EARLY STAGES OF THE UNIVERSE

D. W. SCIAMA

Dept. of Astrophysics, Oxford University, Oxford, England

1. Introduction

When I was young theoretical cosmology was more a series of exercises in geometry than a branch of astrophysics. Moreover these exercises were based on little more than the Hubble law of red shifts and a general impression that on a large scale the Universe was roughly homogeneous and isotropic. Now all that has changed. While geometrical considerations retain their importance through the dominant role played by general relativity, cosmology has also become highly astrophysical. As a result we can now say a great deal about the early stages of the Universe. My aim is to give a brief account of our present understanding of this fascinating subject, at the same time highlighting the most important problems that remain.

Let us first recall the situation as it stood before these recent developments took place, say in the early nineteen fifties. The Hubble law then told us that the red shift z of a galaxy (defined by $\delta\lambda/\lambda_{rest}$) was proportional to the distance of the galaxy. This law had two immediate consequences:

(i) All galaxies would see the same linear law of red shifts relative to themselves and with the same Hubble constant. This strengthened the inference from the observed rough homogeneity of the large-scale distribution of galaxies, that our own Galaxy was in a *typical* place in the Universe and not a special one.

(ii) Taken together with the observed homogeneity and isotropy, and assuming that the red shift was a Doppler shift, it implied that a finite time ago the Universe had a high density (strictly speaking an infinite density, if the symmetries were taken as exact, and Einstein's cosmical constant ignored). The corresponding 'age' of the Universe depended upon the value of the Hubble constant and so upon the distance scale for galaxies, which was being substantially revised in the 1950's and is even now uncertain to at least a factor 2. This important astronomical problem lies outside our present subject; for our purposes it will suffice to use the round figure of 10^{10} yr for the Hubble constant. The time since the initial singularity depended also on the present mean density of matter in the Universe, through its gravitational influence on the deceleration of the expansion. Here the possibilities ranged from a low density universe, in which galaxies made the main contribution to the mean density ($\sim 10^{-30}$ gm cm^{-3}), to a high density universe, in which intergalactic gas (or faint stars, bricks, neutrinos, gravitational waves or black holes) contributed more substantially than galaxies. An upper limit on the possible density came from the requirement that the age of the Universe should not be less than the reasonably well-determined ages of its contents (for instance, the Earth, the Sun, the Milky Way). This upper limit was about 10^{-28} gm cm^{-3}. A low density universe would be destined to expand forever,

while a high density one (with $\varrho > 2 \times 10^{-29}$ gm cm^{-3}) would eventually re-contract into a singularity. (Some of these statements would need modifying if a cosmical constant were retained in Einstein's field equations of general relativity. The resulting new possibilities are discussed in e.g. Bondi (1960).)

An important speculation that had recently been made (Gamow, 1948, 1949) was that in the early stages of the Universe there might have been enough heat present to promote the build-up of heavier elements out of hydrogen by means of thermo-nuclear reactions, a proposal now conjured up in the phrase 'hot big bang'. According to the α-β-γ theory (Alpher et al., 1948) essentially all the elements heavier than hydrogen now observed were built up in this way. In this picture the early heat, present as an initial condition at $t = 0$, would have rapidly attained a black body spectrum by interacting with matter, and would thereafter retain such a spectrum to a good approximation even when, after considerable expansion of the Universe, the interaction with matter became much reduced (for zero interaction the spectrum would remain exactly that of a black body, the temperature dropping in proportion as the scale factor of the Universe expanded). The α-β-γ calculations suggested that the present temperature of the black body radiation field would be measured in tens of degrees absolute, a prediction which it was not then possible to verify (and later became forgotten).

On the theoretical side, the homogeneous and isotropic solutions of Einstein's field equations had been codified into the Robertson-Walker models (also called the Friedman models). This phase of the theory is fully discussed in Bondi (1960). One property of these models, which was first fully elucidated by Rindler (1956), was the occurrence of various kinds of *horizons*. The one which will be important for our later considerations is the *particle horizon*. In a cosmological model with such a horizon a fundamental observer at time t cannot see all the particles (galaxies) in the Universe, but only those within his particle horizon, the size of which increases with t. This means that there has not been time enough for radiation from the more distant galaxies to reach the fundamental observer by the time t. This may seem odd at first sight, and it does depend on the Universe having had a singular origin. Roughly speaking we may say that the initial rate of expansion of the Universe was infinite, so that some fundamental observers were thrown apart to distances such that their light time away exceeded the age of the system. (Rindler shows how this remark can be made rigorous).

So far we have been recalling the conventional picture of the early 1950's, based on an acceptance of general relativity and of the assumption that the matter and radiation content of the Universe obeyed familiar kinds of equation of state. There were also extant a number of unconventional ideas involving deviations from standard theory, such as having a time-varying gravitational constant (Milne, 1948; Dirac, 1938, 1973a, 1973b; Jordan, 1949, 1973; Hoyle, 1973; see also Brans and Dicke (1961) and Dicke (1962)). Of these the most influential and the most closely related to astro-physical questions was undoubtedly the steady state theory of Bondi and Gold (1948) and Hoyle (1948). This theory still has its adherents despite the hostile evidence which

has accumulated during the recent astrophysical phase of cosmology. Indeed one possible lecture I could have given would have involved presenting a critical account of all the modern data relevant to this question. However, I believe that this would not be the appropriate occasion for such an account. I prefer therefore to accept for this lecture what I think is generally agreed to be the most plausible outcome of such a discussion, and to assume that conventional theory is correct at least as far as saying that the Universe *did* have early stages different from its present stage.

This is the essential point because the steady state theory, as its name implies, took the radical view that on a large scale the Universe does not change with time at all. This view was reconciled with the observed expansion by the postulate that matter was being continually created at the appropriate rate, a rate far too small to be subject to direct observational check. The basic question of whether the Universe evolves with time or not is then best answered by seeking clear-cut evidence as to whether the universe *was* different in the past. The nearest approach to such evidence is provided by the 3 K microwave background, whose observational existence was hardly suspected before 1965, although, as we have seen, it was in fact predicted by Gamow in 1948. (The counts of radio sources and quasi-stellar objects also probably convince most astronomers that the Universe is evolving, but since they are subject to greater controversy than the microwave background, I shall not consider them here.)

The story of the great discovery by Penzias and Wilson (1965) of excess radiation at a wavelength of 3 centimeters, of its interpretation as a cosmological black body radiation field at 3 K by Dicke *et al.* (1965), and of the gradual strengthening of our conviction that the microwave background does have an accurately black body spectrum, is too well known to be worth repeating here. (Accounts can be found in Field (1969), Partridge (1969), Peebles (1971), Sciama (1971a, b), and Weinberg (1972)). Suffice it to say that recent claims to have discovered excess radiation with $T > 3$ K which were in conflict with other observations in the submillimetre range have now been withdrawn, but that we still eagerly await (satellite) observations in this decisively important region, which lies beyond the predicted peak at ~ 1 mm. For the purposes of this lecture I shall simply assume that the microwave background does have an accurately thermal spectrum, and shall follow the conventional view that this implies that the Universe was denser in the past than it is to-day. We are then led to a more general question, namely how far back and with what confidence, can we extrapolate into the past? What does observation tell us directly about the early stages of the Universe?

2. Extrapolating into the Past

2.1. THERMALISATION OF THE MICROWAVE BACKGROUND

We are making the assumptions that the microwave background has interacted sufficiently with matter at some stage in the development of the Universe to have been brought into thermal equilibrium and that thereafter this equilibrium has not been significantly disturbed. These assumptions impose a number of important constraints

on the thermal history of the Universe. To understand these constraints it is helpful to consider first the mean free path of a microwave photon as it propagates through the Universe. The most important process determining this mean free path is probably Thomson scattering by intergalactic (or pregalactic) free electrons. We denote by z_0 the red shift of an emission process which occurred one mean free path away. A straightforward calculation shows that this red shift z_0 ranges from about 7 in a high density universe consisting mainly of ionised intergalactic gas, to about 1000 in a low density universe. The corresponding age of the Universe when the photon was emitted ranges from about 10^9 yr to about 10^7 yr.

Now the immediate source of the observed microwave background is the free electrons located on the 'last-scattering surface' at red shift z_0. We may therefore regard ourselves as completely surrounded by such a surface, whose radiation temperature was $3(1+z_0)$ K at the moment of emission, that is, about 24 K in a high density universe and 3000 K in a low density universe. However, to achieve adequate thermalisation requires many scatterings (and also transitions in which photons are absorbed or emitted and not simply scattered, since if photon number were conserved an arbitrary radiation field could not reach thermal equilibrium with matter). Thus if these are the responsible processes thermalisation must have occurred at red shifts considerably greater than z_0 and at correspondingly further times into the past. In fact the most recent thermalisation epoch must have occurred at a red shift of about 4×10^5 in a high density universe, when the age of the system was about 30 years and its density nearly 10^{17} times the present density. In a low density universe the corresponding red shift was greater (up to 4×10^7), the associated density also greater (up to nearly 10^{23} times the present density) and the age of the Universe less (down to 5×10^5 s). Hence if thermalisation is due to the coupling between radiation and ionised material then the present observation that the microwave background has a black body spectrum gives us direct access to processes occurring when the Universe was at most 30 years old. (An alternative possibility is thermalisation by interaction with dust particles (Layzer and Hirely, 1973). In this case thermalisation would have occurred more recently.)

Once thermalisation ceases the black body character of the microwave background is preserved throughout the expansion (with temperature $T \propto 1+z$), unless further interaction with free electrons then distorts the spectrum. This distortion is a real possibility for two reasons. Firstly, after thermalisation ceases the free electrons tend to cool more rapidly than the radiation, and secondly there may be a later phase when the free electrons are heated to a much higher kinetic temperature as a result of interaction with cosmic rays and electromagnetic radiation emitted by galaxies, radio galaxies and quasi-stellar objects. Insofar as distortions from a black body spectrum are not observed we must impose the appropriate constraints on the possible thermal history of the Universe. This is a very interesting and important question, but since it concerns stages later than 30 years from the big bang we shall not consider it here, (see Peebles, 1971; Zel'dovich and Novikov, 1974.)

2.2. THE HELIUM PROBLEM

For possible direct evidence concerning a Universe younger than 30 years old we must turn to the helium problem. As is well known, there is good, but not decisive, evidence that the abundance of helium relative to hydrogen is approximately constant throughout the Galaxy and is about 10% by particle number (Searle and Sargent, 1972; Peimbert, 1974). Attempts to account for this helium abundance in terms of nuclear processes occurring in the Galaxy since its formation have not yet been successful. On the other hand, such processes occurring in the early stages of the Universe do yield a relative abundance of about 10%, the actual value depending slightly on the density of the Universe, being rather more in a high density than in a low density universe (Wagoner, 1973). Elements heavier than lithium, however, are not made appreciably in the hot big bang, so that the α-β-γ theory would not be valid for such elements.

If the helium we now observe is primordial, we thereby have access to processes occurring earlier than the most recent thermalisation epoch. The limiting factor here is that the radiation field must not be so intense that nuclei intermediate between hydrogen and helium are photo-disintegrated before they can be built up into helium. The maximum temperature permitted by this consideration is about 10^9 K, corresponding to a red shift of 3×10^8, and an age for the Universe of about 100 s. Thus if the helium argument could be made firmer, we would have direct access to a very early stage of the Universe indeed.

A similar argument has recently been made in connection with the formation of deuterium (Reeves *et al.*, 1973). This would be particularly interesting because the amount of deuterium formed in the early stages depends more sensitively on the density of the Universe than does the amount of helium. However, this argument is not a strong one at the moment because it appears that the deuterium may have been formed in the Galaxy (Hoyle and Fowler, 1973; Colgate, 1973). Similar remarks apply to the formation of lithium 7 (Audouze and Truran, 1973).

2.3. THE INITIAL SINGULARITY

To study earlier stages still one must turn to other evidence, namely the *isotropy* of the microwave background. We shall see that according to general relativity this isotropy implies that the Universe was singular a finite time ($\sim 10^{10}$ yr) ago. This is so even if one abandons the artificially strict symmetries of the Robertson-Walker models, and deals with a realistic universe. This singularity is widely regarded as physically unacceptable, but it seems likely that the density of the Universe was extremely high in its earliest stages. It will be convenient to defer discussion of this remarkable result until later.

3. Outstanding Problems

So far I have been considering recent developments in our understanding of the early stages of the Universe which, while not definitive, are at least clear-cut. Now it is

time to discuss the most important completely open problems. Amongst these I select the following:

 (a) What is the ultimate origin of the black body radiation field?
 (b) Why is the Universe so symmetric?
 (c) How can we eliminate the initial singularity?

I shall discuss these in turn.

3.1. THE ORIGIN OF THE BLACK BODY BACKGROUND

What requires explanation is not merely the existence of a black body radiation field, but its amount. It is not convenient to specify this amount in terms of the relative energy densities of the radiation field and of matter, because this relation would be time dependent in the expanding universe. What is time *independent* is the number of black body photons per proton in the Universe. This relative number would be about 10^8 in a high density universe and 10^9 in a low density universe. This number appears as a new kind of dimensionless 'constant of nature', and its observed value has to be explained. I must emphasise that the canonical hot big bang picture of Gamow does *not* offer any explanation. In this picture all the black body radiation was present in the Universe from the initial singularity, and the number of photons per proton is *completely arbitrary*.

Amongst the alternative proposals which have so far been made, there are two which stand out because they are amenable to rather detailed analysis, and appear to offer reasonable prospects of actually leading to the observed number of photons per proton. I shall confine myself to these two.

3.1.1. *The Matter-Antimatter Universe*

It has been argued by Harrison (1968), and in most detail by Omnes (1972) and his colleagues, that initially the Universe might have consisted of black body radiation *and nothing else*. Or rather that it consisted of black body radiation plus the thermally excited particle-antiparticle pairs that would be demanded by standard physics. At sufficiently early times, when the radiation temperature exceeded about 10^{13} K, there would have been roughly as many proton-antiproton pairs (and electron-positron pairs) as photons in the radiation field. If these pairs had remained intimately mixed up with one another, they would have annihilated progressively and completely as the Universe expanded and the radiation cooled. However, this end result is by no means what we observe to-day. In fact, we observe one proton left over for every 10^8 or so photons. This means that in the Gamow picture there were, during the earliest stages, $10^8 + 1$ protons for every 10^8 anti-protons. That extra proton was just what was needed to survive annihilation and to provide the material universe that we know to-day.

It would clearly be more attractive to do without the extra proton, if complete annihilation could be avoided by some separation process which kept matter and anti-matter sufficiently apart from one another. This is the problem which Omnes and his collaborators are currently trying to solve. What they must find is a separa-

tion process so efficient that a whole galaxy (and perhaps also a cluster of galaxies) should consist predominantly of matter of one sign, since this is a direct observational requirement. Given a separation process one could calculate, at least in principle, the amount of annihilation that would have taken place, and therefore obtain a definite value for the present ratio of photons to material particles. Thus this theory has within it the potentiality to account for the number 10^8.

This is a beautiful programme, but it suffers at the moment from two main difficulties. The first is that it has not yet been generally accepted that the separation mechanisms so far proposed are efficient enough to guarantee that galaxy-sized objects would separate out. This difficulty might, of course, be resolved by further work (and I personally hope that it will be, in view of the beauty and economy of the theory). The second difficulty is that essentially no helium would have been formed in the early stages (because at that time neutrons would have diffused too rapidly out of regions containing matter of one sign). Thus this theory requires one to find an alternative mechanism for manufacturing helium at a later stage, and would of course benefit from unambiguous evidence that the relative abundance of helium is not uniform throughout the Galaxy.

3.1.2. *Dissipative Processes in the Early Universe*

This explanation is the exact opposite of the previous one in the sense that in it the Universe begins with matter and no radiation, instead of radiation and no (net) matter. However, the potentiality for producing radiation is present in that it is assumed that the matter is undergoing irregular motions whose dissipation would lead to radiation being produced. These irregular motions may also play a decisive role in the formation of galaxies. The possible importance of dissipative processes was first pointed out by Misner (1967, 1968) and by Doroshkevich *et al.* (1967). In its most extreme form it supposes that initially the spectrum of disturbances corresponded in some sense to 'white noise'. In other words 'all' disturbances are assumed to be present initially with comparable amplitudes, and those disturbances not dissipated away should have survived to the present day. According to this approach, sometimes described as chaotic cosmology, any other assumption about the initial conditions would involve special restrictions which would require explicit justification.

Whatever view one might take about this methodological point, the chaotic cosmology programme has the advantage that, like the matter-antimatter programme, it has calculable consequences which can be compared with observation. The most attractive variant proposed so far is that due to Rees (1972) who assumes that the irregular motions to be dissipated involve velocities which are substantial compared with the velocity of light (mildly relativistic turbulence). If dissipation is efficient this would mean that the resulting radiation field would have an energy density ϱ_γ comparable with the rest density of matter ϱ_m. More precisely, as the Universe expands and the size of the particle horizon steadily increases, larger scale irregularities would become available for dissipation, and so the approximate equality $\varrho_\gamma \sim \varrho_m$ would be maintained until the largest-scale irregularity is reached. However, this process must

cut off at the latest when the matter is still capable of thermalising the heat produced by the dissipation; otherwise the microwave background would not now have a black body spectrum. If the cut-off does occur close to the last possible moment then, as Rees points out, the final photon/proton ratio would indeed be about 10^8, and also the largest-scale irregularities would correspond in mass to clusters of galaxies. These two beautiful consequences of the chaotic cosmology programme suggest that it should be investigated in much more detail. In particular its implications for the helium problem need elucidation. At first sight the predicted helium abundance appears to be on the high side, since the photon/proton ratio would have been less than 10^8 when the relevant nuclear reactions occurred. However, the assumed irregularities are likely to change the expansion time-scale in the early stages and this would affect the amount of helium produced. This question deserves further study.

3.2. WHY IS THE UNIVERSE SO SYMMETRIC?

The Robertson-Walker models are based on the assumptions that the Universe is exactly isotropic and homogeneous. These assumptions are certainly not true on a scale of, say, 100 Mpc (de Vaucouleurs, 1971), but the important question for cosmology is how true they are on a scale closer to, say, 1000 Mpc, that is on a scale somewhat smaller than the radius of the Universe itself (3000 Mpc). Here the best evidence comes from the observed isotropy of the microwave background, which is good to about 0.1% on a variety of angular scales (Partridge and Wilkinson, 1967; Conklin and Bracewell, 1967; Partridge, 1969; Parijskij, 1973; Carpenter et al., 1973). This is by far the most accurate observation of importance to cosmology. Regarding the immediate source of the background as the last-scattering surface, we can conclude that the radiation temperature is constant over the surface to a precision of 0.1%, and that the redshift of the surface, and so the rate of expansion of the Universe, is the same in all directions to this precision.

We can also draw conclusions about the large-scale homogeneity of the Universe (Sachs and Wolfe, 1967; Rees and Sciama, 1968). Consider for instance an irregularity on a scale ~1000 Mpc, in which the density were twice the mean density of the Universe. The presence of such an irregularity would influence the temperature of the black body radiation moving through it, and so would impose an anisotropy on the background, which would be of the order of 0.1%. The observed upper limit therefore places constraints on possible large-scale inhomogeneities in the Universe.

We conclude from this that on a large scale the universe is indeed both isotropic and homogeneous to high precision, and that the Robertson-Walker models are adequate representations at least back to the last scattering surface. Indeed they are better representations than mathematical cosmologists had previously had the right to expect. However, what is convenient mathematically may create serious physical problems. Here the problem is, why should the Universe be so symmetric? The urgency of this question derives from the fact that if the Robertson-Walker models are good representations at all epochs, then two points on the last-scattering surface more than, say, 30° apart would have been outside each other's particle horizon up

to the moment that the radiation we now observe was emitted. In other words, no causal influence, even moving with the speed of light, could have linked together the physical processes occurring in the neighbourhood of the two points. Yet their radiation temperatures are the same to within 1 part in a thousand!

This equality represents an unsolved problem. I would like, however, to mention briefly some of the solutions that have so far been proposed.

3.2.1. *Initial Conditions*

According to this proposal, the initial conditions governing the Big Bang simply had the necessary properties. While this may be the correct answer, no self-respecting cosmologist would adopt it while any alternative is conceivable.

3.2.2. *The Mix-Master Universe*

This proposal has the paradoxical property of trying to explain the observed isotropy by assuming that initially the Universe was highly anisotropic (though possibly homogeneous). The idea is that in the presence of a large early deviation from the Robertson-Walker symmetry particle horizons would no longer exist, so that dissipative processes might be able to remove the anisotropies in the time available. The detailed discussion of this idea is rather complicated because in these anisotropic models particle horizons are abolished first in one direction, then in another and so on, so that eventually the whole universe can become smoothed out; hence the name mixmaster universe (Misner, 1969). Unfortunately it now appears that models possessing this property form such a special class amongst all possible anisotropic models, that one is back to appealing to special initial conditions (MacCallum, 1971; Matzer and Misner, 1972).

3.2.3. *The Parabolic Universe*

Amongst the homogeneous (but possibly anisotropic) models of the Universe there are those which expand for ever with energy to spare (hyperbolic cases), those which just have the escape velocity (parabolic), and those which re-contract (elliptic). It has recently been discovered (Collins and Hawking, 1973a) that the hyperbolic models are unstable to the growth of a perturbation in which the anisotropy eventually increases with time. Unless therefore we live at a special time in such a model when the anisotropy happens to be very low we can conclude that the Universe is not hyperbolic. On the other hand in an elliptic universe an arbitrary initial anisotropy would not have time to die away to a low level before the universe re-collapses.

There remain the parabolic homogeneous universes, which do in general tend towards isotropy. However, as Collins and Hawking point out, if we assume that the Universe is parabolic, we are again appealing to a very special initial condition. Is there any independent way of understanding why the universe should have this special property? Collins and Hawking propose that there might be, following a suggestion made by Dicke (1961), Carter (1968) and others. According to this suggestion, one should contemplate the existence of a variety of universes in which the

initial conditions, and perhaps also the laws and constants of nature, take on all conceivable properties. Only in a very restricted subclass of cases would galaxies and stars form, and intelligent life develop. Thus our very presence would impose strict limitations on the laws and constants of nature, and on the initial conditions of the Universe. In particular Collins and Hawking suggest that it might be difficult for galaxies to form in hyperbolic universes, since self-gravitation may not be important during the later stages when the black body radiation itself has ceased to prevent galaxy formation. Their discussion of elliptic universes seems less satisfying since they argue that in this case there is no infinite future available in which an arbitrarily large initial anisotropy would die away, and that small density perturbations would not have time to develop into galaxies before the Universe re-collapses. However, the first of these arguments overlooks the likely fact that intelligent life is only possible for a finite time in a parabolic or hyperbolic universe also. In any case we certainly are observing the Universe at a finite time, and one whose value depends very little on the dynamical character of the Universe. Nevertheless it is an interesting point that a parabolic universe does not have this particular instability to an anisotropic perturbation, and this fact may be relevant to an understanding of the observed isotropy of the Universe.

3.2.4. *The Machian Universe*

According to Mach's principle local inertial frames are determined by the large-scale distribution of matter in the Universe and not by absolute space. One consequence of this principle would be that the Universe could not be in a state of absolute rotation. Astronomers usually take this for granted when they attempt to use extragalactic objects such as galaxies or quasi-stellar objects to define the best 'fixed' reference frame. However, general relativity gives us no grounds for believing it. There are cosmological solutions of Einstein's field equations in which the Universe is homogeneous at any one time, but is rotating relative to absolute space and so is anisotropic. In more technical language we would say that such a Universe possesses vorticity. This would mean that the whole Universe would appear to rotate relative to a local inertial frame of reference (e.g. relative to one in which a Foucault pendulum on the Earth oscillated in a fixed plane, or more correctly in view of the 'dragging of inertial frames' by localised rotating objects, relative to one which was appropriate for discussing dynamics on a scale of, say, 100 Mpc).

We must add a further condition to general relativity if we wish to eliminate this absolute rotation on the grounds that it is unphysical. This is where Mach's principle comes in (Sciama, 1971). Of course one should first ask the purely empirical question, with what precision is the Universe observed not to rotate relative to a local inertial frame? If the observed lower limit on the rotation period of the Universe were less than or of the order of 10^{10} yr, this would hardly constitute good evidence in favour of Mach's principle since 10^{10} yr is also the characteristic time associated with the size of the Universe. This was precisely the situation before the microwave background was discovered. At that time the best limit that could be placed on the rota-

tion period of the Universe derived from the absence of *transverse* Doppler shifts in the spectra of distant galaxies. This led to a lower limit for the rotation period of 10^{10} yr (Kristian and Sachs, 1966), which was too weak to support Mach's principle.

The situation is quite different now that the isotropy of the microwave background has been established to 0.1%. This isotropy imposes a constraint on the transverse Doppler shift associated with any rotation of the last-scattering surface, and so on the present rotation of the universe. This question has been analysed by Hawking (1969). His detailed conclusions are somewhat model-dependent, but in all cases the rotation period has a lower limit of at least 10^{14} yr. This result appears to give strong support for Mach's principle. It could be argued, however, that the comparison between rotation period and expansion time-scale is itself time-dependent, and that even if the rotation period is relatively long now, it need not have been in the past. This question has been studied by Collins and Hawking (1973b), who concluded that in any reasonable model of the universe the rotation period must at all times have been long compared with the expansion time-scale. They also succeeded in placing much sharper limits on the period than did Hawking, at least for the case of closed universes.

We conclude from this that the evidence in favour of Mach's principle is strong enough to justify trying to incorporate it into general relativity. We might expect this to give us a coherent reason for excluding those anisotropies of the universe which are due to vorticity. However, it appears that we can in this way also exclude shear, which would mean that (homogeneous) Machian universes would have to be completely isotropic. The detailed incorporation of Mach's principle into general relativity is a technical question which I cannot go into here. Suffice it to say that, starting out from work by Alt'shuler (1969), Lynden-Bell (1967), Sciama *et al.* (1969), and Gilman (1970), Raine (1974) has gone far towards formulating a rigourous criterion for Mach's Principle in general relativity. In terms of this criterion he is indeed able to show that a homogeneous Machian universe must be isotropic (at least when the matter in the Universe has a perfect gas equation of state, although this restriction is probably inessential). What must now be done is to show, as seems plausible, that a nearly-homogeneous Machian universe is nearly isotropic. If this statement could be made quantitative it would enable one to correlate different aspects of the angular distribution of the microwave background.

If this programme were successful there would still remain the problem of understanding why the Universe is so homogeneous. Mach's principle seems to shed no light on this. So long as, roughly speaking, there is sufficient matter everywhere in the Universe, so that one cannot get arbitrarily far from all the matter that there is, it would seem that Mach's principle could be satisfied whatever the length scale or the amplitude of the inhomogeneities. Their observed limitation to less than 1000 Mpc for amplitudes $\gtrsim 1$ remains a mystery.

3.3. How can we eliminate the initial singularity?

The Robertson-Walker models (with zero cosmical constant) contain a point singu-

larity a finite time ago in which the density of the Universe was infinite. Most cos-
mologists (though not all) regard this singularity as physically unacceptable. Until
recently it was thought that the singularity was an artefact of the assumption that
the Universe is exactly homogeneous and isotropic. In the presence of some irregular
motions one might have expected that, going backwards in time, the material of the
Universe would not be exactly focussed onto one point, but would either bounce or
interpenetrate without a singularity. It turns out that this is not so. In the most im-
portant theoretical development in cosmology since the Robertson-Walker models
were discovered, it has been found that, according to general relativity, self-gravita-
tion was so strong in the early stages of the Universe that the existence of a singularity
is guaranteed. The argument does, however, involve a number of (reasonable) as-
sumptions which I shall discuss in a moment. We owe this development mainly to
work by Penrose, Hawking and Ellis, and a full technical account of it may be found
in a comprehensive monograph by Hawking and Ellis (1973).

There is in fact a variety of singularity theorems, both for collapsing stars and for
expanding universes. Some of these theorems are based on assumptions which, though
plausible, could not actually be checked in practice. Fortunately there is one theorem
in the cosmological case where the basic assumptions can indeed be checked. Re-
markably enough, the existing observations of the microwave background can be
used for this purpose (Hawking and Ellis, 1968). In fact the background plays two
roles in the argument, namely through (i) its isotropy and (ii) its energy density.

The argument is based on the following theorem, which I first state precisely and
then comment on.

Space-time is not singularity-free if the following conditions hold:

(a) Einstein's field equations,

(b) The energy condition $(T_{ab}W^a W^b \geqslant \frac{1}{2}T$ for all time-like vectors $W^a)$,

(c) Strong causality (every neighbourhood of a point contains a neighbourhood of
that point which no non-spacelike curve intersects more than once),

(d) There exists a point P such that all past-directed time-like geodesics through P
start converging again within a compact region in the past of P.

Condition (b) is a very weak and plausible restriction on the energy-momentum
tensor, while a violation of (c) would perhaps be worse than a singularity, being a
global rather than a local breakdown of our ordinary physical concepts. The really
important condition is (d), which is a precise statement of the idea that the gravitation
due to the material in the Universe is sufficiently attractive to produce a singularity.
The behaviour of time-like geodesics required is reminiscent of that of light-rays in
a curved space-time. In the latter case the re-convergence would correspond to the
angular diameter of an object of fixed intrinsic size passing through a minimum as
the distance of the object increased from the point P (which we may take to represent
the Earth). This lens effect certainly occurs in Robertson-Walker models. In fact in
a high density universe the minimum occurs at a red shift of less than 2. Since many
quasi-stellar objects have red shifts exceeding 2 this behaviour is of importance in
the interpretation of measurements of their (radio) angular diameters.

The argument now proceeds by showing that the actual Universe is sufficiently like a Robertson-Walker one for the re-convergence to be preserved. In fact the observed isotropy of the microwave background shows that the universe is Robertson-Walker-like to high precision at least back to the last-scattering surface. The aim of the argument is therefore to show that the re-convergence begins before the last-scattering surface is reached. To simplify the argument we consider only the two extreme cases of a high density and a low density universe. In the former case the last-scattering surface occurs at a red shift of about 7. A straightforward calculation then shows that there is enough matter in this model for its gravitation to start producing re-convergence at a red shift less than 7. In the low density case we can be sure that the Universe is Robertson-Walker-like at least out to the much larger red shift of 1000. However, in this case there may not be enough matter to produce reconvergence at red shifts less than this. Nevertheless the argument is again saved by the microwave background, but now in its role as a contributor to the energy density of the Universe. For this contribution alone is sufficient to guarantee re-convergence before a red shift of 1000 is reached. A detailed investigation shows that models with intermediate density also exhibit re-convergence before the last-scattering surface. (The introduction of a non-zero cosmical constant with a value permitted by observation would not affect the argument, since its contribution to Einstein's field equations would be swamped by that of matter and radiation at and prior to the last-scattering surface).

We conclude that, according to standard theory, the actual Universe was singular at some time or times in the past. We here face a major crisis of fundamental science. The most conservative method of dealing with it would be to argue that when extremely high but finite densities are reached quantum mechanical and elementary particle effects would become important. In particular, it may become necessary to quantise the gravitational field. It has been argued that this might suffice to eliminate the singularity, and certain preliminary calculations do support this contention. However, other calculations suggest that the singularity may survive the quantisation process and therefore that new laws of physics would be needed to resolve the problem. The matter is a highly technical one which is unsettled at the present time. Whatever the outcome, it is clear that radio astronomy, general relativity, quantum field theory and elementary particle physics have combined together to present us with one of the greatest scientific problems of our time. It is a very difficult problem indeed, but I hope one day to listen to an Invited Discourse in which the solution is presented.

Acknowledgements

References

Alpher, R. A., Bethe, H. A., and Gamow, G.: 1948, *Phys. Rev.* **73**, 803.
Alt'shuler, B. L.: 1969, *Soviet Phys. JETP* **28**, 687.

Audouze, J. and Truran, J. W.: 1973, *Astrophys. J.* **182**, 839.
Bondi, H.: 1960, *Cosmology* (2nd ed.), Cambridge.
Bondi, H. and Gold, T.: 1948, *Monthly Notices Roy. Astron. Soc.* **108**, 252.
Brans, C. and Dicke, R. H.: 1961, *Phys. Rev.* **124**, 925.
Carter, B.: 1968, Unpublished Cambridge preprint.
Carpenter, R. L., Gulkis, S., and Sato, T.: 1973, *Astrophys. J.* **182**, L61.
Colgate, S.: 1973, *Astrophys. J.* **181**, L53.
Collins, C. B. and Hawking, S. W.: 1973a, *Astrophys. J.* **180**, 317.
Collins, C. B. and Hawking, S. W.: 1973b, *Monthly Notices Roy. Astron. Soc.* **162**, 307.
Conklin, E. K. and Bracewell, R. N.: 1967, *Phys. Rev. Letters* **18**, 614.
Dicke, R. H.: 1961, *Nature* **192**, 440.
Dicke, R. H.: 1962, *Rev. Mod. Phys.* **34**, 110.
Dicke, R. H., Peebles, P. J. E., Roll, P. G., and Wilkinson, D. T.: 1965, *Astrophys. J.* **142**, 414.
Dirac, P. A. M.: 1938, *Proc. Roy. Soc.* **A165**, 199.
Dirac, P. A. M.: 1973a, in *The Physicists' Conception of Nature*, North-Holland, in press.
Dirac, P. A. M.: 1973b, *Proc. Roy. Soc.* **A333**, 403.
Doroshkevich, A. G., Zel'dovich, Y. B., and Novikov, I. D.: 1967, *Soviet Phys. JETP* **53**, 644.
Field, G. B.: 1969, 'The Growth Points of Physics', *Rivista Nuovo Cimento* **87**.
Gamow, G.: 1948, *Phys. Rev.* **70**, 572.
Gamow, G.: 1949, *Rev. Mod. Phys.* **21**, 367.
Gilman, R. C.: 1970, *Phys. Rev.* **D2**, 1400.
Harrison, E. R.: 1968, *Phys. Rev.* **167**, 1170.
Hawking, S. W.: 1969, *Monthly Notices Roy. Astron. Soc.* **142**, 129.
Hawking, S. W. and Ellis, G. F. R.: 1968, *Astrophys. J.* **152**, 25.
Hawking, S. W. and Ellis, G. F. R.: 1973, *The Large Scale Structure of Space-Time*, Cambridge.
Hawking, S. W. and Penrose, R.: 1970, *Proc. Roy. Soc.* **A314**, 529.
Hoyle, F.: 1948, *Monthly Notices Roy. Astron. Soc.* **108**, 372.
Hoyle, F.: 1972, *Quart. J. Roy. Astron. Soc.* **13**, 328.
Hoyle, F. and Fowler, W. A.: 1973, *Nature* **241**, 384.
Jordan, P.: 1949, *Nature* **164**, 637.
Jordan, P.: 1973, in *The Physicists' Conception of Nature*, North-Holland, in press.
Kristian, J. and Sachs, R. K.: 1966, *Astrophys. J.* **143**, 379.
Lynden-Bell, D.: 1969, *Monthly Notices Roy. Astron. Soc.* **135**, 413.
Layzer, D. and Hively, R.: 1973, *Astrophys. J.* **179**, 361.
MacCallum, M. A. H.: 1971, *Nature Phys. Sci.* **230**, 112.
Matzner, R. A. and Misner, C. W.: 1972, *Astrophys. J.* **171**, 415.
Milne, E. A.: 1948, *Kinematic Relativity*, Oxford.
Misner, C. W.: 1967, *Nature* **214**, 40.
Misner, C. W.: 1968, *Astrophys. J.* **151**, 431.
Misner, C. W.: 1969, *Phys. Rev. Letters* **22**, 1071.
Omnes, R.: 1972, *Physics Reports* **3c**, 1.
Parijsky, Y. N.: 1973, *Astrophys. J.* **180**, L47.
Partridge, R. B.: 1969, *Am. Scientist* **57**, 37.
Partridge, R. B. and Wilkinson, D. T.: 1967, *Phys. Rev. Letters* **18**, 557.
Peebles, P. J. E.: 1971, *Physical Cosmology*, Princeton.
Peimbert, M.: 1974, in J. Shakeshaft (ed.), 'Formation and Dynamics of Galaxies', *IAU Symp.* **58**, in press.
Penzias, A. A. and Wilson, R. W.: 1965, *Astrophys. J.* **142**, 419.
Raine, D. J.: 1974, to be published.
Rees, M. J.: 1972, *Phys. Rev. Letters* **28**, 1669.
Rees, M. J. and Sciama, D. W.: 1968, *Nature* **217**, 511.
Reeves, H., Audouze, J., Fowler, W. A., and Schramm, D. N.: 1973, *Astrophys. J.* **179**, 909.
Rindler, W.: 1956, *Monthly Notices Roy. Astron. Soc.* **116**, 662.
Sachs, R. K. and Wolfe, A. M.: 1967, *Astrophys. J.* **147**, 73.
Sciama, D. W.: 1969, *The Physical Foundations of General Relativity*, Science Study Series.
Sciama, D. W.: 1971a, *Modern Cosmology*, Cambridge.
Sciama, D. W.: 1971b, in R. K. Sachs (ed.), *General Relativity and Cosmology*, Proceedings of the

International School of Physics 'Enrico Fermi'. Course XLVII, Academic Press, New York.

Sciama, D. W., Waylen, P. C., and Gilman, R. C.: 1969, *Phys. Rev.* **187**, 1762.

Searle, L. and Sargent, W. L.: 1972, *Comm. Astrophys. Space Sci.* **14**, 179.

Vaucouleurs, G. de: 1971, *Publ. Astron. Soc. Pacific* **83**, 113.

Wagoner, R. V.: 1973, *Astrophys. J.* **179**, 343.

Weinberg, S.: 1972, *Gravitation and Cosmology*, Wiley.

Zel'dovich, Y. B. and Novikov, I. D.: 1974, *Relativistic Astrophysics*. Vol. II: *Cosmology*, Chicago.