with a velocity that is ultimately proportional to distance. As we have seen this is also the limiting motion of the fluid.

For smaller r the fluid can be expected to lag behind the reference frame in cases (a) and (b) and to be advance in case (c). For instance in case (a) we have seen that at the neutral surface the fluid is at rest in the static reference frame. Accordingly, putting $d\ell/dt = 0$ in (1.42) we find that the fluid has the coordinate velocity $dr/dt = -r/R_s$ in the non static frame, and so is lagging behind the frame.

<u>(iv) The analogue of Birkhoff's theorem</u>

The existence of an analogue of Birkhoff's theorem (<u>10</u>) when the cosmological background is the steady state, instead of empty space, was alluded to in the previous Section. Birkhoff shows that the exterior metric associated with a mass of spherical symmetry embedded in empty space can always be reduced by transformation to the static Schwarzschild solution. Thus spherically symmetrical pulsations could take place in the sphere without leading to loss of energy by gravitational waves.

Consider now a spherically symmetric mass embedded in the smoothed out steady state universe. For a non static system of coordinates with origin at the centre of the sphere we may always convert the general form of metric applicable to

$$ds^2 = e^{\nu} d\tau^2 - e^{\lambda} dl^2 - l^2 d\theta^2 - l^2 \sin^2\theta \, d\phi^2 \quad \ldots.(1.45)$$

where λ and ν are functions of l and τ. An isotropic form of this metric such as in the special case of (1.28) can always be converted by transformation into the form given above. For this metric the T_μ^ν are given by (9),

$$\left.\begin{aligned} KT_1^1 &= -e^{-\lambda}\left(\frac{\nu'}{l} + \frac{1}{l^2}\right) + \frac{1}{l^2} \\ KT_2^2 = KT_3^3 &= -e^{-\lambda}\left(\frac{\nu''}{2} - \frac{\lambda'\nu'}{4} + \frac{\nu'^2}{4} + \frac{\nu'-\lambda'}{2l}\right) + e^{-\nu}\left(\frac{\ddot{\lambda}}{2} + \frac{\dot{\lambda}^2}{4} - \frac{\dot{\lambda}\dot{\nu}}{4}\right) \\ KT_4^4 &= e^{-\lambda}\left(\frac{\lambda'}{l} - \frac{1}{l^2}\right) + \frac{1}{l^2} \\ KT_4^1 &= -e^{-\lambda}\frac{\dot{\lambda}}{l}, \quad KT_1^4 = e^{-\nu}\frac{\dot{\lambda}}{l} \end{aligned}\right\} \ldots.(1.46)$$

All other components vanishing.

If now we suppose that outside the mass the density and pressure are everywhere permanently equal to those of the steady state i.e. $\rho = \rho_s$, $p = p_s$ then, as in (1.16) $T_4^1 = g_{44}(\rho+p)\frac{dl}{ds}\frac{d\tau}{ds} = 0$ Hence $\dot{\lambda} = 0$ by (1.46). Accordingly, the equations (1.46) become identical with those of the static system of metric (1.9) viz. the equations (1.10). Substitution for the T_μ^ν in terms of ρ_s and p_s must therefore lead to the metric (1.8), where m must be a constant independent of τ to make $\dot{\lambda}$ vanish.

It follows that, provided spherical symmetry is maintained and as long as the conditions $\rho = \rho_s$, $p = p_s$ outside the mass are not violated, the metric (1.8) will always apply in the exterior so that pulsations could then take place in the sphere without loss of energy due to gravitational waves. This is the analogue of Birkhoff's theorem in the steady state universe.

In this connection we mention the claim by A. Einstein and G. Straus (11) that expanding space-time has no effect on the gravitational field of a star. This claim was based on their solution for a mass particle in an expanding universe, in which they postulated that the particle was at the centre of a spherical region of space which was otherwise empty; outside this region space was filled with a uniform distribution of matter at uniform pressure. The metrics for the two regions were joined continuously. Einstein and Straus then showed that within the spherical region the metric could be transformed to the static Schwarzschild solution so that apparently the cosmological field had no effect on the field due to the mass particle.

However, that the field in the region immediately surrounding the particle can be reduced to the Schwarzschild solution follows immediately from Birkhoff's theorem (of which Einstein and Straus were then unaware), because of the artificial postulate that this region is empty. Nevertheless their view was apparently concurred in by McVittie (12), based on his own earlier solution already referred to, since he apparently overlooked that his general solution (1.25) could be reduced locally to the Schwarzschild solution by approximation only. It is evident that our derivation of (1.29) from the special case of metric (1.28) is only an approximation. In fact, as we have seen in case (a) of our metrics (1.28) or (1.8), the attractive field of the particle is opposed by cosmic repulsion so that its capturing power is limited to the region of space within a neutral surface of radius given by (1.23).

It should be mentioned that the claim by Einstein and Straus has also been criticised in a note by F. Pirani (13). Pirani, however, bases his argument only on the fact that Einstein and Straus have omitted to take account of the cosmological constant Λ, which would render the Schwarzschild solution dependent on cosmic influence. Nevertheless he confirms our own work here in pointing out that the Schwarzschild solution for positive Λ is equivalent to a mass particle in the steady state universe, and he deduces that the field is attractive in a finite region only, of a radius which would agree with our equation (1.23).

However, as we have pointed out in Sect. (iii), if m is very small, or negative, the field of attraction outside will not exist at all, as in cases (b), (c). The dependence of the field on cosmic influence is therefore abundantly clear.

CHAPTER II: MASS CONCENTRATION IN THE STEADY STATE UNIVERSE - INTERIOR SOLUTION

(i) Derivation of the metric

Our main concern in obtaining an interior solution, corresponding to the stationary exterior solution dealt with in Chapter I, will be with regions of space in which the mass density is different from, but of the same order as, ρ_s the density in the smoothed out steady state universe. It is only in this case that taking explicit account of the cosmological background can be expected to lead to significant information about the behaviour of the mass in such a region; and it is to regions of this order of density (on the average) that a theory of the mechanism of the steady state must be relevant.

Because it is mathematically more tractable, and also, as we shall show, adequate to illustrate the physical principles involved, we shall postulate that the proper density in the mass m is uniform and constant, and of course m is constant by hypothesis. The assumption of an interior static state in which ρ is constant will enable us to draw conclusions regarding quasi-static masses which are expanding or contracting under the prevailing gravitational field, under the same boundary conditions, so that by the analogue of Birkhoff's theorem there is conservation of energy within a sufficiently large fixed region whose boundary lies entirely in the smoothed out steady state substratum.

Let the mass m be a spherical mass of perfect fluid of density ρ and pressure p occupying the region $0 \leq l \leq l_1$, where l is the radial coordinate. For a static solution exhibiting spherical symmetry the metric can be taken in the general form given by (1.9). For this metric the non-vanishing T_μ^ν are given by equations (1.10). Since $T_4^1 = 0$ we must have

$$g_{44}(\rho+p)\frac{dl}{ds}\frac{d\tau}{ds} = 0 \qquad \text{.................. (2.1)}$$

In the interior $\rho+p \neq 0$, in general, and so we conclude that in the interior the fluid must be at rest in the reference frame, the 4-vector of velocity being

$$\left.\begin{aligned} \frac{dl}{ds} = \frac{d\theta}{ds} = \frac{d\phi}{ds} &= 0 \\ \frac{d\tau}{ds} &= e^{-\frac{1}{2}\nu} \end{aligned}\right\} \qquad \text{.................. (2.2)}$$

For the cases of $\rho \doteq \rho_s$ in which we are particularly interested it would be natural to have ρ varying continuously in the mass m, acquiring the value ρ_s at the boundary, and also a continuous variation of 4-velocity through the mass into the steady state substratum. Because we have chosen to consider a static interior and a discontinuous mass distribution, our solution must in general involve discontinuity of 4-velocity at $l = l_1$, since it would not in general be the natural motion of the exterior fluid to be at rest at the boundary. For instance in case (a) of Chapter I, Section (ii) the exterior fluid is at rest only on the neutral surface, and so falls with a 'splash' velocity on the mass m. In cases (b) and (c) the exterior fluid will in general have a non zero velocity outwards at $\ell = \ell_1$. No net energy is of course gained or lost by the sphere, as pointed out in Chapter I, since $\rho + p = 0$ in the exterior.

Substituting from (2.2) into the expression (1.12) for the T_μ^ν of a perfect fluid, we obtain

$$\left.\begin{aligned} T_1^1 = T_2^2 = T_3^3 &= -p \\ T_4^4 &= \rho \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \text{(2.3)}$$

Since ρ is constant, the last equation of (1.10) provides on integration

$$e^{-\lambda} = 1 - \frac{8\pi\rho}{3}\ell^2 + \frac{C}{\ell} \qquad (C \text{ constant})$$

There must be no singularity at $\ell = 0$, so that $C = 0$. Hence

$$e^{-\lambda} = 1 - \frac{\ell^2}{R^2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.4)}$$

where

$$\rho = \frac{3}{8\pi R^2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2.5)}$$

Since $T_1^1 = T_2^2$ the equating of the expressions for T_1^1, T_2^2 in (1.10) leads to the well known relation (<u>14</u>)

$$\frac{dp}{d\ell} = -(\rho + p)\frac{\nu'}{2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \text{(2.6)}$$

which on integration yields

$$\rho + p = De^{-\frac{1}{2}\nu} \qquad (D \text{ constant})$$

Substituting from (1.10) for ρ and p in this result we get finally, on integration,

$$e^{\frac{1}{2}\nu} = A - B\sqrt{1 - \ell^2/R^2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \text{(2.7)}$$

where A and B are constants. The metric for $\ell \leq \ell_1$, can thus be written

$$ds^2 = \left(A - B\sqrt{1 - \ell^2/R^2}\right)^2 dT^2 - \frac{d\ell^2}{1 - \ell^2/R^2} - \ell^2 d\theta^2 - \ell^2 \sin^2\theta \, d\phi^2 \qquad \ldots\ldots \text{(2.8)}$$

Thus far the interior solution is identical with the well known Schwarzschild interior solution for a sphere of perfect fluid of constant density surrounded by empty space (15). The different boundary conditions, however, lead to significantly different values of A and B. These boundary conditions we shall base simply on continuity of pressure and continuity of the metric coefficients at $\ell = \ell_1$. Substituting from (2.4) and (2.7) into the expression for T_1^1 in (1.10) we get for the interior pressure

$$p = \frac{1}{8\pi R^2}\left(\frac{3B\sqrt{1-\ell^2/R^2} - A}{A - B\sqrt{1-\ell^2/R^2}}\right) \qquad \text{...........(2.9)}$$

Putting $p = p_S$ when $\ell = \ell_1$, where $p_S = \dfrac{-3}{8\pi R_S^2}$ (eqn.(1.5)) we find

$$\frac{1}{A - B\sqrt{1-\ell_1^2/R^2}} = \frac{3R^2}{2A}\left(\frac{1}{R^2} - \frac{1}{R_S^2}\right) \qquad \text{.................(2.10)}$$

Continuity of the metric at $\ell = \ell_1$ demands that

$$\left.\begin{aligned}\left(A - B\sqrt{1-\ell_1^2/R^2}\right)^2 &= 1 - \frac{2m}{\ell_1} - \frac{\ell_1^2}{R_S^2} \\ &= 1 - \frac{\ell_1^2}{R^2}\end{aligned}\right\} \qquad \text{.................(2.11)}$$

whence by (2.10), (2.11)

$$1 - \frac{\ell_1^2}{R^2} = \frac{4A^2}{9\left(1 - R^2/R_S^2\right)^2}$$

so that

$$A = \pm\frac{3}{2}\left(1 - R^2/R_S^2\right)\sqrt{1-\ell_1^2/R^2}$$

and

$$B = \pm\frac{1}{2}\left(1 - 3R^2/R_S^2\right)$$

The signs correspond in the order given, so that there is no loss in generality in taking them both positive, in view of (2.8). Hence finally

$$\left.\begin{aligned} A &= \tfrac{3}{2}\left(1 - R^2/R_s^2\right)\sqrt{1 - \ell_1^2/R^2} \\ B &= \tfrac{1}{2}\left(1 - 3R^2/R_s^2\right) \end{aligned}\right\} \qquad \text{.........} \quad (2.12)$$

It is to be noted that our interior solution, which has been joined to the exterior solution of metric (1.8) by the boundary conditions stated, satisfies the boundary conditions which have been established by S. O'Brien and J. L. Synge (16) as the most satisfactory for discontinuous mass distributions in general relativity. For our case of spherical symmetry these boundary conditions require that, whether the metric be static or non-static,

$$g_{\alpha\beta}\,,\quad \frac{\partial g_{rs}}{\partial x^1}\,,\quad T^1_\alpha\,,\quad g_{\alpha\tau}T^\tau_\beta - g_{\beta\tau}T^\tau_\alpha \qquad \text{......} \quad (2.13)$$

should be continuous at the boundary. Here Greek letters represent indices 1,2,3,4 while Latin letters may represent indices 2,3,4 only, corresponding to coordinates θ, ϕ, T respectively. It is easily verified that our solution satisfies these requirements.

(ii) Physical interpretation of the solution

(a) Limitation of density and size

The proper radius of the sphere, as determined by the aggregate of local measurements by observers at rest in the reference frame, is

$$\delta = \int_0^{\ell_1} \frac{d\ell}{\sqrt{1 - \ell^2/R^2}} = R \sin^{-1}(\ell_1/R) \qquad \text{................} \quad (2.14)$$

It follows that

$$\delta \leq \pi R/2 \quad \text{.................(2.15)}$$

where $R = \sqrt{\frac{3}{8\pi\rho}}$ by (2.5). This provides a physical upper limit to the size of a sphere of given density, or an upper limit to the density of a sphere of given size. That is, only within these limits is it possible to find a system of internal stress to keep a uniform sphere in equilibrium in the steady state universe. A particular case of the upper limit is got when $\ell_1 = R = R_s$, so that the sphere of static fluid would become coincident with, and of the same density and pressure as, the whole of the idealised steady state universe within the observational horizon R_s. Since this is a limit of static equilibrium, and since the pressure gradient now vanishes, the contents of such a sphere will actually follow the natural geodesic motion which is now possible and expand.

It is clear that R_s is an upper bound to the bounding coordinate ℓ_1 of any sphere, to preserve the signature of the steady state metric. We must also have $\ell_1 \leq R$ to preserve the signature of the interior metric. In the case when $R > R_s$, corresponding to $\rho < \rho_s$, we should require, consequently, ℓ_1 definitely less than R and so δ definitely less than $\pi R/2$.

<u>(b) Gravitational mass</u>

The quantity m in the case of the Schwarzschild solution for a sphere in empty space is identified as the gravitational mass of the sphere, equal to its inertial mass. But in our case when the steady state cosmic background is taken into account this matter needs review.

By (2.11) we find

$$m = \frac{l_1^3}{2}\left(\frac{1}{R^2} - \frac{1}{R_S^2}\right) \qquad \text{.................(2.16)}$$

so that $m \gtrless 0$ according as $R \lessgtr R_S$, or $\rho \gtrless \rho_S$ ..(2.17)
We have pointed out on the basis of equation (1.22) that the sign of m determines whether a particle, moving on a radial geodesic in the field of uniform stress external to the sphere, has an additional acceleration due to the sphere which opposes cosmic repulsion or assists it. From (2.17) the criterion is therefore whether the sphere is a concentration of mass or a rarefaction relative to the smoothed out steady state model. We shall throw more light on the quantity m by calculating the total gravitational mass within the sphere.

Our metric (2.8) forms a particular case for which the formula for gravitational mass obtained by E. T. Whittaker (<u>17</u>) is valid. Accordingly, the total gravitational mass in the sphere is

$$\int_0^{l_1}\int_0^{\pi}\int_0^{2\pi}\sqrt{-g}\left(T_4^4 - T_1^1 - T_2^2 - T_3^3\right)dl\,d\theta\,d\phi \qquad \text{..........(2.18)}$$

where g is the determinant of the $g_{\mu\nu}$. This is

$$4\pi\int_0^{l_1}\left(\frac{A - B\sqrt{1-l^2/R^2}}{\sqrt{1-l^2/R^2}}\right)(\rho+3p)\,l^2\,dl$$

Now $$\rho + 3p = \frac{3B\sqrt{1-l^2/R^2}}{4\pi R^2\left(A - B\sqrt{1-l^2/R^2}\right)} \qquad \text{.............(2.19)}$$

Hence gravitational mass

$$= \frac{3B}{R^2}\int_0^{\ell_1} \ell^2 d\ell$$

$$= \tfrac{1}{2}\left(1 - 3R^2/R_S^2\right)\frac{\ell_1^3}{R^2} \qquad \text{..............(2.20)}$$

which can be written $= m - \dfrac{\ell_1^3}{R_S^2}$ ....(2.21)

on using (2.11).

We observe that the second term represents the gravitational mass in the sphere which would exist if the sphere were a portion of the steady state universe with the same bounding coordinate ℓ_1 (on putting $R = R_S$ and $m = 0$ in the above integral). Thus the quantity m associated with metric (1.8) represents the excess of ~~its~~ the total gravitational mass over that which would occupy the same coordinate region in the steady state.

The <u>total</u> gravitational mass of the sphere is seen to be positive, zero, or negative according as

$$\rho \gtreqless 3\rho_S \qquad \text{..............(2.22)}$$

We note from the above integral that this statement applies also to any smaller concentric spherical portion of the given sphere. Thus it would follow that at the critical condition $\rho = 3\rho_S$ the gravitational field intensity vanishes at all points of the sphere. The pressure for this case is seen from (2.9) to be uniform and equal to p_S . Thus in this neutral state the sphere is identical with a portion of the Einstein universe (for the case $\Lambda = 0$), and the neutral surface $\ell = \ell_0$

described in Chapter I (equation (1.23)) coincides with the boundary $\ell = \ell_1$ so that the exterior fluid is also at rest there.

If $\rho \neq 3\rho_s$ then there is a gravitational force acting at any point of the sphere and so a gradient of pressure is in general necessary to keep the sphere in the equilibrium that we have postulated. We shall examine this pressure distribution.

(c) Distribution of pressure

From equation (2.9) for the pressure p at any point of the sphere we see that if it is not to become infinite at some point then the expression $A - B\sqrt{1-\ell^2/R^2}$ must never vanish. In view of the metric (2.8) this would involve a singularity in the metric. We can assume that such a circumstance would not be physically possible, that is the physical situation would adjust itself to prevent it. At the boundary $\ell = \ell_1$ the expression is

$$A - B\sqrt{1-\ell_1^2/R^2} = \sqrt{1-\ell_1^2/R^2} > 0$$

Hence we require $A - B\sqrt{1-\ell^2/R^2} > 0$ for all ℓ. That is

$$\tfrac{3}{2}\left(1-R^2/R_s^2\right)\sqrt{1-\ell_1^2/R^2} - \tfrac{1}{2}\left(1-3R^2/R_s^2\right)\sqrt{1-\ell^2/R^2} > 0 \quad \ldots(2.23)$$

(i) If $0 < \frac{R^2}{R_s^2} \leq \frac{1}{3}$ the condition requires

$$\sqrt{1-\ell^2/R^2} < \frac{3\left(1-R^2/R_s^2\right)\sqrt{1-\ell_1^2/R^2}}{1-3R^2/R_s^2}$$

which holds for all ℓ if it holds at $\ell = 0$. That is if

$$\sqrt{1-\ell_1^2/R^2} > \frac{1-3R^2/R_s^2}{3\left(1-R^2/R_s^2\right)}$$

The expression on the right is monotonic decreasing from the value of $1/3$ when $R^2/R_s^2 \simeq 0$ (ρ infinite) to the value zero when $R^2/R_s^2 = 1/3$ ($\rho = 3\rho_s$)

When $R^2/R_s^2 \simeq 0$ therefore we require

$$\ell_1/R < \sqrt{8/9}$$

or by (2.14) $\delta/R < \sin^{-1}\sqrt{8/9}$

while for $R^2/R_s^2 = 1/3$ we require simply

$$\delta/R < \pi/2$$

which is equation (2.15).

(ii) If $\frac{1}{3} \leq \frac{R^2}{R_s^2} \leq 1$, it is clear that (2.23) is always satisfied.

(iii) If $\frac{R^2}{R_s^2} > 1$, we require

$$\sqrt{1-\ell^2/R^2} > \frac{3\left(R^2/R_s^2 - 1\right)\sqrt{1-\ell_1^2/R^2}}{3R^2/R_s^2 - 1}$$

which holds for all ℓ if it holds for $\ell = \ell_1$ which is clearly so.

We shall assume therefore that

$$A - B\sqrt{1-\ell^2/R^2} > 0 \text{, for all } \ell \quad \text{.........} \quad (2.24)$$

From (2.9) we find

$$\frac{dp}{d\ell} = -\frac{AB\ell}{4\pi R^4\sqrt{1-\ell^2/R^2}\left(A - B\sqrt{1-\ell^2/R^2}\right)^2} \quad \text{.........} \quad (2.25)$$

which is $\gtrless 0$ according as $AB \lessgtr 0$ .. (2.26)

Case I: $0 < R^2/R_s^2 < 1/3$

In this case $dp/d\ell < 0$, and since $p = p_s < 0$ at $\ell = \ell_1$ the pressure will vanish at $\ell = \ell^*$ where

$$3B\sqrt{1-\ell^{*2}/R^2} - A = 0$$

i.e.

$$\sqrt{1-\ell^{*2}/R^2} = \frac{(1-R^2/R_S^2)\sqrt{1-\ell_1^2/R^2}}{1-3R^2/R_S^2} \qquad \text{..............(2.27)}$$

Clearly such an ℓ^* always exists in the range $0 \le \ell^* < \ell_1$ The pressure then becomes positive for $\ell < \ell^*$ and increases as ℓ decreases.

Case II: $R^2/R_S^2 = 1/3$

Here, as pointed out in (b), the pressure takes the constant value

$$\left.\begin{aligned} p = p_s &= \frac{-3}{8\pi R_s^2} \\ \text{while } \rho = 3\rho_s &= \frac{9}{8\pi R_s^2} \end{aligned}\right\} \qquad \text{.......(2.28)}$$

The sphere in these circumstances would have the same characteristics as a portion of the Einstein universe (for the case $\Lambda = 0$). In fact the metric of the interior solution would be

$$ds^2 = d\tau^2 - \frac{d\ell^2}{1-\ell^2/R^2} - \ell^2 d\theta^2 - \ell^2 \sin^2\theta\, d\phi^2 \qquad \text{......(2.29)}$$

which is well known to be that of the Einstein universe, if it were supposed that the solution were valid for all $\ell \le R$.

Case III: $1/3 < R^2/R_S^2 < 1$

For this case $dp/d\ell > 0$ so that since $p = p_s < 0$ at $\ell = \ell_1$ the pressure is always negative, its least value being at $\ell = 0$.

Case IV: $R^2/R_S^2 = 1$

The sphere is now identical with a portion of the steady state universe as regards its uniform density ρ_S and uniform pressure p_S. Since the pressure gradient is zero the contents of the sphere are free to take up the natural geodesic motion that is now possible, that is to expand under the repulsive gravitational field which exists when $\rho < 3\rho_S$ (equation (2.22)). We shall return to this point in (d).

Case V: $R^2 > R_S^2$

Here $dp/dl < 0$, so that the pressure might become zero at some internal point, if it were possible for the equation

$$3B\sqrt{1 - l^2/R^2} - A = 0$$

to have a solution for $l < l_1$. This would clearly be impossible when $R^2/R_S^2 > 1$. Thus the pressure in this case is always negative but increases towards the centre.

(d) Positive and negative inertia

It may be regarded as surprising that when ρ decreases through the value ρ_S the pressure gradient dp/dl, necessary to maintain equilibrium, changes sign from positive to negative, despite the fact that the gravitational force, as determined by the gravitational mass within the radius l, varies continuously remaining negative. This is because the density of inertia changes sign from positive to negative when ρ decreases through ρ_S, so that there is an inversion of the effects of motivating forces.

The equations of mechanics in general relativity are expressed by equating to zero the covariant divergence of the $T^{\mu\nu}$ tensor. For a perfect fluid this tensor is given in the form T^{ν}_{μ} by equation (1.12). It is easily shown that in a local system of coordinates which is instantaneously co-moving with the fluid, but geodesic - that is 'freely falling' unaffected by any stress gradient existing in the fluid, three of the resulting equations yield the momentum equation

$$\underline{\nabla} p + (\rho + p)\frac{d\underline{q}}{dt} = 0, \qquad \ldots\ldots\ldots(2.30)$$

where $d\underline{q}/dt$ is the acceleration of the fluid, and $\underline{\nabla} p$ is the gradient of pressure, in the local system. The fourth equation is the energy equation

$$\frac{\partial \rho}{\partial t} + (\rho + p)\, \mathrm{div}\, \underline{q} = 0, \qquad \ldots\ldots(2.31)$$

$\mathrm{div}\, \underline{q}$ being the local spatial divergence of the velocity of the fluid.

These results are well known in various forms, but their application to a medium of negative p of the same order as ρ has not previously been envisaged. The momentum equation shows that the proper density of inertia in a stressed isotropic medium is $\rho + p$; this takes account of the contribution to momentum of the stress in a moving medium. The energy equation takes account of the energy equivalent of the work done by the pressure.

For our interior solution we find by (2.5), (2.9)

$$\rho + p = \frac{A}{4\pi R^2\left(A - B\sqrt{1 - \ell^2/R^2}\right)} \qquad \ldots\ldots (2.32)$$

which, since $A - B\sqrt{1 - \ell^2/R^2} > 0$, takes the sign of A.

Thus $\rho + p \gtrless 0$ according as $\rho \gtrless \rho_s$ ...... (2.33)

The density of inertia in the smoothed out steady state model is therefore zero. In addition, if the density of energy ρ differs but little from ρ_s in a region of concentration or rarefaction of mass in the steady state universe, then only a very small pressure gradient is necessary to keep the mass in equilibrium, or in a state of unaccelerated expansion. Further, equation (2.30) indicates that when $\rho + p < 0$ there must be a gradient of pressure inwards in our sphere of fluid, in order to produce an acceleration in the <u>same</u> direction relative to a system of coordinates following a geodesic outwards. In other words, the fluid is at rest in our coordinate system because the outward gravitational force is balanced, when $\rho + p < 0$, by an outward force of pressure. When $\rho + p > 0$ the outward gravitational force (provided $\rho < 3\rho_s$) is balanced by an inward force of pressure. The familiar case is $\rho > 3\rho_s$ when an inward gravitational force is balanced by an outward force of pressure.

Equation (2.31) indicates that the density of energy ρ remains constant in the steady state, despite the expansion ($\operatorname{div} \underline{q} > 0$), since $\rho + p = 0$. This is because the degradation of energy due to expansion is exactly compensated by the energy equivalent of the work done by the negative pressure. This is the interpretation which was given by W. H. McCrea (<u>1</u>). We note further that when $\rho + p > 0$, ρ decreases with expansion, and when $\rho + p < 0$, ρ increases.

CHAPTER III: THE MECHANISM OF THE STEADY STATE

(i) Introduction

In this chapter we put forward a theory of the mechanism of steady state cosmology according to general relativity. It is designed to provide a physical basis for the zero-point stress introduced by McCrea, and also for an associated zero-point energy, in a manner which seems to be naturally suggested by the provision for negative states of energy in present quantum theory.

A tentative theory of the mechanism of creation of matter, which preserves the principles of conservation of energy and momentum and is based on quantum theory and general relativity, has already been outlined by F. Pirani (4). Pirani accounts for the existence of a zero-point stress and a zero-point energy by postulating the existence of entities of zero rest mass and negative energy which he calls 'gravitinos' and which he tentatively identifies with neutrinos. For this purpose he uses the results obtained by J. L. Synge (18) regarding the contribution to the energy-momentum tensor in general relativity of entities of zero rest mass and negative energy, referred to by Synge as 'attractive impulses'.

The present writer, independently, had formulated similar ideas, suggested by Synge's attractive impulses, of giving a mechanical basis to McCrea's zero-point stress and introducing an associated negative energy term in the energy-momentum tensor. These ideas arose in connection with certain tentative work on ordinary radiation in cosmology. In that work the same expression, with sign changed, was obtained for the rate of emission of radiation per unit proper volume, provided we neglect the pressure of matter as Pirani also does, as has been given by Pirani for the rate of creation of matter. The idea of sustaining the steady state by the conversion of radiation to matter presented itself at that time, but it was realised that unless the density of radiation was negative then expansion would diminish radiation as well as matter.

It was clear from the work of Synge, however, that negative radiation could be treated on the same basis as ordinary radiation, and that expansion would lead to an increase in density (less negative) in this case, due to the positive work done by the negative pressure. The extra available radiation energy could then be converted to replenish the lowered density of matter, by quantum collision processes in which energy and momentum were conserved. If the energy density of negative radiation were of the same order as the average density of matter, this rate of production of energy available for conversion could be adequate to maintain the steady state, if some stabilising mechanism were assumed to sustain it.

In Pirani's theory the creation process occurs, presumably spontaneously, with conservation of energy-momentum secured by the emission of equal amounts, but opposite in sign, of

4-momentum in the form of matter and gravitinos. To represent the presence of gravitinos he incorporates a radiation term into the energy-momentum tensor of general relativity, on the basis of the work of Synge. The usual equation $T^{\mu\nu}_{\ ;\nu} = 0$ then achieves formal conservation of energy, with a physical interpretation of zero-point stress and energy in terms of gravitinos.

The steady state is not singled out for special consideration in Pirani's theory, which does not indicate any particular merit for it. It is not shown whether the rate of creation depends on local conditions, and particularly in the case of the steady state no reason appears why it would be just right to maintain the steady state. Annihilation plays no part in his theory, nor is any allowance made for the possible relation of cosmic rays or ordinary radiation to the creation process.

In this chapter we shall allow for the most general type of collision permitted by quantum mechanics leading to creation or annihilation events. Radiation will be regarded as having a continuous spectrum of energy values from negative to positive, physical effects depending on the sign and intensity at any point. Thus in stars positive radiation will greatly predominate while it will be vice versa in the vast intergalactic regions where the density of matter is small.

The pressure of matter will not be neglected, as by Pirani, since we shall make allowance for the presence of cosmic rays. Since the kinetic energy of cosmic rays would steadily diminish in an expanding universe we make the hypothesis that it is steadily maintained by the creation of matter, at least in part, in the form of cosmic rays.

It being improbable that cosmic rays can arise solely from the stars, we shall find it satisfactory that our theory predicts that the bulk of matter is created where the density is lowest - in intergalactic space.

The balance of a continuous distribution of matter and negative radiation in any quasi-steady state will be deemed to be such that there is dynamic equilibrium in annihilation and creation rates. The stability of the steady state will appear, in the light of this supposition, to be a fundamental consequence of the equations of general relativity.

(ii) The pressure and density of cosmic radiation

We first examine on the basis of general relativity, how the integrated intensity of radiation depends on the particular model adopted from those whose general metric is given by equation (1.1). We shall find the intensity at the spatial origin $r=0$ (arbitrary). The radiation we suppose to include a negative component, whose momentum is directed into the past and whose relative energy is $-h\nu$ per quantum of frequency ν , where h is Planck's constant. It is assumed that the radiation emitted from unit proper volume at r, θ, ϕ, at cosmological time t , will be partially absorbed by the time t_0 when part of it reaches the origin, due to collisions. Such collisions will be those permitted by quantum mechanics, involving photons and material particles. In these collisions radiation will be absorbed by atoms, to increase their intrinsic energy if the radiation is positive, or to reduce it if the radiation is negative. Positive radiation may be materialised also in the form of newly created fundamental particles, while negative radiation may be responsible for the complete annihilation of fundamental particles. The emission itself will be the reversal of these processes.

Let the energy of radiation radiated from unit proper volume in the homogeneous and isotropic model, between epochs $t, t+\delta t$ in the frequency range $\nu, \nu+\delta\nu$ as a result of these general 'collision' processes which occur in that volume, be (suffix r referring to 'radiation')

$$\mathcal{E}_r(\nu,t)\,\delta\nu\,\delta t \qquad \text{........(3.1)}$$

If a photon passing through this volume is involved in collision we suppose it to be completely absorbed, and the residual radiation, if any, actually re-emitted we suppose to be part of the emission spectrum of that volume. Suppose that the proportion of $\mathcal{E}_r(\nu,t)\delta\nu\,\delta t$ not absorbed by the time t_0 is

$$\lambda(\nu,\tau,t)\mathcal{E}_r(\nu,t)\,\delta\nu\,\delta t \qquad \text{.......(3.2)}$$

where $\tau = t_0 - t$. We have of course

$$\lambda(\nu,0,t_0) = 1 \qquad \text{...................(3.3)}$$

corresponding to putting $t = t_0$.

At time t_0, however, the apparent energy of this radiation, relative to fundamental observers located in the spherical shell containing the radiation at that epoch, will be reduced because of redshift to

$$\lambda(\nu,\tau,t)\mathcal{E}_r(\nu,t)\,\delta\nu\,\delta t.\,\frac{R(t)}{R(t_0)} \qquad \text{..........(3.4)}$$

(Part I eqn. (1.7)). The area of the spherical surface (total proper area) will be

$$4\pi R^2(t_0) r^2 \Big/ (1 + kr^2/4)^2 \qquad \text{.........(3.5)}$$

The time for the shell to pass over the origin will be

$$\delta t_0 = \delta t . \frac{R(t_0)}{R(t)} \qquad \text{.........(3.6)}$$

(Part I eqn. (1.5)). Thus the rate at which radiation of all frequencies from unit proper volume at (r, θ, ϕ), emitted at epoch t , passes through unit area at the origin normal to the direction $\theta = 0$ is

$$\frac{\int_0^\infty \lambda(\nu, T, t) \mathcal{E}_r(\nu, t)\, d\nu \ \frac{R^2(t)}{R^2(t_0)} \cos\theta}{\left\{ 4\pi R^2(t_0) r^2 \Big/ (1 + kr^2/4)^2 \right\}} \qquad \text{.......(3.7)}$$

To get the contribution to the total intensity of radiation at the origin from the total volume between the values r, $r + \delta r$ of the radial coordinate we must multiply (3.7) by

$$\frac{R^3(t)\, r^2 \sin\theta\, \delta r\, \delta\theta\, \delta\phi}{(1 + kr^2/4)^3} \qquad \text{.......(3.8)}$$

and integrate with respect to θ between 0 and $\pi/2$, and with respect to ϕ between 0 and 2π . The result is

$$\frac{1}{4} \int_0^\infty \lambda(\nu, T, t) \mathcal{E}_r(\nu, t)\, d\nu . \ \frac{R^5(t)}{R^4(t_0)} \cdot \frac{\delta r}{1 + kr^2/4} \qquad \text{........(3.9)}$$

Now a photon of the radiation through the origin has the world line

$$\int_t^{t_0} \frac{dt}{R(t)} = \int_0^r \frac{dr}{1 + kr^2/4} \qquad \text{.........(3.10)}$$

so that, keeping t_0 fixed we have

$$\frac{\delta r}{1+kr^2/4} = -\frac{\delta t}{R(t)} \qquad \text{...........(3.11)}$$

Thus the contribution from the volume between $r, r+\delta r$ can be written

$$-\frac{1}{4}\int_0^\infty \lambda(\nu,T,t)\,\mathcal{E}_r(\nu,t)\,d\nu \cdot \frac{R^4(t)}{R^4(t_0)}\,\delta t \qquad \text{.........(3.12)}$$

We suppose now that the local time of emission, on the observational horizon of the model, of radiation which will arrive at the origin at time t_0 is $t^*(t_0)$. Then the total intensity of radiation at any point of the model at epoch t_0 is

$$I(t_0) = \frac{1}{4R^4(t_0)}\int_0^\infty \int_{t^*(t_0)}^{t_0} \lambda(\nu,T,t)\,\mathcal{E}_r(\nu,t)\,R^4(t)\,d\nu\,dt \qquad \text{........(3.13)}$$

The isotropic pressure of radiation in the model is therefore

$$p_r(t_0) = \frac{4\,I(t_0)}{3}, \qquad \text{....................(3.14)}$$

using local thermodynamics, while the corresponding energy density of the radiation will be

$$\rho_r(t_0) = 3\,p_r(t_0) = 4\,I(t_0) \qquad \text{...............(3.15)}$$

The sign of these quantities will depend on the average predominance of positive or negative radiation.

<u>(iii) Conservation principles in the cosmic expansion</u>

We now investigate the creation process according to the conservation principles contained in the equations of general relativity. We shall apply these equations directly to the cosmological model of the general type that we are considering. However, it is a remarkable demonstration of the inner consistency of kinematical relations in general relativity that we can achieve local conservation of energy for radiation alone, in the form of the first law of thermodynamics, directly from the previous investigation, as follows.

By (3.13) and (3.15) we have

$$\rho_r(t_0) = \frac{1}{R^4(t_0)}\int_0^\infty\int_{t^*(t_0)}^{t_0}\lambda(\nu,\tau,t)\,\mathcal{E}_r(\nu,t)\,R^4(t)\,d\nu\,dt \qquad \text{...........(3.16)}$$

On differentiating this expression with respect to t_0 , we obtain

$$\frac{d\rho_r(t_0)}{dt_0} = \int_0^\infty \lambda(\nu,0,t_0)\,\mathcal{E}_r(\nu,t_0)\,d\nu - 4\,\rho_r(t_0)\cdot\frac{1}{R(t_0)}\frac{dR(t_0)}{dt_0} - \frac{R^4(t^*)}{R^4(t_0)}\frac{dt^*}{dt_0}\int_0^\infty \lambda(\nu,\tau^*,t^*)\,\mathcal{E}_r(\nu,t^*)\,d\nu + \frac{1}{R^4(t_0)}\int_0^\infty\int_{t^*}^{t_0}\frac{\partial\lambda}{\partial t_0}\,\mathcal{E}_r(\nu,t)\,R^4(t)\,d\nu\,dt \qquad \text{...........(3.17)}$$

where $\tau^* \equiv t_0 - t^*$, and $\frac{\partial\lambda}{\partial t_0}$ means partial differentiation with respect to t_0 keeping t constant. The third term on the right may vanish for several reasons:

(a) $R(t^*) = 0$, which is the case in the steady state model where $R(t) = e^{t/R_s}$ and $t^* = -\infty$, and also for other well known cosmological solutions.

(b) $dt^*/dt_0 = 0$, which is also the case in the steady state, and also in models where $t^* = 0$.

(c) $\int_0^\infty \lambda(\nu,\tau^*,t^*)\,\mathcal{E}_r(\nu,t^*)\,d\nu = 0$ which means that the net amount of radiation reaching the origin from the horizon is zero. This could happen in particular if $\lambda(\nu,\tau^*,t^*) = 0$ for all ν , so that radiation emitted at the horizon is entirely 'absorbed' (i.e. involved in collisions before $t = t_0$, according to our definition). It could also happen if $\mathcal{E}_r(\nu,t^*) = 0$ for all ν , which means that at the cosmological time t^* , the theoretical epoch of emission on the horizon, there was in fact no radiation emitted in the model.

We shall in fact assume that each, or at least the sum, of the last two terms on the right of (3.17) vanishes. ~~We shall assume, therefore, that one or more of these conditions obtain.~~ The alternative violates local thermodynamics as will be evident from what follows, and this provides a criterion for acceptable cosmological models.

Noting that the first term on the right of (3.17) is the total rate of emission from unit volume at time t_0, viz.

$$E_r(t_0) = \int_0^\infty \lambda(\nu, 0, t_0)\, \mathcal{E}_r(\nu, t_0)\, d\nu\,, \quad (\text{noting } (3\cdot 3)) \qquad \ldots\ldots\ldots(3.18)$$

and dropping the suffix o, we get at general time t

$$\frac{d\rho_r(t)}{dt} = E_r(t) - 4\rho_r(t)\cdot\frac{\dot{R}(t)}{R(t)} \qquad \ldots\ldots\ldots(3.19)$$

where $\dot{R}(t) \equiv d\{R(t)\}/dt$. Consider now an element of proper volume v whose boundaries move with the fluid contents of the model (i.e. in its mean motion), then by (1.1)

$$\frac{1}{v}\frac{dv}{dt} = 3\frac{\dot{R}}{R} \qquad \ldots\ldots\ldots\ldots(3.20)$$

Whence by (3.15), (3.19) and (3.20)

$$\begin{aligned} v E_r\, dt &= v\, d\rho_r + \tfrac{4}{3}\rho_r\, dv \\ &= d(\rho_r v) + p_r\, dv \end{aligned}$$

or,

$$dQ_r = d(\rho_r v) + p_r\, dv \qquad \ldots\ldots\ldots\ldots(3.21)$$

where $dQ_r = v E_r\, dt$ is the total energy of radiation emitted in the volume v during the interval $t, t+dt$.

This equation expresses the conservation of energy of radiation in the model as it expands, in that dQ_r is accounted for by the increase $d(\rho_r v)$ of the radiation energy in volume v

plus the work done by the pressure p_r during the expansion.

The total energy density ρ and pressure p in the model can be resolved into

$$\left.\begin{array}{l}\rho = \rho_r + \rho_m \\ p = p_r + p_m\end{array}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.22)$$

where ρ_m is the relative density of material energy (which includes the kinetic energy of matter) as measured by fundamental observers, and p_m is the material pressure. These quantities will include the contribution of cosmic rays. For the metric (1.1) ρ and p are given by (1.2) and (1.3). These equations combine to give

$$\frac{d}{dt}(\rho R^3) + p\,\frac{d(R^3)}{dt} = 0 \qquad \ldots\ldots\ldots\ldots(3.23)$$

or

$$\frac{d(\rho v)}{dt} + p\,\frac{dv}{dt} = 0 \qquad \ldots\ldots\ldots\ldots(3.24)$$

where v is the element of proper volume already introduced. Multiplying by dt we may write

$$d(\rho_r v) + p_r\,dv + d(\rho_m v) + p_m\,dv = 0$$

or

$$dQ_r + dQ_m = 0 \qquad \ldots\ldots\ldots\ldots(3.25)$$

where

$$dQ_m = d(\rho_m v) + p_m\,dv \qquad \ldots\ldots\ldots\ldots(3.26)$$

The last equation is the conservation equation for material energy in the model analogous to (3.21) for radiation, while (3.25) shows how, according to general relativity, all energy is shared inevitably between material energy and radiation energy.

Before the possibility of negative radiation was considered emission of radiation was associated with positive dQ_r and therefore negative dQ_m, so that emission of radiation seemed inevitably to mean a decrease of material energy. But in the event of the average emission of negative radiation in space predominating over that of positive radiation, we have

$$dQ_m = -dQ_r = d(\rho_m v) + p_m dv > 0 \qquad \text{.........(3.27)}$$

Now if p_m is small as appears from observation to be the case then $d(\rho_m v) > 0$. That is there will be a net rate of increase of material energy in any volume v whose boundaries move with the mean motion of matter.

(iv) The Steady state as one of dynamical equilibrium

The density and pressure in the particular case of the steady state are given by (eqn. (1.5))

$$\rho = -p = \frac{3}{KR_s^2} \qquad \text{..................(3.28)}$$

whence, by (3.22), if regard the pressure of matter p_m as small on the average - of the order of the pressure of positive radiation, we can write

$$\left.\begin{aligned} \rho_r = 3p_r &\doteq \frac{-9}{KR_s^2} \\ \rho_m &\doteq \frac{12}{KR_s^2} \end{aligned}\right\} \qquad \text{..............(3.29)}$$

Also

$$\left.\begin{aligned} \rho &\doteq \tfrac{1}{4}\rho_m \\ p &\doteq -\tfrac{1}{4}\rho_m \end{aligned}\right\} \qquad \text{.............(3.30)}$$

Thus the average total density of energy is positive being approximately one quarter of the density of energy in the form of matter (including its associated kinetic energy). The total pressure is the negative of this quantity, due to the predominance of negative radiation as the main source of stress. Thus the admission of negative energy states and in particular negative radiation provides a physical basis for McCrea's zero point stress and an associated zero-point energy. The maintenance of the steady state will now be investigated on this physical basis.

Equation (3.24) is the equation of energy for the cosmological models, which may be written in the form

$$\frac{d\rho}{dt} + (\rho + p)\frac{1}{v}\frac{dv}{dt} = 0 \qquad \text{...............(3.31)}$$

when it is seen to agree with the general energy equation (2.31), applicable in a local geodesic coordinate frame to any physical system of isotropic stress. As pointed out in the context of (2.31), if $\rho + p = 0$ at any instant, as we derive for the steady state metric, then the energy equation shows that ρ remains constant despite expansion. The maintenance of the density has therefore to be interpreted, as by McCrea, as due to the fact that new energy is created as the equivalent of the work done by the zero-point stress. As has been said, no physical basis was given for this mechanism by McCrea; but since the density of radiation was regarded as negligible in McCrea's analysis so that $\rho \doteq \rho_m$, he was led to argue that the mass equivalent of the work done by zero-point stress must appear in the form of newly created matter, in order to maintain a steady state.

According to the development in this thesis the following mechanism suggests itself. Although adiabatic expansion, without radiation emission or absorption, would certainly lead to a decrease in the energy density of matter, this decrease including the adiabatic degradation of the energy of cosmic rays, the negative density of radiation (given by (3.29)) would however be increased towards zero because of the adiabatic effect of its negative stress. That is, supposing there to be no emission or absorption, equations (3.21), (3.26) would yield for a volume v

$$\left.\begin{aligned} dQ_r &= d(\rho_r v) + p_r\, dv = 0 \\ dQ_m &= d(\rho_m v) + p_m\, dv = 0 \end{aligned}\right\} \qquad \text{........(3.32)}$$

Hence

$$\left.\begin{aligned} d\rho_r &= -(\rho_r + p_r)\frac{dv}{v} \\ d\rho_m &= -(\rho_m + p_m)\frac{dv}{v} \end{aligned}\right\} \qquad \text{........(3.33)}$$

Thus

$$d\rho_r = -d\rho_m > 0 \qquad \text{........(3.34)}$$

It follows from (3.34) that the equilibrium between matter and radiation would be upset from the balance which we must logically associate with the steady state if it is to be maintained. The relative densities of matter and radiation for dynamic equilibrium in the steady state are given by (3.29), and will be determined by the collision cross sections of the particles involved. If the steady state is to be maintained it must be assumed that the shift from equilibrium must be in the direction of a net rate of creation of matter, including cosmic ray energy. This will continue with corresponding reduction in radiation energy (i.e. becoming more negative), according to the relation (3.25), until the equilibrium is restored with ρ_r and ρ_m at their steady state values.

In view of the analysis in Chapter I leading to equations (1.14) and (1.15) the relative densities of the steady state can be assumed toprevail in the vast intergalactic regions of space. The activity of negative radiation will therefore predominate over that of positive radiation in these regions. Its emission must be associated with creation of matter and its absorption with annihilation of matter. Since its absorption rate at least would be proportional to the product of ρ_m and $|\rho_r|$ we can understand that the relative change of densities indicated by (3.34) would lead to a reduction in the rate of annihilation. Provided that this is not more than offset by a possible reduction in the rate of emission a net rate of creation of matter would therefore set in, as envisaged above.

(v) The stability of the steady state

To show how the theory given here can provide a basis for the stability of the steady state, because of automatic adjustment of the densities of matter and radiation to secure dynamic equilibrium, we shall consider small fluctuations of density from that of the steady state value ρ_s. We make use of the stationary interior and exterior solutions established in Chapters I, II for concentrations or rarefactions of mass in the steady state universe.

The basic interpretation of these solutions that we make for our purpose is that since they are stationary there must be dynamic equilibrium between the matter and radiation present. Accordingly, the interior solution provides the equilibrium pressure distribution, due to matter and radiation, corresponding to a given uniform distribution of mass within a sphere outside of which steady state conditions obtain. These results were admittedly established only for static conditions in the interior of the sphere, but we suppose that even if the sphere were expanding the rate of expansion in any limited region (intergalactic say) of the expanding universe must be so slow that these equilibrium relations between density and pressure, established in Chapter II, would be approximately applicable at any instant. In fact we found in Chapter II that in the particular case when $\rho = \rho_s$, even if the fluid was regarded as static, the pressure could only be p_s, as if the fluid were actually moving geodesically in the steady state.

Suppose, therefore, that in a limited region of the steady state universe there is a small fluctuation of density which does not upset the general expansion of the region. Then by (2.33) we find that for dynamic equilibrium the density of inertia $\rho+p$ satisfies the relation

$$\rho+p \gtrless 0 \qquad \text{according as} \quad \rho \gtrless \rho_s \qquad \ldots\ldots\ldots(3.35)$$

The local energy equation (2.31), or (3.31), therefore leads correspondingly to

$$\frac{d\rho}{dt} \lessgtr 0 \qquad \text{according as} \quad \rho \gtrless \rho_s \qquad \ldots\ldots\ldots(3.36)$$

The interpretation is that if $\rho+p>0$ then the expansion carries energy outwards relative to a static reference frame set up locally i.e. $T_4^1>0$ (c.f.(1.16)), and vice versa if $\rho+p<0$. It follows that the effect of expansion in either case is to restore the equilibrium density, and hence the pressure, to the steady state value.

In addition we remark that the interior solution of Chapter II leads to a non zero value for the parameter m if $\rho\neq\rho_s$. This parameter was identified as the excess of gravitational mass in a fixed region of space over that corresponding to an equal volume of the steady state universe. According to the analogue of Birkhoff's theorem, dealt with in Chapter I, Section (iv), this excess remains despite expansion, unless there is an actual radiation transfer to the exterior. Such a radiation transfer of energy would, in the nature of things, be in the sense determined by the sign of $\rho+p$, so that the process would be a part of the restoration of the steady state. In any case the ratio of the excess to the total gravitational mass in a region whose boundaries move with the fluid must tend to zero with expansion.

(vi) The rate of creation of material energy in the steady state

The density of matter ρ_m in the steady state, according to our theory, has been given in (3.29) as approximately $12/\kappa R_S^2$, if we neglect the pressure of matter. This agrees with Pirani's result for the particular case of the steady state model. Since McCrea on the other hand makes no physical interpretation of zero-point stress he assumes that $\rho = \rho_m$ so that, as given for our ρ in equation (3.28), he obtains $\rho_m = 3/\kappa R_S^2$. As Pirani remarks this discrepancy would provide in principle a basis for testing the hypothesis of radiation of negative energy, assuming the steady state. It should be noticed that Hoyle's result for ρ_m agrees with McCrea's, since he also assumes $\rho = \rho_m$ and neglects any form of stress in his model.

According to our equation (3.26) the rate of creation of material energy in a volume v is

$$\frac{dQ_m}{dt} = \frac{d(\rho_m v)}{dt} + p_m \frac{dv}{dt} \qquad \text{......(3.37)}$$

for the general models of metric (1.1). Consequently for the steady state model in which ρ_m is constant, taking $v = 1$, we find for the rate of creation of material energy in unit volume

$$E_m = 3(\rho_m + p_m)/R_S \qquad \text{................(3.38)}$$

on using (3.20). By equation (3.25) this must be equal to the rate of decrease of radiation energy in unit volume, so that the rate of creation of radiation energy, derived similarly to be

$$E_r = 3(\rho_r + p_r)/R_s = 4\rho_r/R_s \quad \ldots\ldots\ldots\ldots(3.39)$$

also satisfies

$$E_m = -E_r \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.40)$$

Thus $$E_m = -4\rho_r/R_s \quad \ldots\ldots\ldots\ldots\ldots\ldots(3.41)$$

The result given by (3.41) agrees with that of Pirani for the steady state. However the result (3.38) does not, since Pirani neglects the pressure of matter. By neglecting p_m he ignores that part of the material energy which has to be created to restore the energy dissipated in the degradation of cosmic rays due to the expansion. If however we regard the cosmic ray pressure as relatively small and neglect p_m then from (3.38) and (3.29) we obtain Pirani's result

$$E_m \simeq 36/\kappa R_s^3 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.42)$$

Although it is not pointed out by Pirani there is also a discrepancy here between (3.42) and the rate of creation of matter given by McCrea and by Hoyle. Hoyle, neglecting the pressure altogether and identifying ρ_m with ρ, obtains what would be obtained from (3.38) on the same suppositions viz. $E_m = 9/\kappa R_s^3$. McCrea gets the same result by identifying ρ_m with ρ and calculating the rate of working of the negative stress p on the boundary of unit volume. That is McCrea works in terms of the equation for total energy (3.24) and calculates

$$\frac{1}{v}\frac{d}{dt}(\rho v) = -p\,\frac{1}{v}\frac{dv}{dt}$$

as the rate of creation of energy. This is evidently a different interpretation of created energy from ours, and represents in fact the adiabatic increase in energy which a fluid of negative stress gains on expansion, without energy transfer from outside sources. In this sense an ordinary gas of positive pressure would experience creation of energy on contraction. This does not however seem to be the correct use of the word 'creation' in the physical context of the problem.

In our theory the quantity dQ_m of material energy is supplied from an outside source - namely the field of radiative energy from which it is materialised or created. It is accounted for by the increase in material energy actually witnessed in a volume v, plus the amount used to make good the loss of material energy which would arise, due to the material pressure, even in an adiabatic expansion. The total energy, material plus radiative, supplied from 'outside' sources must always be zero in general relativity i.e. $dQ_m + dQ_r = 0$ (eqn.(3.25)), in agreement with the principle of conservation of energy. That is, the expansions and contractions of a homogeneous fluid are always adiabatic in respect of total energy.

We now turn to the important question of how the rate of creation of matter in the substratum is affected by the proximity of a mass concentration or rarefaction, assuming that dynamic equilibrium between matter and radiation maintains the steady state density and pressure. For this purpose we shall employ the stationary exterior solution of metric (1.8) obtained in Chapter I.

The generally covariant equation of conservation of energy in a perfect fluid is

$$(\rho\lambda^{\nu})_{;\nu} + p\lambda^{\nu}_{;\nu} = 0 \qquad \text{...........(3.43)}$$

where ρ and p are the proper density and pressure respectively, λ^{μ} is the unit tangent to the world line of its motion, and a semi-colon indicates covariant differentiation. It is easily verified that in a local geodesic system of coordinates this equation reduces to that already expressed by (2.31). According to (3.22) we may write this equation in the form

$$(\rho_m\lambda^{\nu})_{;\nu} + p_m\lambda^{\nu}_{;\nu} = -\left\{(\rho_r\lambda^{\nu})_{;\nu} + p_r\lambda^{\nu}_{;\nu}\right\} \qquad \text{. ....(3.44)}$$

The left hand side of this equation is the rate of creation of matter in unit proper volume, referred to an arbitrary coordinate system, which we shall call E_m. The right hand side is $-E_r$, or the rate of decrease of radiation energy in unit proper volume. For the steady state undisturbed by local masses these expressions reduce to those given in (3.38) and (3.39).

For the model of metric (1.8) we have shown in Chapter I, Section (ii) that ρ and p have the constant values ρ_s and p_s at all points outside the concentration. Accordingly, because of dynamic equilibrium ρ_m and p_m can be assumed to have the constant values obtaining in the undisturbed steady state. Hence by (3.44)

$$\begin{aligned} E_m &= (\rho_m + p_m)\lambda^{\nu}_{;\nu} \\ &= (\rho_m + p_m)\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^{\nu}}\left(\sqrt{-g}\,\lambda^{\nu}\right) \qquad \text{....(3.45)} \end{aligned}$$

Hence for the metric (1.8)

$$E_m = (\rho_m + p_m)\frac{1}{\ell^2}\frac{d}{d\ell}\left(\ell^2 \frac{d\ell}{ds}\right) \qquad \ldots.(3.46)$$

The velocity $d\ell/ds$ of the substratum is given by (1.21), so that on simplification

$$E_m = \frac{3(\rho_m + p_m)}{R_s}\left\{\frac{1 + \frac{2}{3}\frac{(\alpha^2-1)R_s^2}{\ell^2} + \frac{mR_s^2}{\ell^3}}{\sqrt{1 + \frac{(\alpha^2-1)R_s^2}{\ell^2} + \frac{2mR_s^2}{\ell^3}}}\right\} \qquad \ldots(3.47)$$

We consider first the case when the mass concentration m is positive and sufficiently large for a neutral surface $\ell = \ell_0$ to exist. For this case α is given by (1.24). Whence by (1.23)

$$E_m = \frac{3(\rho_m + p_m)}{R_s}\left\{\frac{1 - 2\frac{\ell_0^2}{\ell^2} + \frac{\ell_0^3}{\ell^3}}{\sqrt{1 - 3\frac{\ell_0^2}{\ell^2} + \frac{2\ell_0^3}{\ell^3}}}\right\}$$

$$= \frac{3(\rho_m + p_m)}{R_s}\left\{\frac{1 + \frac{\ell_0}{\ell} - \frac{\ell_0^2}{\ell^2}}{\sqrt{1 + \frac{2\ell_0}{\ell}}}\right\} \qquad \ldots\ldots\ldots(3.48)$$

We note that the factor outside the curly bracket is by (3.38) the rate of creation in the undisturbed steady state, corresponding to $m = 0$ and $\alpha = 1$.

The behaviour of the disturbance factor is easily investigated by putting $\ell_0/\ell = x$. We omit the details and pronounce as follows. Due to the effect of the mass concentration the rate of creation in the substratum decreases monotonically from the undisturbed steady state value at great distance to $1/\sqrt{3}$ of this value at the neutral surface, decreases further monotonically inside $\ell = \ell_0$ to zero at $\ell = \{2/(1+\sqrt{5})\}\ell_0$ $\simeq .62\,\ell_0$ after which annihilation sets in for smaller ℓ. The variation of the rate of creation is monotonic throughout.

We can obtain a general appreciation of these results from the following considerations. Outside the neutral surface the acceleration of the substratum is reduced at any point by the effect of the mass m, when compared with the undisturbed steady state (eqn. (1.22)). Thus neighbouring particles in the line of motion move apart with decreased velocity so that a reduced rate of creation is necessary to maintain constant ρ_m. Inside the neutral surface the acceleration may actually exceed the undisturbed value, so that neighbouring particles in the line of motion move apart with increased velocity sufficiently near m, but this is more than counteracted by the lateral contraction of the substratum as the particles converge towards the mass m.

For the cases when m is positive, but so small that the neutral surface does not exist, and when m is negative, corresponding to cases (b) and (c) of Chapter I, Section (ii), it would be reasonable to expect that α should still be given by (1.24) although there appears to be no clear argument establishing this. However the exact analysis of (3.47) is unnecessary in these cases. For in both cases cosmological repulsion prevails everywhere in the substratum (eqn.(1.22)), being decreased in case (b) and increased in case (c) relative to the undisturbed steady state. Accordingly, the velocity of separation of the particles in the line of motion, and with it the rate of creation, is decreased in case (b) and increased in case (c). In both cases however the rate of creation is positive, being least nearest the mass m in case (b) and greatest nearest the 'hole' in the substratum in case (c).

We conclude, therefore, that creation of matter in the steady state universe, treated according to general relativity, is by no means a statistically uniform random process independent of local conditions, but highly sensitive to local gravitational influence.

Although our analysis has been applied to the substratum outside a stationary disturbance in the steady state it is evident that the same qualitative results will apply within expanding temporary concentrations and rarefactions in the steady state universe. Thus if in any region ρ is slightly greater than ρ_s then we have a positive mass concentration m. Within and around this region the expansion will be slowed up relative to the undisturbed steady state, and the rate of creation necessary to maintain equilibrium between matter and radiation will decrease. Conversely if ρ is slightly less than ρ_s the rate of creation of matter will increase. We have seen in Section (v) (eqn.(3.36)) that in either case expansion has the effect of restoring the steady state density and pressure.

(vii) Formation of galaxies in the steady state universe

By the analysis in Chapter I of the exterior solution of metric (1.8), or (1.28), it was found that a mass concentration in the otherwise smoothed out steady state universe will provide an attraction to counter cosmic repulsion only if $m > 0$ (eqn.(1.22)). If m is sufficiently large then cosmic repulsion will be overcome, and the fluid substratum will converge upon the mass m inside a certain neutral surface $\ell = \ell_0$, given by (1.23). The assumption that the physical situation was stationary led to the result that the steady state density ρ_s and pressure p_s were maintained in the exterior. In Section (vi) of this Chapter we have seen that this in turn requires a reduced rate of creation of matter outside and inside $\ell = \ell_0$, and in fact annihilation of matter for $\ell < .62\, \ell_0$.

In the particular case of the mass m being a sphere of uniform density ρ, the minimum density for the neutral surface to exist in the substratum was found in Chapter II, Section (ii) (b) to be $\rho = 3\rho_s$, and in this case this surface coincided with the boundary of the mass m. In this limiting case the exterior fluid was at rest at the boundary, and the gravitational mass of the sphere was then actually zero, the density and pressure being those of the Einstein universe. For densities

$\rho > 3\rho_s$ the gravitational mass of the sphere becomes positive, while the region of contracting fluid inside the neutral surface increases in size as ρ increases.

As we have remarked, in the analysis of Chapters I, II we were concerned with stationary solutions only, corresponding to dynamic equilibrium between matter and radiation. However in the case when there is a very massive central condensation, such as a galaxy, the rate of contraction of the substratum might become so large in its neighbourhood (eqn.(1.21)) that there would not be time for equilibrium to be maintained between matter and radiation. In these circumstances the density of matter in the region $\ell < \cdot 62\,\ell_0$ would build up in the contracting fluid while the density of negative radiation would also intensify. Much of this radiation would escape before it had annihilated much matter. This would increase the self gravitation of the contracting fluid whose contraction would therefore accelerate. In such circumstances, sufficiently near the central mass the substratum, even if the central mass were removed, would be able to contract under its own influence, and so to form a new source of attraction to the substratum further afield.

A temporary perturbing influence of this magnitude might be provided by the passage of a galaxy into a region in which the steady state was being maintained, as has been suggested by D.W. Sciama (19). Sciama of course assumes that the intergalactic fluid is gaseous hydrogen, with a negligible amount of ordinary radiation present. We follow Sciama in supposing that the fluid of the substratum would contract in the wake of such a galaxy. Since, as we have seen, any negative radiation not absorbed in the contracting substratum would be lost to outer space, the mechanics of the condensation will be along the lines discussed by Sciama for a contracting gas.

In particular, since the parameter m in the case of the original perturbing galaxy will represent nearly the entire mass of the galaxy, it follows that if the process is to lead to the formation of another galaxy the amount of mass attracted into the wake must be of order m. Since the entire mass within the neutral surface will be drawn in, it is satisfactory to note that even if we include the negative radiation this is of order $\rho_s \ell_0^3$, which by (1.5) and (1.23) is indeed of order m. It is easy to see from (1.21) that this process of contraction into the wake from the neutral surface ($\alpha \simeq 1$) will take a time of order T, where $cT = R_s$ so that T is the reciprocal of the Hubble parameter. This time is of the required order for the maintenance of a steady state as regards galaxies.

In this time the 'hole' left inside the neutral surface, filled predominantly with negative radiation when the material contents are first abstracted, will be restocked with newly created matter. This is made possible by the fact that in general the peculiar velocities of the parent galaxy and its child will have carried them out of the original neutral surface. Consequently the negative radiation left in the hole will begin to expand more rapidly than in the steady state, since the gravitational mass is negative. This rapidly increases the density of radiation, and, as we have seen earlier in Section (v) for the case when $\rho + p < 0$, this leads to the restoration of the steady state by the equilibrium process.

Sciama has shown that in about 50% of the galaxial births the parent galaxy captures its child, so that clusters of galaxies are gradually built up. We shall be content with remarking that according to our theory this process can continue only so long as the total gravitational mass of the cluster remains positive. If the average mass density ρ falls below $3\rho_s$, corresponding to the Einstein state, then the cluster cannot capture any more galaxies and may even lose some of its own members.

The problem arises, when a galaxy has sufficiently condensed for stars to be formed, are we to regard a star as in sufficiently quasi-static state for equilibrium between matter and radiation to exist in its interior, as we have assumed when the density was of order ρ_s? In thermodynamics a physical system which receives no energy from outside sources is regarded as being in chemical and thermodynamic equilibrium when its entropy is a maximum. This criterion, for the ordinary processes of conversion of matter in one form of chemical arrangement into another, leads to a knowledge of the relative proportions of the different constituents of the system which are present in equilibrium. The assumption usually made is that the zero of entropy for matter occurs when it is in pure crystalline form at the absolute zero of temperature.

However, in reactions of the type that we are considering involving the complete annihilation or creation of matter in equilibrium with radiation, the above assumption regarding the zero of entropy would ignore the possibility of large sources of entropy within the nucleus itself, as has been pointed out previously by R.C. Tolman (20). Thus, while the classical analysis of equilibrium would show that the concentration of matter in the stars is far too high for equilibrium with radiation, it may be that further knowledge of the atomic nucleus would lead to a different view.

(viii) The nature of negative radiation

Mathematically the notions of negative radiation, and of negative energy, stress, and inertia in general, introduce no extra difficulties in general relativity, and in fact they appear as a natural extension of the theory. This is also true of quantum mechanics where the equations permit of the existence of negative states of energy. It would be highly desirable if in this way a link was found between the quantum and gravitational field theories.

With such an aim evidently in mind Pirani, in his paper already referred to, tentatively ascribed the name gravitinos to the quanta of negative radiation emitted simultaneously with the creation of matter, but did not go further than this. From the point of view of quantum theory he suggested that the radiation might possibly be identified with neutrinos, and pointed out that from the experimental point of view there would be no difference between the emission of a neutrino of positive energy and the absorption of one of negative energy.

We remark that it has been suggested by F. Reines and C. L. Cowan (21) that a large part of the energy of the universe may exist in the form of neutrinos, because of their association with nuclear decay and their great penetrability of matter without being reabsorbed. Reines and Cowan have calculated that at energies near the threshold of energy for absorption by protons the mean free path of neutrinos would be of the order of the radius of the universe. This threshold energy is a few MeV. However, if our creation theory is correct, the negative radiation emitted with the creation of a proton and anti-proton would be far in excess of the threshold energy for subsequent partial annihilation by absorption, so that the cross section for collision could be expected to be considerably increased. The dynamic equilibrium envisaged in our theory might well be feasible, especially as the theory leads to a density of negative radiation of the same order as that of matter.

Whether negative radiation can be identified with neutrinos or not, we shall for our part put forward the view in this final section that negative radiation must in fact be associated with the energy of the gravitational field due to matter; that contrary to what is generally believed the gravitational field energy and stress contribute to the energy tensor in general relativity - in the form of Maxwellian electromagnetic radiation of negative energy and stress. This would explain why investigations in the past of gravitational waves, based on their propagation in empty space ($T_{\mu\nu} = 0$), have not been particularly fruitful.

Whittaker's theorem for gravitational mass as defined by Gauss' theorem was quoted in Chapter II, Section (ii) (b) and yields

$$M_G = \iiint \left(T^4_4 - T^1_1 - T^2_2 - T^3_3 \right) \sqrt{-g}\, dx\, dy\, dz \qquad \ldots.(3.49)$$

integrated over the spatial region concerned. The theorem is based on the gravitational force being the geodesic acceleration of free particles relative to observers fixed in the reference frame. The result is not in any way restricted to matter as such, and in fact (3.49) involves integration over all parts of the region where $T_{\mu\nu} \neq 0$.

Whittaker's theorem was proved only for a specialised type of time-orthogonal metric. It has been shown by G. L. Clark (<u>22</u>) that a generalisation of Whittaker's theorem by H. S. Ruse (<u>23</u>), applicable to an arbitrary metric, involves, in general, terms additional to those in Whittaker's result. Using Ruse's theorem Clark showed that the gravitational mass of an isolated system of freely moving particles, in an approximation to order m^2 where m is the mass of a typical particle (gravitational units), was equal to the total energy of the system, viz.

$$M_G = M_0 + K + \Omega \qquad \ldots\ldots\ldots\ldots\ldots(3.50)$$

In (3.50) M_0 is the sum of the invariant rest masses of the particles, K is the classical kinetic energy and Ω is the classical gravitational energy of the system. The latter contributes negatively as field energy arising from the disposition of the particles in space. The kinetic energy may also be regarded as field energy arising from the motion of the particles in space. It is noteworthy that Clark showed that the terms leading to this result included non zero contributions from 'empty' space.

For a static, spherical, and continuous distribution of cohesive fluid Whittaker's form of the theorem is applicable. For this case Clark shows (<u>24</u>) that $M_G = M_0$, to order M_0^2 as before. The reason why in this case $M_G \neq M_0 + K + \Omega$ must be put down to the fact that non gravitational stress and energy inside the mass, necessary to keep it in equilibrium, cancel out the gravitational stress and energy contributions when the integration is performed. Thus looked at from the point of view of (3.49) it seems to be indicated that the T^{ν}_{μ} must include contributions from the gravitational field.

It might be argued that gravitational field energy enters via the factor $\sqrt{-g}$ only. But it should be noted that even in a locally comoving geodesic system of coordinates there is a gravitational mass density σ , which for an isotropic medium is, according to Whittaker's result,

$$\sigma = \rho + 3p \qquad \text{..................(3.51)}$$

where ρ and p are the proper density and pressure. In fact, with a more satisfactory definition of gravitational force in terms of the relative acceleration of particles on adjacent geodesics, J. L. Synge (25) obtains precisely (3.51) for the local gravitational mass density relative to the world line of matter at the point in question.

In the form (3.51) Synge's result applies to the general cosmological models of metric (1.1). For that metric we see from equations (1.2), (1.3) that

$$K\sigma = -6\ddot{R}/R \qquad \text{..................(3.52)}$$

(K is now of course the gravitational constant 8π)

This equation must determine the actual motion of the substratum since p is uniform and hence the world lines are geodesics. Accordingly the expansion accelerates if $\sigma < 0$. This is the case in the steady state in which $\rho + p = 0$, $\rho > 0$. It seems natural to conclude that the cosmic counterpart of field energy to that in the isolated system considered by Clark is contained in ρ and p.

We have used the notion of gravitational mass to establish the conclusion of the last paragraph. We now turn to the consideration of invariant mass which is of greater theoretical importance. The invariant mass density is $T = T^{\nu}_{\nu}$ which on integration over a spatial region leads to the invariantly defined mass

$$M_0 = \iiint T \sqrt{-g}\, \frac{dt}{ds}\, dx\, dy\, dz \qquad \ldots\ldots\ldots\ldots\ldots(3.53)$$

It is this quantity that has to be regarded as the <u>rest</u> mass of a particle held together under stress - such as an electron, since the integral leading to M_G in (3.49) is not invariantly defined. The two quantities are however indistinguishable to order M_0^2, as Clark has shown.

In a comoving geodesic coordinate system in a perfect fluid the invariant mass density will be

$$T = \rho - 3p \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3.54)$$

For all kinds of radiation this quantity vanishes. Accordingly, T gives the invariant density of energy in the form of matter in any region. It provides the measure in that region of the non-Maxwellian energy in the form of fundamental particles, or that part of their energy which is non-Maxwellian. For a purely Maxwellian field, whether isotropic or not, T must vanish, as is well known.

We note also that T is the quantity that accounts entirely for the invariant Gaussian curvature of space-time, by virtue of the relation derived from the field equations

$$G = \kappa T \qquad \text{..........(3.55)}$$

Radiation makes no contribution to this curvature. The creation of matter in a region of space must therefore lead inevitably to the propagation of a variation of gravitational curvature G within and beyond this region.

Since we have seen that the creation of matter in the steady state is accompanied inevitably by the emission of negative radiation, according to the relation $d\rho_m = -d\rho_r$ (eqn. (3.34)), it follows that the propagation of a change in the Gaussian curvature G is simultaneous with the propagation of radiation of negative energy. It is natural to conclude that the propagation of gravitational waves occurs with the speed of light in the form of radiation, which is negative if matter is created and positive if matter is annihilated (ordinary radiation). The energy carried by the waves is the negative of the material energy created, in the form of matter and its associated kinetic energy, or it is material energy destroyed. In the case of creation the momentum of the radiation is opposite in direction to the sense of propagation, while in annihilation the momentum and direction of propagation are in the same sense.

The mechanism of the steady state may therefore be looked upon as follows. The geodesic expansion of the fluid leads to a decrease in the density of material energy, including the kinetic energy of matter in the form of cosmic rays, in a given spatial region. There is a corresponding equal increase in 'gravitational potential energy' in the form of negative isotropic radiation whose density and stress contribute to the energy tensor. This leads to a transfer of energy from a state of radiation to one of matter, in order to maintain a dynamic equilibrium which exists between them. Concentrations of matter, or rarefactions in the steady state lead automatically to adjustment of the energy density of the gravitational medium.

We remark that the 'hole' theory of P. Dirac (26) seems to find some expression in our theory. For our theory implies that in the steady state a 'sea' of positive energy is in dynamic equilibrium with a sea of negative energy. There is, in order to maintain this equilibrium as the universe expands, a steady passage of energy from the negative sea to the positive sea i.e. a hole of positive energy appears in the negative sea. This is equivalent to saying that the expansion increases the energy of the negative sea and decreases that of the positive sea and the transfer of energy restores the status quo. Furthermore the introduction of states of energy in a negative sea means that the local density of inertia can become negative with a reversal of the effects of motivating forces; it also means the possible existence of particles of negative mass.

REFERENCES

(1) W. M. McCrea. Proc. Roy. Soc. A 206 562 (1951)

(2) H. Bondi and T. Gold. Mon. Not. R.A.S. 108 252 (1948).

(3) F. Hoyle. Mon. Not. R.A.S. 108 372 (1948).

(4) F. Pirani. Proc. Roy. Soc. A 228 455 (1955)

(5)(J. Lemaitre. J. Math. and Phys. (M.I.T.) 4 188 (1925)
(H. P. Robertson. Phil. Mag. 5 835 (1928).

(6) R. C. Tolman. Relativity, Thermodynamics, and Cosmology; Oxford: Clarendon Press, 1934, p. 244.

(7) R. O. Kapp. (i) Nature 165 68 (1950)
(ii) Nature 165 687 (1950)

(8) G. C. McVittie. Mon. Not. R.A.S. (93) p. 325 (1933)

(9) R. C. Tolman. Op. Cit. (6) p. 250.

(10) G. D. Birkhoff. Relativity and Modern Physics; Harvard: University Press, 1923, p. 253

(11) A. Einstein and G. Straus (i) Rev. Mod. Phys. 17 120 (1945)
(ii) Rev. Mod. Phys. 18 148 (1946)

(12) G. C. McVittie. Reported in loc. cit. (11) (ii)

(13) F. Pirani. Proc. Camb. Phil. Soc. 50 637 (1954)

(14) R. C. Tolman. Op. cit (6) p. 243

(15) K. Schwarzschild. Berlin Sitzungsberichte p. 424 (1916).

(16) S. O'Brien and J. L. Synge. Comm. Dublin Inst. Adv. Studies A No. 9 1952.

(17) E. T. Whittaker. Proc. Roy. Soc. A.149 384 (1935)

(18) J. L. Synge. Trans. Roy. Soc. Canada 28 127 (1934)

(19) D. W. Sciama. Mon. Not. R.A.S. 115 1 (1955)

(20) R. C. Tolman. Op. cit. (6) p. 150.

(21) F. Reines and C. L. Cowan. Nature 178 446 (1956).

(22) G. L. Clark. Proc. Roy. Soc. Edin. A 62 412 (1946)

(23) H. S. Ruse. Proc. Edin. Math. Soc. (2) 4 144 (1935)

(24) G. L. Clark. Proc. Roy. Soc. Edin. A 62 424 (1946)

(25) J. L. Synge. Proc. Edin. Math. Soc. (2) 5 93 (1937)

(26) P. A.M. Dirac. Proc. Camb. Phil. Soc. 31 150 (1934)

Reprinted from the MONTHLY NOTICES OF THE ROYAL ASTRONOMICAL SOCIETY,
Vol. 117, *No.* 2, 1957

GENERAL RELATIVITY AND MACH'S PRINCIPLE

W. Davidson

(Communicated by W. H. McCrea)

(Received 1957 January 25)

Summary

Sciama, among others, has taken the view that general relativity has failed to account satisfactorily for the inertial properties of matter. This paper shows that general relativity is entirely consistent in principle with Sciama's ideas of inertia as an inductive effect predominantly of distant matter, and that therefore his remarks concerning general relativity are not justified. It is shown that general relativity provides a superior presentation of his idea of Mach's principle and appears to be the general tensor theory he was looking for.

Arguments are put forward to show that general relativity may fully incorporate Mach's principle contrary to Einstein's own belief. This paper emphasizes the fitness of the steady state theory as a cosmological solution which permits this possibility.

---

1. *Introduction*.—A tentative theory has been presented by Sciama (1), with Maxwell type equations, which is designed to provide a combination of Newton's laws of motion and of gravitation with the inertial frames determined by Mach's principle. In the introduction to his paper Sciama states that general relativity has failed to provide an adequate theory of inertia. He claims that his theory differs from general relativity principally in the following respects:

(i) it enables the amount of matter in the universe to be estimated from a knowledge of the gravitational constant;

(ii) the principle of equivalence is a consequence of his theory, not an initial axiom; and

(iii) it implies that gravitation must be attractive.

The chief characteristic of Sciama's theory is that "in the rest frame of any body the gravitational field of the universe as a whole cancels the gravitational field of local matter so that in this frame the body is 'free'. Thus in this theory inertial effects arise from the gravitational field of a moving universe." For this purpose Sciama employs a scalar potential Φ and a vector potential $\mathbf{A}$ to calculate gravitational effects, using Maxwell type field equations in flat space-time.

It is the purpose of this paper to show that general relativity is fully consistent with this interpretation of Mach's principle by Sciama, and to indicate that general relativity may fully incorporate Mach's principle.

2. *Free particle in general relativity*.—The motion of a free particle in general relativity, when the gravitational field is weak and when the reference frame is such that the spatial velocity of the particle is small compared with the velocity of light, can be described by a Maxwell-type pondermotive equation. This idea

is not new and has in fact, with limited application, been presented by Einstein (**2**). But since Einstein's derivation of the result appears to contain errors of detail we give our own derivation here, before investigating its significance for Mach's principle.

Let Latin letters indicate space coordinates running over indices 1, 2, 3 while Greek letters cover the space-time indices 1, 2, 3, 4. The world line of a free particle in general relativity is a geodesic in the field having equations, in a standard form,

$$\frac{d}{ds}\left(g_{\mu\alpha}\frac{dx^\alpha}{ds}\right)=\tfrac{1}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\mu}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} \tag{1}$$

where

$$ds^2=g_{44}(dx^4)^2+2g_{4p}dx^4dx^p+g_{pq}dx^pdx^q. \tag{2}$$

For $\mu=i$ equations (1) may be written

$$\left(\frac{ds}{dx^4}\right)^2\frac{d}{ds}\left(g_{ip}\frac{dx^p}{ds}\right)=\tfrac{1}{2}\frac{\partial g_{44}}{\partial x^i}-\left(\frac{ds}{dx^4}\right)^2\frac{d}{ds}\left(g_{i4}\frac{dx^4}{ds}\right)+\frac{\partial g_{p4}}{\partial x^i}\frac{dx^p}{dx^4}+\tfrac{1}{2}\frac{\partial g_{pq}}{\partial x^i}\frac{dx^p}{dx^4}\frac{dx^q}{dx^4}.$$

Write now $x^4\equiv t$, $dx^p/dt\equiv v^p$, and neglect squares and products of the spatial coordinate velocities v^p, getting

$$\frac{ds}{dt}\frac{d}{dt}\left(g_{ip}\frac{dt}{ds}v^p\right)=\tfrac{1}{2}\frac{\partial g_{44}}{\partial x^i}-\frac{ds}{dt}\frac{d}{dt}\left(g_{i4}\frac{dt}{ds}\right)+\frac{\partial g_{p4}}{\partial x^i}v^p. \tag{3}$$

We shall now represent the metric (2) as that of a weak field in the form

$$ds^2=(1+\gamma_{44})dt^2+2\gamma_{4p}dx^pdt-(1-\gamma_{11})(dx^1)^2-(1-\gamma_{22})(dx^2)^2-(1-\gamma_{33})(dx^3)^2 +\gamma_{pq}dx^pdx^q \qquad (p\neq q) \tag{4}$$

Here we take the velocity of light, c, as unity. We see that the $\gamma_{\mu\nu}$ are the deviations of the $g_{\mu\nu}$ from the Galilean values in the so-called inertial frames. They are the $\gamma_{\mu\nu}$ of Einstein's treatment of the problem except for sign due to his employment of imaginary x^4.

We make the assumption that the squares and products of the $\gamma_{\mu\nu}$ and those of their derivatives can be neglected. In solving the field equations to this approximation Einstein showed (**3**) that the $\gamma_{\mu\nu}$ were the solutions of the equations

$$\left\{\left(\frac{\partial}{\partial x^1}\right)^2+\left(\frac{\partial}{\partial x^2}\right)^2+\left(\frac{\partial}{\partial x^3}\right)^2-\left(\frac{\partial}{\partial t}\right)^2\right\}\gamma_{\mu\nu}=2\kappa(T_{\mu\nu}-\tfrac{1}{2}\delta_{\mu\nu}T) \tag{5}$$

provided

$$\frac{\partial^2}{\partial x^\nu\partial x^\alpha}(\gamma^\alpha_\mu-\tfrac{1}{2}\delta^\alpha_\mu\gamma^\beta_\beta)+\frac{\partial^2}{\partial x^\mu\partial x^\alpha}(\gamma^\alpha_\nu-\tfrac{1}{2}\delta^\alpha_\nu\gamma^\beta_\beta)=0 \tag{6}$$

to the order of the approximation. Here $\kappa=8\pi G/c^2$ where G is the Newtonian constant of gravitation, and $\gamma^\alpha_\mu=\delta^{\alpha\beta}\gamma_{\mu\beta}$ where $\delta^{\alpha\beta}$ are the Galilean values of the $g^{\alpha\beta}$. Assuming the contribution of stress to the energy momentum tensor to be vanishingly small compared with the density and momentum components for the case considered by Einstein, equation (5) yields Einstein's solution:

$$\left.\begin{array}{ll}\gamma_{11}=\gamma_{22}=\gamma_{33}=\gamma_{44}=-\dfrac{\kappa}{4\pi}\displaystyle\int\frac{[\rho]}{r}dV, & \gamma_{4p}=\dfrac{\kappa}{2\pi}\displaystyle\int\frac{[\rho u^p]}{r}dV \\ \gamma_{pq}=0, & p\neq q\end{array}\right\}. \tag{7}$$

and

In this solution ρ is the mass density, u^p the space velocity, of the element of mass in volume dV at distance r from the point where the $\gamma_{\mu\nu}$ are evaluated, all quantities being measured by observers at rest in the reference frame. Square brackets indicate retarded values corresponding to the propagation of the field with the unit velocity.

The integrals, supposed convergent, are over all matter producing the field. Such a solution of the wave equation of Lorentz (equation (5)) is well known to be valid only if the quantities solved for (the $\gamma_{\mu\nu}$) tend to zero in a suitable way. Thus it is clear that the case considered by Einstein involves mass concentrations only in the neighbourhood of the space origin and a field metric which is Galilean at "infinity". The integrals in (7) evaluated over such mass concentrations are therefore evidently convergent. The solution has to be consistent with conditions (6) which will be satisfied if the expressions

$$\frac{\partial}{\partial x^\alpha}(\gamma^\alpha_\mu - \tfrac{1}{2}\delta^\alpha_\mu \gamma^\beta_\beta)$$

vanish, for all μ, to the first order in the $\gamma_{\mu\nu}$. Using (7) and the fundamental equations $T,^{\mu\nu}_{\nu}=0$, it is easily seen that for integrals over a finite region of mass these expressions are indeed second order quantities.

To this approximation therefore, retaining only first order terms, we can reduce equation (3) to

$$\frac{d}{dt}\left(\frac{g_{ii}v^i}{\sqrt{g_{44}}}\right) = \tfrac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{d}{dt}(g_{i4}) + \frac{\partial g_{p4}}{\partial x^i}v^p,$$

on using (2) to find the appropriate approximation for ds/dt in each term. There is of course no summation over i on the left. Rearranging we can write:

$$\frac{d}{dt}\left(-\frac{g_{ii}v^i}{\sqrt{g_{44}}}\right) = -\tfrac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{\partial}{\partial t}(-g_{4i}) + \left\{\frac{\partial}{\partial x^p}(g_{4i}) - \frac{\partial}{\partial x^i}(g_{4p})\right\}v^p \tag{8}$$

so that

$$\frac{d}{dt}\{(1-\gamma_{ii}-\tfrac{1}{2}\gamma_{44})v^i\} = -\tfrac{1}{2}\frac{\partial\gamma_{44}}{\partial x^i} - \frac{\partial}{\partial t}(-\gamma_{4i}) + \left\{\frac{\partial}{\partial x^p}(\gamma_{4i}) - \frac{\partial}{\partial x^i}(\gamma_{4p})\right\}v^p. \tag{9}$$

Write now

$$\phi = -G\int\frac{[\rho]dV}{r},\quad A^p = -4G\int\frac{[\rho u^p]}{r}dV \tag{10}$$

so that by (7),

$$\gamma_{11}=\gamma_{22}=\gamma_{33}=\gamma_{44}=2\phi,\ \gamma_{4p}=-A^p. \tag{11}$$

Equation (9) can therefore be written in vector form covering $i=1, 2, 3$

$$\frac{d}{dt}\{(1-3\phi)\mathbf{v}\} = -\operatorname{grad}\phi - \frac{\partial\mathbf{A}}{\partial t} + \mathbf{v}\wedge\operatorname{curl}\mathbf{A}. \tag{12}$$

Equation (12) is the required Maxwell-type pondermotive equation of the field. The assumptions made during its derivation are:

(i) The particle velocity $\mathbf{v}$ in the reference frame is assumed small such that v^2/c^2 is negligible compared with v/c.

(ii) The deviations of the $g_{\mu\nu}$ from the Galilean values are small such that their squares and products and those of their derivatives can be neglected.

(iii) The deviations $\gamma_{\mu\nu}$ vanish at "infinity" so that the quantities $\mathbf{A}$, ϕ are defined in terms of convergent integrals. If *in addition* we now further assume that

(iv) The source velocities of the field are also small in the reference frame so that the same remark as in (i) applies for them, then equation (12) reduces to

$$\frac{d\mathbf{v}}{dt} = -\operatorname{grad}\phi - \frac{\partial\mathbf{A}}{\partial t} + \mathbf{v}\wedge\operatorname{curl}\mathbf{A}, \tag{13}$$

The equation obtained by Einstein was (our notation)

$$\frac{d}{dt}[(1-\phi)\mathbf{v}] = -\operatorname{grad}\phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \wedge \operatorname{curl}\mathbf{A}.$$

Since he assumed condition (iv) as well as (i), (ii), (iii), his result is incorrect to the order he was considering, and misleading. In obtaining this result he put

$$[p4, i] = \tfrac{1}{2}\left(\frac{\partial g_{4i}}{\partial x^p} - \frac{\partial g_{4p}}{\partial x^i}\right)$$

thereby neglecting the term $\partial g_{ip}/\partial x^4$ which, when $p=i$, contributes to our result in equation (9) as the term $d(-\gamma_{ii})/dt$ in the coefficient of v^i in the left hand side. The neglect of this term is of course consistent with condition (iv), but on the other hand the retention of the term $d(-\phi)/dt$ in the coefficient of $\mathbf{v}$, arising in our approximation from the term $d(-\frac{1}{2}\gamma_{44})/dt$ in the coefficient, is not consistent with Einstein's assumptions.

3. *Interpretation of the pondermotive equation.*—As Einstein pointed out, equation (12) indicates that general relativity goes far towards incorporating Mach's principle. It may be compared with the Newtonian equation, viz.,

$$\frac{d\mathbf{v}}{dt} = -\operatorname{grad}\phi.$$

The additional terms are small in the quasi-Galilean frame considered by Einstein, and, as he said, beyond physical measurement. Nevertheless they show in the sense of Mach's principle how concentrated matter affects the inertial mass of a freely moving particle; and the acceleration of its locally inertial rest frame relative to the given frame, in the following respects:

(i) The inertial mass is apparently proportional to $1-3\phi$.

(ii) The locally inertial rest frame of the particle is accelerated by means of:

(*a*) gravitational attraction towards the local mass concentrations indicated by the term $-\operatorname{grad}\phi$;

(*b*) an inductive effect of local accelerating matter in the same sense as the acceleration, indicated by the term $-\partial\mathbf{A}/\partial t$;

(*c*) an inductive effect of matter which is rotating relative to the compass of inertia (to use Gödel's phrase) at "infinity", in the sense of the rotation, as indicated by the term $\mathbf{v} \wedge \operatorname{curl}\mathbf{A}$. This is of the same type as the "fictitious" Coriolis force familiar in Newtonian dynamics, when a reference frame is used which is rotating relative to the compass of inertia. Centrifugal force also arises in this case as a fictitious gravitational force.

It is clear therefore that general relativity certainly incorporates in detailed manner the aspects of Mach's principle indicated above. For a satisfactory theory of Mach's principle, however, Einstein realised the necessity of showing how inertia depended on the entire cosmic distribution of matter. Because he assumed that the metric was Galilean at "infinity" and therefore excluded any contribution to $\mathbf{A}$ and ϕ other than that of the matter concentrated near the space origin, he was unable to examine the cosmic influence on inertia.

We shall here put forward an analysis to show that general relativity actually permits the same interpretation of inertia which has been presented by Sciama as the inductive effect of the whole universe.

4. *Inductive effect of the universe in general relativity.*—We shall investigate the extent to which we may generalize the circumstances when the motion of a free particle may be described by a Maxwell-type pondermotive equation. For this purpose we make the assumptions less restrictive than in the quasi-Galilean case as follows:

(i) The particle velocity $\mathbf{v}$ in the reference frame is assumed small such that v^2/c^2 is negligible compared with v/c.

(ii) The velocities of the sources of the field, in the region of space-time coordinates with which we shall be concerned, are also small of the same order so that the same remark applies.

(iii) The deviations of the $g_{\mu\nu}$ from the Galilean values are small in the above quoted range of space-time coordinates, such that their squares and products and those of their derivatives can be neglected. We do not however assume that these deviations vanish at "infinity", nor that they even remain small outside the specified range.

It is clear from equation (8) that the equation of motion of a free particle can in these circumstances be written

$$\frac{dv^i}{dt} = -\tfrac{1}{2}\frac{\partial g_{44}}{\partial x^i} - \frac{\partial}{\partial t}(-g_{4i}) + \left\{\frac{\partial}{\partial x^p}(g_{4i}) - \frac{\partial}{\partial x^i}(g_{4p})\right\} v^p \tag{14}$$

for $i=1, 2, 3$.
This equation is generally covariant, in the sense that, in all reference frames and regions of space-time which do not violate the assumptions above, it describes the space motion of the free particle in terms of the derivatives of the $g_{\mu\nu}$ involved. We now generalize the quantities $\mathbf{A}$, ϕ occurring in the quasi-Galilean analysis by defining

$$(\mathbf{A}, \Phi) \equiv (-g_{4i}, \tfrac{1}{2}g_{44}). \tag{15}$$

The three equations in (14) may then be written concisely

$$\frac{d\mathbf{v}}{dt} = -\operatorname{grad}\Phi - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \wedge \operatorname{curl}\mathbf{A}. \tag{16}$$

The vector notation implies the vector character of the terms for purely spatial transformations. For space-time transformations however the quantities $(\mathbf{A}, \Phi)$ do not transform as a 4-vector but as components of the tensor $g_{\mu\nu}$. This is because, unlike the corresponding electromagnetic pondermotive equation, the permitted transformations are not necessarily between inertial frames and therefore not in general linear.

It is to be noted that here we have not as in the quasi-Galilean case identified $\mathbf{A}$, Φ with the deviations of the $g_{\mu\nu}$ involved, from their Galilean values but, consistent with our endeavour to account for the whole of inertia according to Mach's principle, in terms of the total $g_{\mu\nu}$. The covariance of (16) is secured by the tensor character of the total $g_{\mu\nu}$ involved; the deviations do not transform as tensors for general transformations. Indeed according to the field equations it is the total $g_{\mu\nu}$ field that is related inseparably to the distribution of matter in the whole universe.

Bearing in mind therefore the physical interpretation of the quantities $\mathbf{A}$, ϕ in the quasi-Galilean case we should expect analogous interpretation of $\mathbf{A}$, Φ in (16) which would, if Mach's principle is to be satisfied, take account of the distribution and motion of matter in the whole universe, relative to the particular reference frame being used. It would be an immediate consequence of such an interpretation of the terms in (16) that the "fictitious" forces of Newtonian mechanics in accelerating or rotating reference frames would become directly attributable to the inductive effect of a moving universe.

In particular, in a reference frame in which a freely moving particle was permanently at rest, equation (16) would reduce to

$$-\operatorname{grad}\Phi-\frac{\partial\mathbf{A}}{\partial t}=0, \tag{17}$$

holding at the particle. This is the equation *postulated* by Sciama. To use Sciama's expression the "gravoelectric" field of the whole universe would be zero at the particle and it would be gravitationally "free" in its own rest frame.

For a reference frame at rest relative to the averaged motions of the rest of the matter in the universe (the "smoothed-out" universe) we should expect by Mach's principle that, in the neighbourhood of the space origin, the derivatives of $\mathbf{A}$, Φ on the right of (16) would vanish and therefore that the left hand side must vanish. The real existence of such frames which are locally inertial is the basis of Newtonian mechanics. This aspect of Mach's principle is built into general relativity theory since the field equations predict that such a reference frame will be Galilean near the space origin, because of the spherical symmetry about it. Thus in this neighbourhood the metric will approximate to

$$ds^2=dt^2-dx^2-dy^2-dz^2. \tag{18}$$

It is emphasized that in this paper we attach importance for Mach's principle to the total $g_{\mu\nu}$ involved in equation (16) and not just their derivatives. General reasons for this have already been given and further justification provided in Sections 5, 6. Accordingly it is important to obtain the total value of Φ. It follows from equation (18) that the static potential Φ_0, at the origin of such a frame, of the whole universe would be

$$\Phi_0=\tfrac{1}{2}g_{44}(0)=\tfrac{1}{2}\ (\text{or } \tfrac{1}{2}c^2 \text{ in general units}). \tag{19}$$

The dimensions of Φ and the significance we are trying to associate with it would require Φ_0 to be of order $-GM/R$ where M is the effective gravitational mass of the universe and R its effective radius. Sciama's approach is to *define* Φ_0 as $-\int_{r=0}^{r=R}\frac{\sigma dV}{r}$ where σ is the gravitational mass density, and he gets

$$G\Phi_0=-c^2.$$

Both results are numerically of the same order. The discrepancy in sign will occupy us later. Before investigating to what extent general relativity theory justifies this tentative physical interpretation of $\mathbf{A}$, Φ, we give some applications of our theory.

5. *Applications of the inductive theory in general relativity.*—(i) Sciama considers the case of a free particle in rectilinear motion in the gravitational field of a mass M which is at rest relative to the smoothed-out universe. If we choose a reference frame at rest relative to the smoothed out universe with this mass M at the space origin, we can neglect in that neighbourhood the deviations from the Galilean values of the $g_{\mu\nu}$ as far as they arise from the universe as a whole, and

include only the deviations due to the mass M. Thus to this approximation the metric will be

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 + \frac{2GM}{r}\right)(dx^2 + dy^2 + dz^2). \tag{20}$$

It is to be noted that, according to the ideas presented in this paper, the contribution to the $g_{\mu\nu}$ potentials from the universe as a whole is present in the Galilean terms of the $g_{\mu\nu}$.

Since the universe is at rest in this frame we have

while

$$\left.\begin{aligned} \mathbf{A} &= 0 \\ \Phi &= \tfrac{1}{2}\left(1 - \frac{2GM}{r}\right) \end{aligned}\right\}. \tag{21}$$

Suppose the particle is moving freely towards the mass M along the x axis. If its space coordinates are $(x_1, 0, 0)$ at coordinate time t then its coordinate speed is $dx_1/dt = -v$, where $v > 0$. Make now the transformation to a suitable rest frame for the particle, by means of the relations

$$x = X + x_1, y = Y, z = Z, t = T \tag{22}$$

yielding $dx = dX - vdT$, $dy = dY$, $dz = dZ$, $dt = dT$. We get therefore to sufficient order for the covariance of (16)

$$ds^2 = \left(1 - v^2 - \frac{2GM}{r}\right)dT^2 + 2vdXdT - \left(1 + \frac{2GM}{r}\right)(dX^2 + dY^2 + dZ^2). \tag{23}$$

Thus in the particle's rest frame

$$\left.\begin{aligned} \mathbf{A} &= (-v, 0, 0) \\ \Phi &= \tfrac{1}{2}\left(1 - v^2 - \frac{2GM}{r}\right) \end{aligned}\right\}. \tag{24}$$

Apply now equation (17) in the particle's rest frame, yielding

$$-\frac{\partial}{\partial X}\left\{\tfrac{1}{2}\left(1 - v^2 - \frac{2GM}{r}\right)\right\} - \frac{\partial}{\partial T}(-v) = 0 \tag{25}$$

leading to

$$-\frac{GM}{r^2} + \frac{dv}{dt} = 0 \tag{26}$$

on substituting the original coordinates. This is the Newtonian equation of motion of the particle and is also the equation which would follow from the general pondermotive equation (16), applied in the original frame, using (21).

On examining (24) and (25) we see that the origin of the inertial term

$$-\frac{\partial}{\partial T}(-v)$$

in (25) lies in the relative motion of the universe, yielding $\mathbf{A} = (-v, 0, 0)$ in the particle's rest frame, and thus creating an inductive field at the particle which balances the local gravitational attraction due to the mass M, thus connecting with Sciama's ideas.

We note also that the $\mathbf{A}$, Φ in (24) arise by transformation of the *whole* $g_{\mu\nu}$ and not just their deviations from the Galilean values, in accordance with our tentative interpretation of the Galilean values as the static potentials of the whole universe.

(ii) The other case considered by Sciama is that of a particle moving with uniform motion in a circle under the attraction of a mass M at the centre, this mass being again at rest relative to the smoothed-out universe.

Transform therefore from the metric (20) to a suitable rest frame for the particle according to the relations

$$\left.\begin{aligned} x&=X\cos\omega T-Y\sin\omega T\\ y&=Y\cos\omega T+X\sin\omega T\\ z&=Z\\ t&=T \end{aligned}\right\} \tag{27}$$

so that to sufficient order

$$\begin{aligned} ds^2=&(1-2GM/R-\omega^2R^2)\,dT^2-2\omega\,(-YdXdT+XdYdT)\\ &-(1+2GM/R)\,(dX^2+dY^2+dZ^2) \end{aligned} \tag{28}$$

with $R^2=X^2+Y^2$.

Thus in this frame

$$\left.\begin{aligned} \mathbf{A}&=(-\omega Y,\omega X,0)\\ \Phi&=\tfrac{1}{2}\left(1-\frac{2GM}{R}-\omega^2R^2\right). \end{aligned}\right\} \tag{29}$$

The equation (17) then yields

$$-\frac{GM}{R^2}+\omega^2R=0 \tag{30}$$

which is the Newtonian equation of motion, and also, putting $R=r$, what would be given by (16) in the original frame.

Connecting with Sciama's ideas we say that the gravitational attraction by M is balanced by the gravitational field induced by a rotating universe, whose rotational momentum is indicated by $\mathbf{A}$ in (29).

(iii) As a final example we shall show how, by means of the covariance of (16), the Newtonian "fictitious" forces may be attributed to the inductive effect of a moving universe in the most general Newtonian motion of the reference frame relative to a locally inertial frame.

Consider a free particle at rest in a reference frame which is locally inertial, so that the metric is approximately as given by (18) in that region. Let $\mathbf{r}$ be the position vector of the particle in that frame. Then by (15), (16) we have

$$\mathbf{r}=\text{constant}. \tag{31}$$

Transform to a second frame whose space origin has variable velocity $\mathbf{V}$ and which has variable spin $\boldsymbol{\omega}$ relative to the first frame. If the position vector of the particle in this frame is $\mathbf{R}$, then a well-known kinematic result of Newtonian motion gives

$$\dot{\mathbf{r}}=\mathbf{V}+\dot{\mathbf{R}}+\boldsymbol{\omega}\wedge\mathbf{R} \tag{32}$$

$$\ddot{\mathbf{r}}=\dot{\mathbf{V}}+\boldsymbol{\omega}\wedge\mathbf{V}+2\boldsymbol{\omega}\wedge\dot{\mathbf{R}}+\dot{\boldsymbol{\omega}}\wedge\mathbf{R}+\boldsymbol{\omega}\wedge(\boldsymbol{\omega}\wedge\mathbf{R})+\ddot{\mathbf{R}} \tag{33}$$

differentiation being with respect to the common Newtonian time of either frame. Thus for the particle in the second frame

$$\ddot{\mathbf{R}}=-[\dot{\mathbf{V}}+\boldsymbol{\omega}\wedge\mathbf{V}+2\boldsymbol{\omega}\wedge\dot{\mathbf{R}}+\dot{\boldsymbol{\omega}}\wedge\mathbf{R}+\boldsymbol{\omega}\wedge(\boldsymbol{\omega}\wedge\mathbf{R})]. \tag{34}$$

The transformation connecting the two frames is, by (32), in differential form

$$\left.\begin{aligned} d\mathbf{r} &= (\mathbf{V} + \boldsymbol{\omega} \wedge \mathbf{R})dT + d\mathbf{R} \\ dt &= dT \end{aligned}\right\}. \tag{35}$$

Hence $ds^2 = dt^2 - d\mathbf{r}^2$

$$= [1 - \mathbf{V}^2 - 2\mathbf{V}\,.\,(\boldsymbol{\omega} \wedge \mathbf{R}) - (\boldsymbol{\omega} \wedge \mathbf{R})^2]dT^2 - 2(\mathbf{V} + \boldsymbol{\omega} \wedge \mathbf{R})\,.\,d\mathbf{R}dT - d\mathbf{R}^2 \tag{36}$$

so that in the second frame

$$\left.\begin{aligned} \mathbf{A} &= \mathbf{V} + \boldsymbol{\omega} \wedge \mathbf{R} \\ \Phi &= \tfrac{1}{2}[1 - \mathbf{V}^2 - 2\mathbf{V}\,.\,(\boldsymbol{\omega} \wedge \mathbf{R}) - (\boldsymbol{\omega} \wedge \mathbf{R})^2] \end{aligned}\right\}. \tag{37}$$

Now

$$\operatorname{grad} \Phi = \boldsymbol{\omega} \wedge \mathbf{V} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{R})$$

$$\frac{\partial \mathbf{A}}{\partial T} = \frac{\partial \mathbf{V}}{\partial T} + \frac{\partial \boldsymbol{\omega}}{\partial T} \wedge \mathbf{R}$$

$$= \dot{\mathbf{V}} + \dot{\boldsymbol{\omega}} \wedge \mathbf{R}$$

while curl $\mathbf{A} = 2\boldsymbol{\omega}$.

Hence by (16) $\ddot{\mathbf{R}} = -[\dot{\mathbf{V}} + \boldsymbol{\omega} \wedge \mathbf{V} + 2\boldsymbol{\omega} \wedge \dot{\mathbf{R}} + \dot{\boldsymbol{\omega}} \wedge \mathbf{R} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{R})]$

giving complete agreement with (34).

Thus our theory gives an exact treatment of the fictitious forces as the inductive effect of a moving universe.

6. *Physical interpretation of* **A**, Φ *in general relativity*.—The analysis in this section is intended to be of a tentative nature, since complete rigour cannot be claimed for it.

In Section 4, equation (19), we obtained the result

$$\Phi_0 = \tfrac{1}{2}c^2 = \tfrac{1}{2}g_{44}(0)$$

for that value of the gravitational potential Φ of the whole universe which enters into the pondermotive equation (16), when evaluated at the space origin of a reference frame locally inertial there. In order to interpret this result in terms of Mach's principle we recall the expressions for ϕ in the quasi-Galilean case given by (7) and (11). Since the field equations are relations for the whole $g_{\mu\nu}$ in terms of the matter in the whole universe, we make the tentative inference that in some way the Galilean terms themselves are related to world gravitation, so that inertia would arise in accordance with Mach's principle. To what extent does general relativity provide justification of this inference?

In all cosmological models of general relativity in which the average inertial density ρ does not vanish there is an effective radius R of the model which is the distance, measured in a suitable way, to the horizon of the model where the velocity of the matter relative to the space origin equals the velocity of light. For an observer at the origin matter which goes beyond this distance virtually ceases to exist because of the Doppler effect on its light and presumably on its gravitation. This is the case whether the model be of the homogeneous rotating type (Gödel's models) or the isotropic expanding or contracting types. We shall discuss the latter as an example. These have the general metric

$$ds^2 = c^2dt^2 - e^{g(t)}\frac{(dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2)}{(1 + r^2/4R_0^2)^2} \tag{38}$$

where R_0^2 may be positive, negative, or infinite, and the fundamental particles have constant r, θ, ϕ.

The distance to the particle at (r, θ, ϕ), measured in the simultaneity of the fundamental observers at cosmological time t, from $r=0$ is

$$l=e^{\frac{1}{2}g(t)}\int_0^r \frac{dr}{1+r^2/4R_0^2}.$$

Its radial velocity is therefore

$$\dot{l}=\tfrac{1}{2}\dot{g}l \tag{39}$$

so that

$$|\dot{l}|=c \text{ when } l=\frac{2c}{|\dot{g}|}.$$

Thus

$$R=\pm\frac{2c}{\dot{g}} \quad \text{according as } \dot{g}\gtrless 0 \tag{40}$$

and

$$\dot{l}=\pm\frac{cl}{R}. \tag{41}$$

The importance for Mach's principle is that R is related to σ, the cosmological density of gravitational mass. In general relativity theory, gravitational mass density is defined so as to lead to Gauss' flux theorem for small regions of space (see, for example, Synge (**4**)), and for the isotropic cosmological models $\sigma=\rho+3p/c^2$ where p is the pressure. The gravitational "force" on unit mass due to the field is in this case the proper acceleration relative to the space origin. With these definitions McCrea (**5**) has shown that the equations of general relativity for the isotropic models are consistent with the variation of the gravitational force according to the Newtonian inverse square law, using proper radial distance but Euclidean geometry, for spatial regions of any size. The field equations applied to the metric (38) give, with $\Lambda=0$, in general units,

$$\left.\begin{aligned} \frac{8\pi G}{c^2}p &= -\frac{c^2}{R_0^2}e^{-g(t)}-\ddot{g}-\frac{3}{4}\dot{g}^2 \\ 8\pi G\rho &= \frac{3c^2}{R_0^2}e^{-g(t)}+\frac{3}{4}\dot{g}^2 \end{aligned}\right\} \tag{42}$$

so that

$$8\pi G\sigma=-3(\ddot{g}+\tfrac{1}{2}\dot{g}^2).$$

Now for $\dot{g}>0$, $\quad \dot{g}=\dfrac{2c}{R},\ \ddot{g}=-2c\dot{R}/R^2$ by (40). Hence

$$8\pi G\sigma=-\frac{6c^2}{R^2}(1-\dot{R}/c)$$

$$\left.\begin{aligned} &\text{whence, if } \dot{g}>0, & G\sigma R^2 &= -\frac{3c^2}{4\pi}(1-\dot{R}/c) \\ &\text{and, if } \dot{g}<0, & G\sigma R^2 &= -\frac{3c^2}{4\pi}(1+\dot{R}/c) \end{aligned}\right\} \tag{43}$$

which are the required relations between σ and R, at time t. For $\dot{g}>0$ we see that $\sigma\gtrless 0$ according as $\dot{R}\gtrless c$. If $\dot{R}>c$ matter is entering the region bounded by the defined horizon; if $\dot{R}<c$ matter is passing beyond this horizon. For $\dot{g}<0$, $\sigma\gtrless 0$ according as $\dot{R}\lesseqgtr c$.

By (41) and (43) we get the Newtonian type equation

$$\ddot{l}=-\tfrac{4}{3}\pi G\sigma l \tag{44}$$

relating gravitational force and proper distance at time t.

Comparing equation (44) with the pondermotive equation (16) we see that, for a fundamental observer whose radial space coordinate is the proper distance l and whose time is the cosmological time t, $\mathbf{A}=0$ and $-\text{grad}\ \Phi = -\frac{4\pi}{3}\pi G\sigma\mathbf{l}$. Equation (44) however holds for all $l \leqslant R$ and not just in the neighbourhood of the origin. Consider therefore the gravitational "work" done by the field when a particle of unit mass is moved from its actual position at time t to the horizon and therefore beyond influence of the origin. This will be

$$\Phi_l = -\frac{4}{3}\pi G \int_l^R \sigma l dl. \tag{45}$$

This may be regarded as the analogue of the Newtonian potential at the distance l, in the gravitational field as witnessed by an observer at the origin. Both σ and R will vary with l in this integral as the motion proceeds, according to (42), (43). However Φ_l may be evaluated as

$$\begin{aligned}\Phi_l = \int_l^R \ddot{l} dl &= \tfrac{1}{2}c^2 - \tfrac{1}{2}\dot{l}^2 \\ &= \frac{c^2}{2}(1 - l^2/R^2).\end{aligned} \tag{46}$$

This is the potential at $\mathbf{l}$ at time t.

Putting $l=0$ we get $$\Phi_0 = c^2/2 \tag{47}$$

which therefore provides a physical identification, in a natural way, of the potential Φ_0 arising in equation (19).

The result given by (47) for Φ_0 is got as the limit of Φ_l when l tends to zero irrespectively of the sign of σ or $\dot{g}$. For instance if $\sigma > 0$ and $\dot{g} > 0$ then $\dot{R} > c$, so that for the field to carry the particle to the horizon of the space origin would mean going backwards in time. The discrepancy in sign between our Φ_0 and Sciama's, referred to at the end of Section 4, arises because of Sciama's arbitrary definition of Φ for an expanding universe. His definition appears to ignore the above considerations and in particular to presuppose the identity of the cosmological gravitational mass density and the inertial density.

For cogent reasons which have been put forward elsewhere (**6**) a stationary cosmological solution is to be preferred. The only known stationary solution which does not contradict observational results (expansion, spatial isotropy) is the steady state theory proposed by Bondi and Gold (**6**). This has the metric

$$ds^2 = c^2dt^2 - e^{2ct/R}(dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2) \tag{48}$$

where R is a constant which is the effective radius for the model. The steady state model, unlike the general cosmological models of metric given by (38) for which equation (16) vanishes identically, allows a static metric to be used so that the motion of a particle, relative to the observer at the origin, is measured by the rate of change of the spatial coordinates. This is the De Sitter metric

$$ds^2 = c^2(1 - l^2/R^2)d\tau^2 - \frac{dl^2}{1 - l^2/R^2} - l^2d\theta^2 - l^2\sin^2\theta d\phi^2 \tag{49}$$

connected to (48) by a well known transformation. In this metric l is the distance from the origin, in the simultaneity of the fundamental observers, of our general analysis. The theory of Section 4 defines the Φ involved in (16) as $\frac{1}{2}g_{44}$ which in the case of the metric (49) gives

$$\Phi = \frac{c^2}{2}(1 - l^2/R^2) \tag{50}$$

agreeing with (46) and therefore having the physical interpretation associated with (46).

It is to be noted that the steady state forms a natural cosmological background to mass concentrations. For instance the *exact* solution of the field equations for an isolated mass m superimposed on the steady state is

$$ds^2 = c^2\left(1 - \frac{2mG}{c^2 l} - \frac{l^2}{R^2}\right)d\tau^2 - \frac{dl^2}{1 - \frac{2mG}{c^2 l} - \frac{l^2}{R^2}} - l^2 d\theta^2 - l^2 \sin^2\theta d\phi^2. \tag{51}$$

For this metric

$$\Phi = \frac{c^2}{2}(1 - l^2/R^2) - \frac{mG}{l} \tag{52}$$

to be interpreted physically as the work done by the field in removing unit mass from the point in question to the horizon of the model, regarding mG/R as negligible.

A Newtonian type integral for Φ in terms of the distribution of the mass does not follow simply in the case of the general models because of the stated dependence of σ and R on cosmological epoch. However, for the steady state, σ and R are constant and we may write for the potential of unit mass, at distance l from the mass σdV constantly in the volume element dV,

$$\begin{aligned} d\Phi &= -G\int_l^R \frac{\sigma dV}{l^2}\, dl \\ &= -G\sigma dV\left(\frac{1}{l} - \frac{1}{R}\right). \end{aligned} \tag{53}$$

Thus

$$\begin{aligned} \Phi_0 &= -G\sigma\int_0^R \left(\frac{1}{l} - \frac{1}{R}\right) 4\pi l^2 dl \\ &= -\frac{2\pi}{3} G\sigma R^2. \end{aligned}$$

Equation (43) gives for the steady state $G\sigma R^2 = -3c^2/4\pi$ so that

$$\Phi_0 = c^2/2$$

in agreement with (47) and justifying our physical interpretation of Φ_0 as the gravitational potential of all the matter in the universe apparent to an observer at the origin and having influence there.

The quantity **A** of our theory defined in equation (15) of Section 4 is zero for the cosmological metric (38). On making a transformation such as that of the first example in Section 5, equation (22), a non-zero **A** arises by transformation of the $g_{\mu\nu}$. If we accept the association of the Galilean values of the $g_{\mu\nu}$ with world gravitation, according to the tentative analysis presented above, the association of **A** with the relative momentum of the universe would also follow. While the Galilean g_{44}, viz. c^2, is associated with Φ_0 as $2\Phi_0$, the spatial Galilean $g_{\mu\nu}$ are associated with **A**. Thus in the first example in Section 5 to get **A** in the particle's rest frame we have to multiply the Galilean g_{11} in the original frame, viz. $-1 = g_{11}(0)$, by v. According to the subsequent application of the pondermotive equation $-g_{11}(0)$ is proportional to the inertial mass of the particle, the whole equation indicating equality of gravitational and inertial mass in the case of a particle. Since $-g_{11}(0) = g_{44}(0)/c^2 = 2\Phi_0/c^2$ we see that inertial mass can therefore be associated with the influence of the whole universe, in accordance with Mach's principle.

7. *Comparison with Sciama's theory.*—In this final section we shall remark briefly on the three principal differences claimed by Sciama between his theory and general relativity, enumerated (i), (ii), (iii) in the introduction to this paper.

(i) It is evident from the analysis in this paper that a knowledge of G, occurring in the integral (45) leading to (47), together with R given as cT where T is the reciprocal of the Hubble parameter, leads to an estimate of the amount of matter in the universe in general relativity as well as in Sciama's theory.

(ii) Sciama states that in general relativity the principle of equivalence predicts that one gravitating mass in an otherwise empty universe produces the same inertial effects as in his theory, and since there is no universe in this case to give rise to the inductive field "it is difficult to see why the principle of equivalence should be true". Such an argument however implies a solution of the field equations involving the use of coordinates for all points of space-time in a universe which, except for the isolated mass, is empty. Such coordinates are purely conceptual, defined without reference to matter and restoring to space an objective substance, independent of matter, which general relativity has sought to deny. The logical course for general relativity, according to the field equations, is to relate the Galilean $g_{\mu\nu}$ of special relativity to world gravitation in a full universe. That general relativity may be capable of doing so has been indicated in this paper, where, independently of the value of ρ as long as it does not vanish, $\frac{1}{2}\, g_{44}(0)$ has been identified as $\Phi_0 = c^2/2$, the potential of the universe. The case of the empty universe can only logically be approached as a limit where $\rho \to 0$ and $R \to \infty$ (equation (43)), so that inertia is always accounted for.

(iii) Sciama's determination of the sign of the field in his theory is of doubtful significance as it depends on his Φ as defined turning out to be negative. Reasons for questioning this arbitrary definition of Φ have been given in Section 6.

In general relativity the coefficient of $T^{\mu\nu}$ in the field equations is chosen so as to make gravitation attractive on the small scale (the pressure being then relatively negligible). However, on the cosmological scale this leads to gravitational mass being interpreted as negative if the expansion is accelerating (equation (44)). Thus there would appear to be no intrinsic importance to be attached to the sign of the field in any theory that allows for factors at present unknown, that is whether the gravitational effect of a lump of matter is more primary than that of a cosmological region containing "zero-point" stress.

I am grateful to Professor W. H. McCrea for opportunities for discussion of this paper and for his criticisms which have stimulated several improvements. Acknowledgments are due also to the authorities of the Battersea College of Technology for a relaxation of lecturing duties to perform this research.

Royal Holloway College (University of London),
Englefield Green,
Surrey:
1957 *January* 24.

References

(1) D. W. Sciama, *M.N.*, **113**, 34, 1953.
(2) A. Einstein, *The Meaning of Relativity* (Methuen) 5th Edition, p. 97, 1951.
(3) A. Einstein, *op. cit.*, p. 83.
(4) J. L. Synge, *Proc. Edin. Math. Soc.*, 2nd Series, **5**, 93, 1937.
(5) W. H. McCrea, *Proc. Roy. Soc.* A, **206**, 562, 1951.
(6) H. Bondi and T. Gold, *M.N.*, **108**, 252, 1948.