Cosmological models and their observational validation

1966 Rep. Prog. Phys. 29 539

(http://iopscience.iop.org/0034-4885/29/2/304)

Cosmological models and their observational validation

W. DAVIDSON† AND J. V. NARLIKAR‡

† Mathematics Department, University of Otago, Dunedin, New Zealand
‡ King's College, Cambridge

Contents

Abstract. The observational data of cosmology are presented in considerable detail and their interpretation in terms of various cosmological models is described. The data survey the nature and cosmic distribution of luminous matter (galaxies), radio sources, quasars, intergalactic matter, the radio and optical background radiation, x rays, γ rays and cosmic rays. The theoretical basis of cosmological models is described in its general aspects and in particular cases. Topics discussed include the homogeneous anisotropic models, isotropic models, Mach's principle, relativistic cosmology, Newtonian cosmology and the various versions of the steady-state theory and creation of matter. Theoretical observable relations are derived or described for the isotropic models of relativistic cosmology and the steady-state theory.

The theoretical relations between observables are compared with the data from various observational tests. These include the red-shift–apparent-magnitude relation, counts of radio sources, the radiation background at radio and optical frequencies and the x-ray and γ-ray background. The possibility of intergalactic cosmic rays is discussed and an account is given of recent observational tests for intergalactic hydrogen. The relevance to particular cosmologies of neutrino degeneracy, the absorber theory of radiation, the age distribution of galaxies and the origin of the chemical elements is analysed in some detail.

The most important conclusion is that several tests, among them the measurements of red shift and apparent magnitude, the counts of radio sources, the radio background at high frequency and the present He/H abundance ratio, all support a universe that was denser in the past. Furthermore, recent tests suggest that the density of hydrogen in intergalactic space is below the mean density of luminous matter (galaxies) by several orders of magnitude. Thus the evidence is now extremely strong against the steady-state cosmology in its original simple form, while relativistic cosmology to the above extent finds support. On the other hand, there is the difficulty that the evolutionary ages of some star clusters are estimated to exceed the predicted ages of galaxies in both cosmologies, while the absorber theory of radiation gives consistent retarded solutions in the steady-state theory but not in relativistic cosmology.

Although contrary arguments exist, there is strong theoretical evidence that a singularity in the relativistic cosmology is inevitable, and no way has yet been found in conventional physics either to prevent it or to describe it. However, the discovery of new force fields in physics at high density, on the lines of the negative energy C field of the steady-state theory, may show how a contracting universe may be reversed into expansion without singularity. This might imply a finitely oscillating universe, for which there are special difficulties. The theoretical and observational study of the recently discovered quasars may throw light on this issue. But from the point of view of present-day physics the evidence points to a fundamental singularity of the observable Universe that occurred about ten thousand million years ago. There is also considerable evidence that this was followed by a hyperbolic expansion.

1. Introduction to the present situation in cosmology

Both theoretical and observational cosmology have made very significant and encouraging progress in the last ten years. On the one hand, this period has been one of steady accumulation of relevant observational data, often obtained by feats of technical ingenuity on the part of radio and optical astronomers. On the other hand, it has been a time of severe testing and adjustment of cosmological theories to fit these observational data. The final solution to the cosmological problem is not yet in sight, but sufficient progress has been made to warrant optimism that it may not be far off. What is certain is that cosmology has become fully established as a science, in the correct sense of the word, in which theory and observation work constructively together to advance knowledge. Moreover, it has become clear that cosmological issues are likely to react, sooner or later, on most other branches of fundamental physics.

The principal protagonists among competing theories of the cosmos are still the models of relativistic cosmology and the steady-state theory, but the last few years have witnessed some attempts to close the gap between these two theories, at least superficially. This development has been inspired partly by the counts of radio sources by Ryle and his co-workers at Cambridge. These have shown decisively that a homogeneous steady-state universe, in which radio sources are evenly distributed throughout space, must be ruled out. Prior to these results, Gold and Hoyle had already envisaged a scale of inhomogeneity in a steady-state universe, prompted by difficulties in the theory of formation of galaxies in that cosmology. These ideas were extended by Hoyle and Narlikar to accommodate the Cambridge counts. By postulating the existence of large age-correlated regions in a universe which is overall in a steady state, Hoyle and Narlikar showed that under certain circumstances such a model might make predictions radically different from those of the steady-state theory. In particular, if the potentiality of a galaxy to become a radio source were age-correlated, then for a certain distance radio sources might increase in number density as one looked further into space from our neighbourhood.

However, since then the original Gold–Hoyle theory of galaxy formation has been criticized on the ground that the very high-temperature state of intergalactic matter that it requires is contradicted by observational evidence.

Another instance of an apparent rapprochement between the two kinds of cosmology is provided by the recent field theory derived for the steady-state model by Hoyle and Narlikar. This indicates that the creation of matter in that cosmology does not take place uniformly throughout space as previously assumed, but is strongly influenced by any inhomogeneities present in the Universe; it now appears that creation occurs mainly in the neighbourhood of existing concentrations of matter. It is tentatively suggested that creation of matter may be taking place in the recently discovered, extraordinarily powerful quasars, in radio sources and in massive galaxies. Further, owing to an essential instability in the creation process, the creation of matter in any region will die out if it expands and intensify if it contracts. Hoyle and Narlikar therefore suggest that we live locally (on the scale of the Hubble radius) in an expanding 'bubble' of galaxies, in which there is now negligible creation and which bears considerable similarity to the exploding universes of general relativity.

It might then be argued, perhaps, that it would be simpler to make a search first for a suitable model suggested by relativistic cosmology. Indeed, considerable success has been achieved in this direction. The work of Davidson and Davies has shown that a restricted range of exploding models, in which the existence of radio sources is necessarily limited in past time, can fit the radio counts very closely. This result depends on the plausible assumption that the mean power of the sources increases as we look into the past. Again, the observations of optical astronomers, such as Humason, Mayall and Sandage, and Baum, and again very recently, Sandage, indicate that the measurements of red shift (fractional shift of spectral wavelength) of distant galaxies, as a function of apparent magnitude (brightness), are consistent with the retarded expansion of the exploding models, rather than the accelerated expansion predicted by the earlier form of the steady-state theory. There seems every likelihood that, as recently indicated by Sandage, the very powerful, and therefore very distantly observed, quasars will allow this programme to be much extended. The greatest red shifts observed by Schmidt, Burbidge and others for these objects now exceed $z = 2$. One difficulty here will be to find an exact value for the intrinsic luminosity of these objects at different red shifts, without begging the solution to the cosmological problem. An alternative course may be to make an apparent-magnitude count of quasi-stellar galaxies, the radio-silent but optically brilliant objects recently discovered in some profusion by Sandage.

Further evidence of a high-density, high-temperature state of the Universe in the past has been very recently reported by Penzias and Wilson, who find a present background radiation temperature at radio wavelengths considerably greater than that expected in an expanding universe from the galaxies and radio sources as observed locally now. With this question there is associated the fascinating problem of the neutrino density in the Universe, and of course the still unanswered query: where do the high-energy cosmic rays come from? These problems have also a possible link with the quasars and with the powerful x-ray sources and the background of x rays recently investigated by Giacconi, Friedmann and others.

If a model of the Universe, involving a presently retarded expansion from an earlier high-density state, continues to receive observational support, it would be very satisfactory if, as already adumbrated, steady-state theory and relativistic cosmology could both countenance this situation theoretically, even if from different points of view. But the task of providing any satisfactory theoretical treatment of such a state of the Universe remains one of the most formidable in cosmology. In the case of relativistic cosmology this question raises issues that are receiving most energetic investigation at the present time.

For example, is an apparent singular state in the finite past an inevitable theoretical result in relativistic cosmology? Or can the possibility of local angular momentum, increasing the spin velocity as the mass density increases, prevent this situation, as has been suggested by Heckmann and Shucking but denied by Narlikar? Alternatively, can local inhomogeneities in a contracting universe controvert a singular state, as argued by Lifshitz and Khalatnikov and countered by Hawking? To what extent, in any case, are these considerations relevant to an exploding universe which is not oscillatory? Granted that the singularity always exists, is it due to an inadequacy of general relativity, or to a gap in our knowledge of basic physics, or both? In this connection, can that conundrum, the creation of primary matter in the Universe, be resolved in the context of general relativity and contemporary physics, or must we also admit some of the ideas of the steady-state theory? Finally, can galaxies in an expanding universe governed by general relativity form more easily than in one governed by steady-state cosmology, and by what mechanism?

Although extremely difficult, many of these questions have lost the imponderable or intractable character that they possessed a few years ago. The properties of an ultra-high-density state of matter, in a region of sufficient extent for gravitation to be important, have only recently begun to be investigated, but already there are signs that it is here that the key to many of the above questions may lie. The recent discovery of massive, compact aggregates of extremely luminous matter, the quasi-stellar sources (quasars) and quasi-stellar galaxies, have stimulated novel and thought-provoking theories to explain them. Can an almost catastrophic conversion of a substantial portion of their rest mass, apparently necessary to account for their enormous power, be explained by a super-supernova gravitational collapse, as has been suggested by Hoyle and Fowler, Wheeler and others? Is a collapsing quasar a test case for a collapsing universe?

Apart from these questions relating to the large-scale structure of the Universe there are others which connect cosmology with the local laws of physics. Is the local inertial frame (in which Newton's laws of motion hold) determined by cosmological considerations (Mach's principle)? Are the irreversible phenomena in our locality (which serve to define an arrow of time) strongly connected with the expansion of the Universe as suggested by Hogarth and by Hoyle and Narlikar, or the contrary as maintained by Percival, Penrose and others?

These and other important questions, the progress so far in answering them, together with the basic phenomenon which gives rise to them, will be considered in some detail in the following pages.

2. Observational data in cosmology

2.1. *The cosmic distribution of luminous matter*

The basic features of the cosmic scene are now well known—the approximately even distribution of galaxies, and clusters of galaxies, in all directions about our own Galaxy, and the recession of the galaxies from us at a speed proportional to their distance. These remarkable discoveries by Hubble 35 years ago have been well substantiated in recent years, with added detail rather than drastic revision.

Hubble found that, in a given range of apparent brightness, elliptic (E) and lenticular (SO) galaxies constituted about 20% of the total population of galaxies, while the spirals (Sa, Sb and Sc of increasing openness), and the barred spirals (SBa, SBb, SBc) and irregular galaxies (Ir) accounted for the remainder. The same ratio has been found by Van den Bergh (1961 a) in a sample of galaxies brighter than the twelfth apparent magnitude†. Out of a total of 768 galaxies Sandage (1961 a)

† Apparent magnitude m in astronomy is connected to the apparent luminosity l, that is the rate at which radiant energy is received per unit area outside the Earth's atmosphere, by the relation

$$m = -2\cdot5 \log l + \text{const.}$$

A zero apparent magnitude corresponds to

$$l = 2\cdot5 \times 10^{-5} \text{ erg cm}^{-2} \text{ sec}^{-1}.$$

This may be compared with the total apparent luminosity of the Sun (the solar constant), viz.

[Continued overleaf

found that E and SO galaxies accounted for 25% of the total. Of the remainder, Sc galaxies are the most common, about 34%, the Sa, Sb together accounting for 23%. About 15% are barred spirals and irregular galaxies constitute about 3%. A typical bright galaxy has a diameter of 25–30 kparsec†, including the outermost region (halo), and a mean density of matter (stars and gas) of about 10^{-24}–10^{-23} g cm^{-3}.

As has been stressed by Van den Bergh (1961 b) and Oort (1958), the gathering of galaxies into clusters is an undoubted phenomenon on a scale of 50 Mparsec, varying from 20 member galaxies in our local group to 2500 in the Virgo cluster. However, above this scale up to a surveyed distance of 10^9 parsec Oort (1958) and Abell (1958) find an apparently homogeneous and isotropic distribution of clusters interspersed by field galaxies. Oort concludes that "all available evidence supports the concept of a universe which is homogeneous on a large scale". Within a distance of 10^9 parsec there are perhaps five million clusters and a thousand million galaxies.

The distance scale is derived by the method of Hubble, through the successive links: Cepheid variables in the nearest galaxies, brightest stars in fainter galaxies, and mean standard galaxy assumption for the most distant galaxies. This scale has been made more accurate owing to investigations by Baade and Swope (1955) and by Sandage (1958), particularly regarding the first two steps of the chain. Regarding the last step, recent statistical examinations of the luminosity function of galaxies have been made by Van den Bergh (1961 a) and Kiang (1961). It is indicated by Kiang's work that the mean absolute photographic magnitude‡ of a galaxy, observed in a given range of apparent magnitude, is

$$\overline{M}_{pg} = -20{\cdot}3 \pm 1{\cdot}6. \qquad (2.1)$$

This corresponds to a mean emission L given by

$$L = 4 \times 10^{43} \text{ erg sec}^{-1} \qquad (2.2)$$

or $10^{10}\, L_{\odot}$.

On the other hand, Kiang finds that the number of galaxies in unit volume of space varies as $(M-M_0)^3$ for the brighter galaxies, and as $10^{0{\cdot}2(M-M_0)}$ for the fainter galaxies, M_0 being the brightest magnitude observed (about -22) with M varying between -22 and -15. Thus there appears to be a very large number of faint galaxies. Indeed, the total density of galaxies in space is found by Kiang to be

$$n = 0{\cdot}46 \text{ galaxies/Mparsec}^3 \qquad (2.3)$$

and his results show that the total emission of light from space, in terms of galaxies

$1{\cdot}4 \times 10^6$ erg cm^{-2} sec^{-1}. Apparent magnitudes may be bolometric (m_{bol}) corresponding to total apparent luminosity, visual (m_v) corresponding to apparent luminosity in the visual part of the spectrum, or photographic (m_{pg}) corresponding to that part of the spectrum registered on a photographic plate. They may also be measured for selected colours; for example, U denotes the apparent magnitude in the ultra-violet, B in the blue and V in the visual or yellow.

† 1 parsec = 3·26 light years.

‡ Absolute magnitude, denoted by M, measures intrinsic luminosity and is defined as the apparent magnitude at a distance of 10 parsec from the object. A zero absolute magnitude corresponds to an intrinsic luminosity L equal to 3×10^{35} erg sec^{-1}. This may be compared with the Sun's total emission $L_{\odot} = 3{\cdot}9 \times 10^{33}$ erg sec^{-1}. As for apparent magnitude the absolute magnitude may be bolometric, photographic or visual, etc.

with $M = \overline{M} = -20{\cdot}3$, is equivalent to 0·03 galaxies/Mparsec³. In terms of galaxies with $M = -15$ it is 3·7 galaxies/Mparsec³. The total emission is therefore

$$E = 3{\cdot}2 \times 10^{-32}\ \mathrm{erg\,sec^{-1}\,cm^{-3}} \tag{2.4}$$

or $2{\cdot}8 \times 10^{8}\, L_{\odot}$ Mparsec^{-3}, which is a result in close agreement with conclusions by Van den Bergh (1961 a) and Oort (1958). Therefore, while the mean luminosity of galaxies observed in a given range of apparent magnitude is $10^{10}\, L_{\odot}$, the mean luminosity of galaxies in space is only about $6 \times 10^{8}\, L_{\odot}$.

The estimates of the mass of a galaxy depend on the application of the virial theorem to clusters of galaxies, which gives the total mass of the cluster, or on the dynamics of double and triple galaxies, or on the rotation of a single galaxy. However, these do not always agree. For example, by using the virial theorem Van den Bergh (1960 a) reports a mass-to-light ratio in terms of $M_{\odot}/L_{\odot}$ of 900 for the Coma cluster, which consists mainly of E and SO galaxies. But an examination of the dynamics of multiplets of E and SO galaxies by Page (1959) shows that the mass-to-light ratio is about 40. A similar discrepancy has been found by Van den Bergh (1960 b) for the Virgo cluster. It may well be that the virial theorem does not apply to clusters of galaxies, since this assumes a relaxed state of negative energy. It has been suggested by Ambartsumian (1958) that many large clusters, including the Coma and Virgo clusters, are systems of positive energy in the process of dispersing. Zwicky, on the other hand, attributes the discrepancy to the presence, perhaps 95%, of mass in the form of non-luminous matter.

From rotation of galaxies and dynamics of multiplets Oort adopts a mass-to-light ratio of 50 for E and SO galaxies, 20 for Sa and Sb spirals and 7 for Sc spirals. With these results the giant ellipticals have a typical mass of $5 \times 10^{11}\, M_{\odot}$, while that of the Sc spirals is about $4 \times 10^{10}\, M_{\odot}$. Taking account of the luminosity function of galaxies he finds that the average mass-to-light ratio in space is 21. This implies a mean emission ratio

$$\frac{L}{M} = 0{\cdot}1\ \mathrm{erg\,g^{-1}} \tag{2.5}$$

for the luminous matter of the Universe. From the previous determination of the mean light emission from space (equation (2.4)) one obtains the value

$$\rho\ (\text{luminous matter}) = 3{\cdot}2 \times 10^{-31}\ \mathrm{g\,cm^{-3}} \tag{2.6}$$

for the extragalactic mean density of observed matter in the Universe. Van den Bergh (1961 a) assumes a mean mass-to-light ratio of 30 and, using his own luminosity function, obtains a value $6{\cdot}8 \times 10^{-31}\ \mathrm{g\,cm^{-3}}$ for the mean density. Zwicky and others would favour a value of at least $10^{-30}\ \mathrm{g\,cm^{-3}}$ and higher values are predicted for the total density in several cosmological models. Present determined efforts to estimate the amount of intergalactic hydrogen in the Universe may help to resolve this issue (§ 6.5).

2.2. *The red shift*

Following Hubble, measurements of the fractional shift z of wavelength of spectral lines of a galaxy (relative to laboratory or solar wavelengths) have been systematically extended in recent years.

Interpreted formally in terms of a velocity of recession, in accordance with the view now accepted by most astronomers, the classical Doppler relation

$$V = cz \tag{2.7}$$

yields the velocity of a galaxy showing red shift z. Since a generally accepted cosmological model is lacking, a more accurate formula for the large values of z recently measured is that given by special relativity, namely

$$V' = \frac{c(2z+z^2)}{2+2z+z^2}. \tag{2.8}$$

This, of course, gives (2.7) if z is small. Evidently $V' \to c$ as $z \to \infty$, which corresponds to the observational 'horizon' of the Universe from where all electromagnetic radiation is red-shifted to zero frequency by the time it reaches the observer.

The 'luminosity distance' D of a galaxy is measured by its apparent magnitude m in accordance with the conventional inverse square law of diminution of its apparent luminosity. It follows that if D is in parsecs the definitions of m and absolute magnitude M yield the relation (for galaxies of small red shift)

$$\log D = 0{\cdot}2(m-M)+1.$$

For more distant sources m, as measured, must be corrected for the selective effect of the photographic plate on the received spectrum, since the latter is now red-shifted compared with a standard nearer source. The relation is now

$$\log D = 0{\cdot}2(m-K-M)+1 \tag{2.9}$$

where K is the 'red-shift correction' depending on z and calculable empirically if the emission spectrum is assumed to be known.

Evidently if we are given the measured m and z, V and D follow. Hubble found that up to Doppler velocities of 40 000 km sec^{-1}, V was proportional to D, viz.

$$V = HD \tag{2.10}$$

where H is Hubble's constant.

A large number of measurements of z and m were published by Humason *et al.* (1956) carrying z to a value of 0·2, corresponding to a classical velocity of one-fifth of the velocity of light. These were extended by Baum (1957, 1961 a, b), using multicolour photoelectric methods, to $z = 0{\cdot}46$. This is the red shift of the cluster containing the very powerful radio galaxy 3C 295 identified by Minkowski (1960). These astronomers still find a nearly linear relation between $\log(cz)$ and m_c $(= m-K)$, and hence approximately between the defined V and D (figure 5). The coefficient H in (2.10) is found to have a reciprocal

$$T = H^{-1} \simeq 10^{10} \text{ year} \tag{2.11}$$

with an error of perhaps 30%. This implies a general expansion of the Universe at a rate of 100 km sec^{-1} Mparsec^{-1}. With this value of H a red shift of 0·46 corresponds to a distance of approximately 1600 Mparsec, and a time lapse between emission and receipt of the radiation of about 3500 million years.

As mentioned in § 2.1, the errors inherent in Hubble's early criteria for determining the distances of the nearer galaxies were cleared up in the 1950's. These corrections evidently affect the assumed standard value of M for the more distant galaxies and, in turn, their deduced distances via (2.9). Finally, it affects H in

(2.10) and the net result has been a sixfold increase in the distance scale of the Universe. The new scale seems likely to be nearly correct, since it makes the size of the nearby spirals, as deduced by angular diameter measurements, comparable with that of the Galaxy, instead of being considerably smaller. It also considerably reduces the difficulty of the time scale of the Universe raised by Hubble's results, which made the age of the exploding models of general relativity less than the geological age of the Earth. This question will be considered further in § 6 when we deal with the information provided by higher-order terms of the (m, z) relation.

For normal galaxies measurements of apparent magnitude and red shift become extremely difficult for $z > 0{\cdot}3$. However, the recent discovery of the quasi-stellar sources (quasars) and quasi-stellar galaxies (§ 2.4) of a brilliance 100 times that of a normal galaxy has led to a staggering increase in the measured values of z. Apparent magnitudes for about 70 quasi-stellar sources (Sandage 1965, Bolton *et al.* 1965, Sandage *et al.* 1965, Wyndham 1965, Calif. Inst. Technol. Radio Obs. Yellowback, No. 13) and red shifts for 22 of these objects (Schmidt 1965 b, Lynds *et al.* 1965, Oke 1965, Burbidge 1965 a, b, 1966) are now available. Nine of the red-shift values, obtained mainly from emission lines displaced from the far ultra-violet, exceed unity, so that in the absence of a known cosmological solution the formula (2.8) (for V') must be used. One of the greatest red shifts is that of the quasi-stellar source 3C 9, viz. $z = 2{\cdot}012$, which implies by (2.8) a recessional velocity of four-fifths of the velocity of light. However, more measurements of red shift will have to be made for these objects before they can have adequate statistical accuracy for use in the (m, z) relation (see § 6).

2.3. *Extragalactic radio sources*

The rapid development of radioastronomy in the last fifteen years has had profound influence on most other branches of astronomy, and on cosmology in particular in the last seven years. Since the discovery in the early 1950's that extremely powerful sources of radio emission were located outside the Galaxy, it has become confirmed that these sources are distributed in space with even greater large-scale homogeneity and isotropy than are the galaxies, of which they appear to form an abnormal sub-set. Moreover, because of their enormous power and their relatively flat spectra, and also because of the increasing sensitivity of interferometric telescopes formed by linked arrays of antennae, these objects can be clearly detected out to distances of cosmological significance (of order cT where $T = H^{-1}$).

A very large number of sources has now been catalogued or recorded by the Cambridge group, the Sydney group and the California Institute of Technology Radio Observatory. The Cambridge surveys are confined to the northern hemisphere and, in particular, the revised 3C survey (Bennett 1962) at 178 Mc/s covers that area of the sky for which the declination is greater than $-05°$ and contains 328 sources, the flux density S of which exceeds 9 flux units:

$$1 \text{ flux unit} = 10^{-26}\ \mathrm{w\,m^{-2}\,(c/s)^{-1}}.$$

The 4C survey (Scott and Ryle 1961), which is nearly complete, will cover a similar area at 178 Mc/s down to $S = 2$ flux units. This survey indicates that down to this level there are about 6500 radio sources in the whole sky. Preliminary results have been published (Ryle and Neville 1962) of a more sensitive survey in the north polar area of the sky, using a new version of the technique of aperture synthesis (moving one telescope of an interferometer relative to the other). These show that down to $S = 0{\cdot}25$ flux units there are 10 000 sources/sterad or 125 000 in the whole sky.

By comparing the counts in depth in different areas of the sky Scott and Ryle have demonstrated the remarkable isotropy in the large-scale distribution of the sources about the Galaxy, and therefore presumably about any point in space. Shakeshaft *et al.* (1955), Mills *et al.* (1958, 1960) and Edge *et al.* (1959) have shown that there is no detectable clustering on the large scale of an angular size 3–30°. Leslie (1961) has found an extraordinary degree of isotropy and absence of clustering by sampling numerous sky areas of dimensions 3′ to 200′. All these results have immensely encouraged theoreticians to believe that cosmological models assuming a large-scale homogeneity and isotropy of space are on the right lines.

The once difficult question of the distance scale of the radio source distribution has now been answered. To establish this one must somehow ascertain the intrinsic power of the sources. If $\bar{n}_0$ (Mparsec^{-3}) is their mean space density at the present epoch t_0, and $\bar{P}_0$ (w sterad^{-1} (c/s)$^{-1}$) is the median power of the sources observed in a given range of high flux density ($S \to \infty$), then in any homogeneous cosmological model the number $N(S)$ (sterad^{-1}) for S large is (Ryle and Clarke 1961, Davidson 1962 a)

$$N(S) = \tfrac{1}{3}\bar{n}_0 \bar{P}_0^{3/2} S^{-3/2} \tag{2.12}$$

as in a static Euclidean universe. The actual number counts for high S give the coefficient $\bar{n}_0 \bar{P}_0^{3/2}$ (figure 7). Also, a maximum value of $\bar{n}_0 \bar{P}_0$ was deduced by Ryle and Clarke from the upper limit to the intensity of background radiation that could be attributed to the radio sources. The combination of these data showed that $\bar{P}_0 > 10^{24}$ w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s, and incidentally that the sources could not be located in the Galaxy without making artificial assumptions. Subsequent extensive measurements of the angular diameter of the sources, mainly by the Jodrell Bank group and the California Institute of Technology, coupled with the information provided by optical identifications, have shown that the median power of the sources at high flux density is

$$\begin{aligned} \bar{P}_0 &= 4 \times 10^{25} \text{ w sterad}^{-1} (\text{c/s})^{-1} \\ &= 4 \times 10^{32} \text{ erg sterad}^{-1} (\text{c/s})^{-1} \quad \text{at 178 Mc/s} \end{aligned} \tag{2.13}$$

with a dispersion which is approximately logarithmically Gaussian of standard deviation unity. A very recent analysis by Longair and Scott (1965) gives full weight to the identification of quasars, and this would revise the mean power upwards to 8×10^{25} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s (figure 1). However, most of the quasars are so distant that it may not be justified to regard them as strictly 'local'. Indeed, they may be regarded as evidence for the higher mean source power at earlier epochs which seems to be necessary to explain the radio-source counts in evolutionary cosmology (§ 6.2). Calculations in this report will therefore be based on the value of $\bar{P}_0$ given in (2.13).

This value of $\bar{P}_0$ and the known value of $\bar{n}_0 \bar{P}_0^{3/2}$ now lead to

$$\bar{n}_0 = 3 \times 10^{-7} \text{ Mparsec}^{-3} \tag{2.14}$$

so that, referring to (2.3), galaxies are $1{\cdot}5 \times 10^6$ times more numerous than radio sources, the latter having an average spacing of 240 Mparsec compared with 2 Mparsec for galaxies. The magnitude of the distance scale becomes obvious, if we consider that at the limiting flux density, 2 flux units, of the 4C survey even a source of the median power 4×10^{25} w sterad^{-1} (c/s)$^{-1}$ would be at a distance of order $1{\cdot}5 \times 10^{9}$ parsec, with a red shift of 0·4.

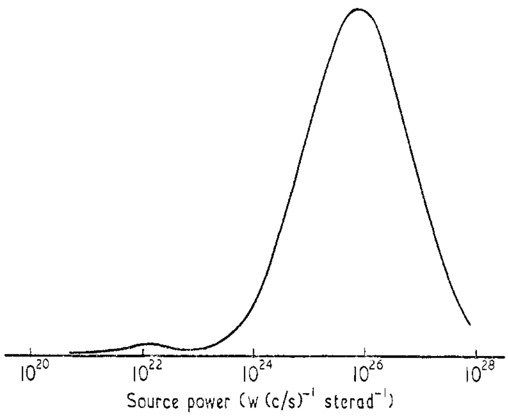

Figure 1. The power distribution of radio sources is shown at 178 Mc/s. The ordinate indicates to an arbitrary scale the relative number of sources (observed in a given range of flux density) of the power P_{178} (w (c/s)$^{-1}$ sterad^{-1}) shown in the abscissa. The median power $\bar{P}_{178}$ is found to be 8×10^{25} w (c/s)$^{-1}$ sterad^{-1}. This distribution takes full account of the identified quasars. (From Longair and Scott 1965.)

Since radio positions of the sources can now be given accurate to 2″ of arc a large number of optical identifications has been made. Of the 328 sources in the revised 3C catalogue ($S > 9$ flux units, $\delta > -05°$) about 160 have been fairly certainly identified with extragalactic objects, of which about 100 are galaxies of abnormal appearance and spectroscopic peculiarities and 60 are the newly discovered quasars. There are also about 25 sources of large angular diameter identified with H_{II} regions or supernova remnants of the Galaxy, such as Cassiopeia A and the Crab nebula. Another 40 sources are unlikely to be easily identified as they lie near the obscuring plane of the Galaxy. There are only about 20 normal galaxies in the catalogue since, although many nearby galaxies are known to radiate at radio wavelengths (Hanbury Brown and Hazard 1961 a, b), as in the case of the Milky Way, their total power is only about 10^{21} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s. Thus a normal galaxy would have to be nearer than 3×10^{6} parsec to have $S > 9$ flux units.

Over 100 sources have also been identified with galaxies and quasars in the southern hemisphere, using the Parkes catalogue at Sydney, and about 88 identifications have been made from the California Institute of Technology (CTA) catalogue.

Of the identified extragalactic sources red shifts have been obtained in the last three years for about 65 radio galaxies, mainly by Schmidt, Minkowski, Greenstein, Morgan and Matthews of Mount Palomar Observatory and the California Institute of Technology Radio Observatory. Red shifts have also been obtained for 22 quasars.† These achievements have helped greatly to establish the power distribution of the sources.

† See references given in § 2.2 earlier.

The measurements of angular diameter by interferometer at Jodrell Bank (Allen *et al.* 1962, Rowson 1963) and at the California Institute of Technology (Moffett and Maltby 1962) have been made for a large number of sources having $S > 12$ flux units. The majority of these sources have angular diameters between 1″ and 300″ of arc, with an average of 30″, and this has led to estimates of their surface temperatures in the range 10^5–10^8 °K. About three-quarters of the sources consist of two radio components, the distance measurements by red shift indicating that on average each component has a diameter of about 25 kparsec separated by a distance of 100 kparsec.

Theory and all the observational evidence go to prove that these radio components are clouds of high-speed electrons and other charged particles emitting synchrotron radiation in a local magnetic field. The analysis of the spectra of power at different frequencies ν reveals that $P(\nu) \propto \nu^{-x}$, where $0{\cdot}6 < x < 0{\cdot}8$ (Conway *et al.* 1963) except at high frequencies. For $\nu > 1420$ Mc/s, x approaches unity or more, and 10% of the sources have strongly curved spectra. By integration of the spectral emission over different frequencies it is found that many sources have a total radio output of about 10^{45} erg sec^{-1} which exceeds the optical emission of the brightest galaxies. The colossal total energy in the form of electron motion and magnetic fields in a typical radio galaxy is estimated to lie in the range 10^{60}–10^{62} erg, the energy equivalent of 10^6–10^8 solar masses. At the observed power this would give a lifetime of 10^7 or 10^8 years.

The optical identifications indicate that the radio-emitting clouds have almost certainly been ejected violently from the associated galaxy, which is frequently of the spherical or elliptic type and lies between the clouds. Numerous galaxies which are radio sources have been observed with luminous jets protruding from them, such as M 87 in Virgo. Others like the remarkable M 82, photographed in the red light of $H\alpha$ emission by Sandage (Burbidge *et al.* 1963, Sandage 1964), seem to have undergone a gigantic explosion about a million years earlier. It appears highly likely that the radio galaxies are closely related to the quasi-stellar sources and quasi-stellar galaxies discussed below.

2.4. *Quasi-stellar sources (quasars) and quasi-stellar galaxies*

The fact that not all of the radio sources in the 3C catalogue have been identified is not surprising. Such sources are likely to be at the higher end of the power distribution. For to have $S > 9$ flux units stronger sources need not be so near as those less powerful. For example, a source of power $P = 10^{26}$ w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s and of flux density $S = 10$ flux units will be at a distance 10^9 parsec. Assuming it has the average optical absolute magnitude of $-20{\cdot}3$, which in fact is the average for the identified sources, it would have an apparent magnitude of about 20, which is approaching the limit of the 200 in. telescope. On the other hand, this difficulty would not arise for intrinsically much brighter objects (optically), and such are the quasars which have turned out to constitute about 30% of all identified sources.

The first few of these objects were identified in 1963 after exact positions, in some cases by the technique of lunar occultation, had been provided for certain unresolved radio sources of small angular diameter, less than 1″ of arc. The compact spherical images identified in the radio positions were thought to be stars (hence 'quasi-stellar source') until Schmidt (1963) recognized the spectrum of 3C 273B (the 'stellar' component of 3C 273, the component 3C 273A being a luminous jet providing most of the radio emission at low frequency) as one of an object of red shift 0·16. As its apparent photographic magnitude was 13 this meant that its absolute magnitude was $-25{\cdot}4$ corresponding to an optical emission of 4×10^{45} erg sec^{-1}, or 100 times that of the Galaxy. Similar powers have been shown to belong to the other quasi-stellar sources of known red shift (cf. figure 6). Their radio powers are comparable with those of the strongest radio sources.

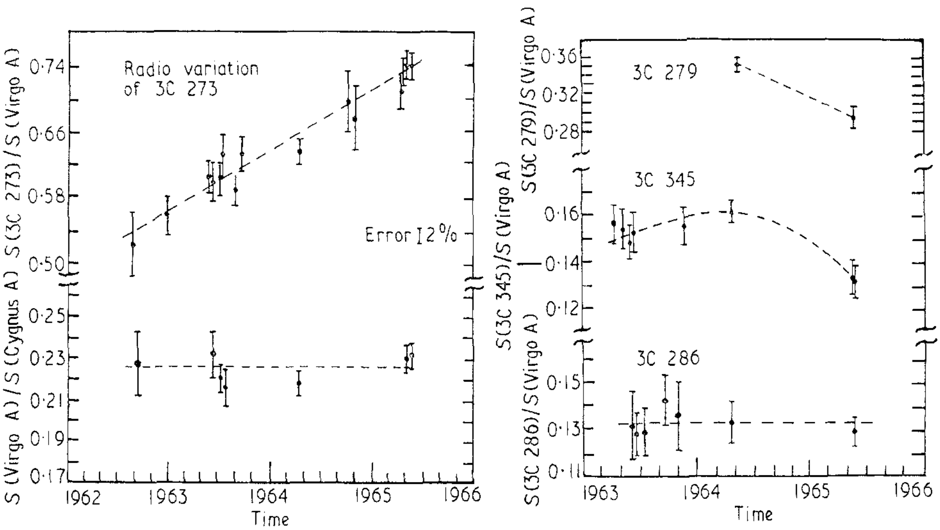

Figure 2. This shows the variation of radio power received from certain quasi-stellar sources at 8000 Mc/s. The graph on the left shows a 40% increase in the flux density S of 3C 273 over the last three years. This was measured as a ratio of the flux density of 3C 273 to that of Virgo A (upper left), the latter having remained essentially constant in power when measured against Cygnus A over the same period (lower left). Similar variations of power are shown on the right for the quasi-stellar sources 3C 279 and 3C 345. All these sources have flat spectra ($x \simeq 0$). On the other hand, quasi-stellar sources with curved spectra such as 3C 286 (lower right) and 3C 147 (not shown) indicate no such variations. (From Dent 1965.)

The quasi-stellar galaxies are the objects very recently discovered by Sandage (1965), which were formerly thought to be blue halo stars of the Galaxy. They are very similar to the quasi-stellar sources optically but are silent, or have relatively weak emission, at radio frequencies. Three red shifts for the quasi-stellar galaxies have been obtained by Sandage, the largest being $z = 1{\cdot}24$. They appear to be at least 50 times more numerous than the quasi-stellar sources and the two categories may represent different stages of a single evolutionary process (figures 6 and 13).

Despite their phenomenal power the physical size of the quasi-stellar sources is amazingly small. Very recent determinations of some of their radio angular diameters by Adgie, Palmer and others at Malvern and Jodrell Bank are of the order of 0·1″ of arc (Adgie *et al.* 1965). For the known red shifts this indicates a linear diameter of 200 parsec. Furthermore, several quasi-stellar sources that have been studied in detail fluctuate very considerably in both radio brightness at high frequency and in optical brightness. For example, as recently reported by Dent (1965), 3C 273B has increased its radio emission at 8000 Mc/s by 40% in the last three years (figure 2), while it is known that its optical brightness has an irregular period of about 10 years (Smith 1963). Again, the light from 3C 48 has been shown to vary by 30% in one year, and according to Sholomitsky (1965, Int. Astr. Un. Commission 27, Inform. Bull. No. 83) the source CTA 102 has a radio period of 100 days at 940 Mc/s. If one supposes, therefore, that in times of order 1 year the associated disturbances are propagated from the centre of the quasi-stellar sources at a velocity less than c, say $10^{-2}c$, then the diameter of the region responsible for the variable part of the emission would be of order 10^{16} cm.

The fact that the quasi-stellar sources show a continuous optical spectrum with ultra-violet excess indicates that the synchrotron mechanism may be operating on electrons of very high energy. The fact that the spectrum also contains forbidden emission lines, and fluctuates in brightness, shows that they are unlikely to be assemblages of stars, although nearby stars may inject the necessary energy, for example, supernovae. Rather does it suggest that the actual emitters are massive gaseous objects of low surface density and high temperature. Evidently, their radio and optical powers would imply that 10^{60} erg or 10^{6} solar masses are available for emission over a time of order 10^{6} years. The theories that have been put forward to explain the quasi-stellar sources will be described in § 9.

2.5. *The background radiation in the Universe at radio and optical frequencies*

With rapid advances in technique the detection of radiation due to cosmic sources is now becoming a practical feasibility. This is most fortunate as it seems likely that comparison of these observations with theory can have the greatest importance for cosmology (§ 6).

In the radio range the minimum temperature in the sky is available at one or two frequencies. Evidently the sky temperature due to cosmic sources such as normal galaxies and radio galaxies has this minimum value as an upper limit, since contributions from a more or less isotropic emission in the Galaxy would be included. A galactic contribution might arise from the interaction of cosmic rays with interstellar hydrogen (bremsstrahlung), or from the synchrotron radiation of cosmic-ray electrons or any other general distribution of high-speed electrons.

At 178 Mc/s the minimum sky temperature is 80 °K (Turtle and Baldwin 1962). On the assumption that the Galaxy and other normal galaxies have a different spectrum of emission from the average radio galaxy, Turtle *et al.* (1962) have deduced from their observations that the temperature due to radio galaxies is about 28 °K at 178 Mc/s. However, from the counts of the sources it seems clear that normal galaxies must contribute by far the greater proportion of the background at this frequency. This is because the background depends on source space density times power, viz. nP, and therefore (2.3) and (2.14), together with the fact that $P = 10^{21}$ w sterad $(c/s)^{-1}$ for normal galaxies and $P = 4 \times 10^{25}$ w sterad^{-1} $(c/s)^{-1}$ for radio galaxies, show that this product is greater for the first category by a factor of 40. Therefore, it seems safer to take the upper limit of temperature due to extragalactic radiation as 80 °K at 178 Mc/s.

On the lower side of this frequency Kenderdine (1963) obtains a minimum sky temperature of 5400 °K at 38 Mc/s. On the higher side Pauliny-Toth and Shakeshaft (1962) working at 404 Mc/s find a minimum temperature of 15 °K. Very recently, Penzias and Wilson (1965) have found an isotropic sky temperature of 3·5 °K at 4080 Mc/s. In terms of energy density we have therefore as upper limits for the extragalactic background (egb)

$$\rho_{\mathrm{egb}}(\mathrm{radio}) \leqslant \begin{Bmatrix} 9{\cdot}7 \times 10^{-28}\ \mathrm{erg\,cm^{-3}\,(c/s)^{-1}} & \text{at} & 38\ \mathrm{Mc/s} \\ 3{\cdot}2 \times 10^{-28}\ \mathrm{erg\,cm^{-3}\,(c/s)^{-1}} & \text{at} & 178\ \mathrm{Mc/s} \\ 3{\cdot}1 \times 10^{-28}\ \mathrm{erg\,cm^{-3}\,(c/s)^{-1}} & \text{at} & 404\ \mathrm{Mc/s} \\ 7{\cdot}5 \times 10^{-27}\ \mathrm{erg\,cm^{-3}\,(c/s)^{-1}} & \text{at} & 4080\ \mathrm{Mc/s} \end{Bmatrix}. \tag{2.15}$$

Below 404 Mc/s the above variation of density with frequency is roughly consistent with a power law spectrum of index $x \simeq 0{\cdot}6$. Actually this appears to be a typical spectral index for normal galaxies in this range of frequency. It is known that the background intensity steadily increases to a maximum at about 3 Mc/s below which it drops sharply. It has been suggested that the drop is due to absorption by interstellar ionized hydrogen (Smith 1964, Alexander and Stone 1965).

On the other hand, the index evidently goes positive between 404 and 4080 Mc/s. This may suggest that another process dominates over synchrotron radiation at high frequency in the background sources. Alternatively, the important discovery by Penzias and Wilson may reveal a residual Planck distribution of primeval radiation in an exploding-type universe (§ 6.3).

The information on the background radiation at optical frequencies is rather meagre. By an extrapolation of the local counts and apparent luminosities of galaxies de Vaucouleurs (1949) obtained a density at photographic (pg) frequencies

$$\rho_{\mathrm{egb}}(\mathrm{pg}) = 10^{-29}\ \mathrm{erg\,cm^{-3}\,(c/s)^{-1}}. \tag{2.16}$$

This would correspond to a total optical background due to external galaxies of approximately

$$\rho_{\mathrm{egb}}(\text{total optical}) \simeq 3{\cdot}4 \times 10^{-15}\ \mathrm{erg\,cm^{-3}}. \tag{2.17}$$

As reported by Whitrow and Yallop (1965) the total optical sky background due to external galaxies plus scattered starlight is, according to Sandage, equivalent to 4 tenth-magnitude stars per square degree. On conversion this gives

$$\rho_{\mathrm{egb}}(\text{total optical}) < 1{\cdot}37 \times 10^{-14}\ \mathrm{erg\,cm^{-3}}. \tag{2.18}$$

2.6. *The* X*-ray and* γ*-ray background and cosmic rays*

Because of absorption by the Earth's atmosphere X rays cannot be received at the surface of the Earth from outer space. However, in the early 1950's X-ray emission from the Sun was registered by Geiger counters in rockets flying above the atmosphere. This work was carried out by H. Friedmann of the Naval Research Laboratory, Washington, and shows that the solar hard X-ray emission (at a few kev) is about 10^{23} erg sec^{-1}. Since 1961 X rays of energy in the range 1–50 kev have been detected by rockets from 10 discrete sources outside the solar system. These sources are registered above an apparently isotropic background of X rays.

The first source, a particularly strong one in Scorpius, was discovered by Giacconi *et al.* (1962) of the Massachusetts Institute of Technology. The X rays were detected in the wavelength range 1–10 Å or energy range 1–10 kev. The flux of energy received from this source, about 10^{-27} erg cm^{-2} sec^{-1} (c/s)$^{-1}$, is comparable with that from the Sun at the same frequency. Remarkably, however, no luminous object is visible in the neighbourhood.

Only one X-ray source has been identified optically and this is the Crab nebula, the supernova of A.D. 1054. This has a flux one-eighth as strong as the Scorpius source in the range 1–10 kev but is the stronger near 50 kev. By lunar occultation of the source Friedmann has been able to show that the source is a region of about 1 light year in diameter within the nebula. This appears to rule out the conjecture that X-ray sources are so-called 'neutron stars'. A neutron star is a theoretical state of matter under enormously high pressure, such as might be obtained in the collapsed core of a supernova. The compression would be such as to confine a solar mass within a sphere of a few kilometres' radius. Another suggestion is that the X-ray emission is due to synchrotron radiation from electrons of extremely high energy (of order 30 GeV), such as might result from a supernova explosion. The Scorpius source may be an extinct supernova. An objection to this theory is that the electron energies would decay below the necessary threshold in a much shorter time than the age of the Crab nebula. An alternative theory is that the necessary injection of electron energy is maintained by the decay of radioactive elements produced in the original supernova explosion (see § 8.2.1).

Of great cosmological interest is the fact that an isotropic background of X rays exists. The analysis of the spectrum of the X-ray background by Giacconi *et al.* (1965) indicates that the total X-ray intensity is about 2×10^{-8} erg cm^{-2} sec^{-1} in the range 1–40 kev with a peak at 10 kev (figure 3). Although the possibility of its origin in the Galaxy is by no means ruled out, the above-mentioned authors are able to rule out a thermal origin for the spectra of individual sources in the Galaxy. There is the possibility of emission from optically thin regions of the Galaxy by π^0 meson decay generated by cosmic-ray protons, or by the bremsstrahlung or synchrotron radiation of cosmic-ray electrons. However, all these mechanisms appear to be quite inadequate for the observed background (§ 6.4). An extragalactic origin is more than possible, however. Since it is estimated that the total emission from the discrete sources of the Galaxy is 3×10^{37} erg sec^{-1} (Clark 1965, unpublished), a similar emission from all normal galaxies would give a background intensity not far below the observed order. But is is conceivable that abnormal galaxies such as radio sources and quasars play a large part, and it is also possible that the nature of the intergalactic medium might lead to an emission of X rays (see § 6.4). If we *assume an extragalactic origin* for the observed X-ray background we can write for the energy density

$$\rho_{\text{egb}}(\text{X rays}) = 2{\cdot}6 \times 10^{-18} \text{ erg cm}^{-3}. \qquad (2.19)$$

At γ-ray energies an isotropic background of photons of energy exceeding 100 MeV has been observed by Kraushaar and Clark (1962) and Kraushaar *et al.* (1965). These experiments were made using the satellite Explorer XI. The authors report that the isotropic flux of γ rays, in the range 100 MeV and upwards, has an intensity

$$I = (3{\cdot}3 \pm 1{\cdot}2) \times 10^{-4} \text{ photons/cm}^2 \text{ sec sterad}. \qquad (2.20)$$

Since no anisotropy was detected owing to sources in the Galaxy, this result was interpreted as an upper limit only for possible γ rays of cosmological origin. In terms of energy density this would give

$$\rho_{egb}(\gamma \text{ rays}) \leqslant 2{\cdot}2 \times 10^{-17} \text{ erg cm}^{-3}. \tag{2.21}$$

A possible source in the Galaxy for γ rays is once again the ubiquitous cosmic rays. Their action could be via π^0 meson decay, bremsstrahlung or inverse Compton effect, although the first process is estimated to give a flux 10–20 times too small to match the total background (Kraushaar *et al.* 1965). It might also be due to these

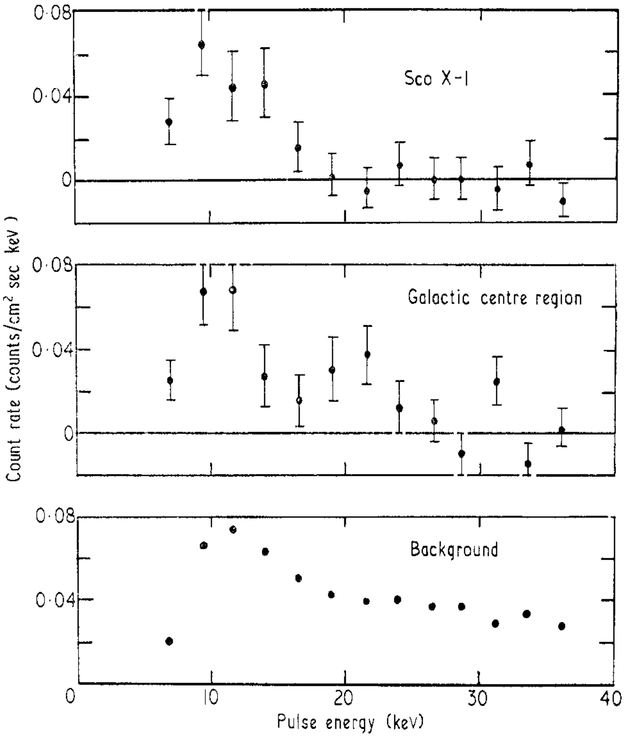

Figure 3. The spectra of X-ray sources and the X-ray background are given. The differential count rates of X rays, measured in counts/cm^2 sec kev, as a function of pulse height (in kev) are given for (*a*) the source Scorpius X-1, (*b*) the Galactic centre region and (*c*) the isotropic background. In (*a*) and (*b*) the background contribution is subtracted from the count rate. The detector had a lower threshold of about 8 kev. (From Giacconi *et al.* 1965.)

mechanisms arising from extragalactic cosmic rays and intergalactic hydrogen if both exist in sufficient quantities. Or it could originate in the emission of exploding galaxies, strong radio sources or quasars (§ 6.4).

Finally, the cosmic distribution of matter may include the cosmic rays, as already mentioned. About 70% of these are protons, nearly 30% heavier nuclei and 0·2% electrons. As is well known and well established, the cosmic rays have no detectable anisotropy and have a total energy flux close to or exceeding that of starlight. The space density is about 10^{-10} particles/cm^3, and the total energy density gives an upper limit to the possible *extragalactic* component

$$\rho_{egb}(\text{cosmic rays}) < 10^{-12} \text{ erg cm}^{-3}. \tag{2.22}$$

The mean total energy of the particles is 10^{10} ev, the complete spectrum ranging from 10^9 to 10^{20} ev. It seems that in the range 10^9–10^{15} ev the cosmic rays may be satisfactorily accounted for by ejections from ordinary stars (low energy), by accelerating mechanisms in the gaseous envelopes of supernovae and possibly by a gigantic explosion at the centre of the Galaxy in past ages. No generally accepted theory has yet been proposed to account for the origin in the Galaxy of the particles of highest energy. However, the recent cosmic discoveries already mentioned may well lead to establishing their origin as extragalactic (§ 6.5).

3. Introductory theoretical background

3.1. *Cosmological space–time*

The first real impetus to a theoretical discussion of cosmology came from Einstein's general theory of relativity. Even today general relativity provides the best framework for cosmology. It is not difficult to see why this should be so. Cosmology has to deal with the large-scale distribution of matter in the Universe and, as such, gravitational interactions form a basic part of the subject. Newton's theory of gravitation, with all its successes, has one drawback. It is a theory of *instantaneous* action at a distance and therefore inconsistent with Einstein's special theory of relativity which is now so well established in local physics. In cosmology great distances and large velocities are involved, and their discussion under Newtonian concepts is naturally suspect. What is needed is a theory which can count on the successes of Newton's theory and which at the same time is consistent with the special theory of relativity. General relativity meets these requirements. We shall be concerned here mainly with general relativity, except for a brief mention of Newtonian cosmology at the end of § 4.

According to general relativity the geometry of space–time is Riemannian. We shall use coordinates x^i $(i = 1, 2, 3, 4)$ and metric tensor g_{ik} to describe the line element in the general form

$$ds^2 = g_{ik}\,dx^i\,dx^k. \tag{3.1}$$

The g_{ik} are in general functions of x^i. In the infinitesimal neighbourhood of any point in space–time the line element can be reduced to that of special relativity:

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2 \tag{3.2}$$

by a coordinate transformation $x^i \rightarrow (x, y, z, t)$. In (3.2) dx, dy, dz are in length units, dt in time units, and c is the velocity of light. Therefore ds^2 has the dimensions of $[\text{length}]^2$. We shall stick to this convention throughout the article.

The quantities g_{ik} are determined from the distribution of matter and energy by Einstein's field equations

$$R_{ik} - \tfrac{1}{2} g_{ik} R = -\frac{8\pi G}{c^4} T_{ik} \tag{3.3}$$

where R_{ik} is the Ricci tensor, R the scalar curvature and T_{ik} the energy–momentum tensor of matter and any other sources of energy present in the Universe. G is the Newtonian constant of gravitation. The T_{ik} therefore have the dimensions of $(\text{mass})\,(\text{time})^{-2}\,(\text{length})^{-1}$.

The equations (3.3) are second-order non-linear partial differential equations and are therefore very difficult to handle. For this reason even after more than fifty years since the origin of general relativity very few *exact* solutions of these equations are available today. Such solutions, as are available, are possible because some simplification is introduced by the physical nature of the problem. Fortunately, such a simplification is possible in cosmology.

Observations described in the previous section indicate that there is a remarkable degree of regularity in the large-scale structure of the Universe. The distant galaxies appear to be receding from us almost radially. Also the distribution of galaxies in space appears to be uniform. These somewhat vague observational results have led to the well-known Weyl (1923) postulate and the cosmological principle, which together simplify the cosmological problem considerably. These are briefly described below.

(i) *Weyl postulate.* This states that the world lines of galaxies form a bundle of geodesics diverging from a point in the finite or infinite past.

This postulate enables us to simplify the line element (3.1) by a special choice of coordinates in the following way. Through each point in space–time there passes a unique geodesic of the family described in the postulate. We can therefore use x^1, x^2, x^3 as three parameters to label these geodesics. And along each geodesic we choose t as the parameter which measures the *proper* time of an observer following the geodesic. Such an observer has x^μ = constant ($\mu = 1, 2, 3$), and will be called a fundamental observer. The line element can then be written in the form

$$ds^2 = c^2\,dt^2 + 2g_{\mu 4}\,dt\,dx^\mu + g_{\mu\nu}\,dx^\mu\,dx^\nu. \tag{3.4}$$

(The Greek indices run from 1 to 3 and the Latin ones from 1 to 4.) The requirement that x^μ = constant is a geodesic leads to

$$\frac{\partial g_{\mu 4}}{\partial t} = 0. \tag{3.5}$$

Thus $g_{\mu 4}$ depend on x^μ only. The parameter t is called the cosmic time.

(ii) *Cosmological principle.* This states that at a given cosmic time the Universe presents the same large-scale view to all fundamental observers.

Physically, if we identify fundamental observers with galaxies, then at any cosmic time the Universe should look the same when viewed from any galaxy. Thus the phenomenon of the expansion of the Universe, which we observe from our own Galaxy, would also be observed from any other galaxy.

The above statement of the cosmological principle, in its most general form, is equivalent to saying that the Universe is homogeneous *space-wise*. That is, if $P(x^\mu, t)$ and $P'(x'^\mu, t)$ are any two points in the sub-space t = constant, it is possible to make a continuous transformation of the Universe into itself such that P goes to P'.

Different cosmologies have interpreted the cosmological principle in different ways. Stated in the general form above, the Universe is homogeneous but, in general, anisotropic. That is, it may not present the same large-scale view in all directions to a fundamental observer. As mentioned in the previous section, there is no strong evidence for anisotropy from the present observations. We shall consider anisotropic models at a later stage (§ 9.4).

If we insist on spatial isotropy as well as homogeneity, the line element (3.4) is further simplified. It can be shown that $g_{\mu 4} = 0$ and that the spaces t = constant admit a six-parameter group of motions and are therefore spaces of constant curvature. The line element then reduces to the form

$$ds^2 = c^2\,dt^2 - R^2(t)\left\{\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\right\}, \quad k = 0, \pm 1 \tag{3.6}$$

where the coordinate r has value zero at an arbitrarily chosen fundamental observer. The surface r = constant has the geometry of the surface of a sphere and θ, ϕ are the usual spherical polar coordinates. The Weyl geodesics are given by r = constant, θ = constant and ϕ = constant.

The line element (3.6) was first used by Friedmann (1924). It was deduced rigorously by Robertson (1935, 1936) and Walker (1936) independently. The line element (3.6) is sometimes written in a different form:

$$ds^2 = c^2\,dt^2 - \frac{R^2(t)}{(1+\frac{1}{4}kr^2)^2}\{dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\}, \quad k = 0, \pm 1. \tag{3.7}$$

The r coordinate used in (3.7) is not the same as that used in (3.6); t, θ, ϕ have the same meaning. Both forms are identical in the case $k = 0$.

It is possible to take the cosmological principle a stage further by requiring that the Universe present the same large-scale view to all fundamental observers *at all times*. This is the perfect cosmological principle of Bondi and Gold (1948), which will be discussed in § 5. This leads to a line element of the form

$$ds^2 = c^2\,dt^2 - e^{2Ht}\{dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\} \tag{3.8}$$

where H is a constant. This is the line element first considered by de Sitter (1917) and is also used to describe the steady-state cosmology.

The unknowns in (3.6) or (3.7), i.e. the function $R(t)$ and the value of k, can be determined from Einstein's equations, provided T_{ik} is known. What exactly does T_{ik} stand for? Cosmological theories differ in their interpretation of T_{ik} in the actual Universe. From a historical point of view T_{ik} stands for the energy–momentum tensor of matter alone. In this form the Einstein equations can be reduced to the Newtonian ones provided gravitational fields are weak. Divergence of (3.3) leads to the conservation law

$$T^{ik}{}_{;k} = 0. \tag{3.9}$$

Thus the energy–momentum of matter in the Universe is conserved. Cosmologies based on this hypothesis—which in general lead to an evolving universe—are described in the next section.

A wider interpretation of T_{ik} requires it to include the energy tensors not only of matter, but of all other quantities. Thus if an electromagnetic field is present, its energy tensor will have to be included in T_{ik}. It then becomes possible to imagine T_{ik} as made up of matter tensor and another tensor, the two together being conserved according to (3.9); but they do not have to be conserved separately. Such a device was used by Hoyle (1948) to account for continuous creation of matter in the steady-state cosmology. This will be described in § 5.

3.2. *Red shift of spectral lines*

The phenomenon of red shift observed by Hubble admits of a very natural explanation in the Robertson–Walker space–time. Consider a galaxy at (r_1, θ, ϕ) which emits a light signal at time $t = t_1$, so as to reach the Galaxy at $r = 0$ at time $t = t_0\ (> t_1)$. The path of a light ray is given by the equation $ds = 0$, i.e. using (3.6), by

$$\frac{dr}{dt} = -\frac{c}{R(t)}(1-kr^2)^{1/2} \tag{3.10}$$

which leads to

$$\int_0^{r_1} \frac{dr}{(1-kr^2)^{1/2}} = c\int_{t_1}^{t_0} \frac{dt}{R(t)}. \tag{3.11}$$

A light signal emitted at $t_1 + \Delta t_1$ will reach $r = 0$ at $t_0 + \Delta t_0$, where

$$\int_0^{r_1} \frac{dr}{(1-kr^2)^{1/2}} = c\int_{t_1+\Delta t_1}^{t_0+\Delta t_0} \frac{dt}{R(t)}. \tag{3.12}$$

For small Δt_1, Δt_0, we have

$$\frac{\Delta t_0}{R(t_0)} = \frac{\Delta t_1}{R(t_1)}. \tag{3.13}$$

Now suppose that light of a specific wavelength λ is being emitted at $r = r_1$, and suppose Δt_1 stands for the period of one oscillation corresponding to the wavelength λ, i.e.

$$\lambda = c\Delta t_1. \tag{3.14}$$

Similarly the wavelength of light received at $r = r_0$ is given by

$$\lambda + \Delta\lambda = c\Delta t_0 \tag{3.15}$$

since Δt_0 is the period of oscillation of the light received. Hence

$$\frac{\lambda+\Delta\lambda}{\lambda} = 1 + z = \frac{\Delta t_0}{\Delta t_1} = \frac{R(t_0)}{R(t_1)} \tag{3.16}$$

where $z = \Delta\lambda/\lambda$ is called the spectral shift. 'Red shift' as observed by Hubble corresponds to $z > 0$, i.e. $R(t_0) > R(t_1)$. Thus the nature of spectral shift is simply related to the behaviour of the function $R(t)$ of the Robertson–Walker line element.

An element of proper three-volume at time t is given by

$$dV = R^3(t)\frac{r^2}{(1-kr^2)^{1/2}}\sin\theta\, dr\, d\theta\, d\phi. \tag{3.17}$$

Thus $R(t_0) > R(t_1)$, i.e. R increasing with time, indicates that the proper volume of any region bounded by galaxies increases with time. This is another way of saying that the Universe is expanding. A 'blue shift' would similarly indicate a contracting universe.

3.3. *Luminosity and flux density of radiating objects*

Consider a source of radiation at (r_1, θ, ϕ) of luminosity L, assumed constant for the time being. At time t_0, the radiation emitted at t_1 will be crossing a surface area $4\pi r_1{}^2 R^2(t_0)$. If there were no red shift the rate at which radiation crossed this area would be simply $L/4\pi r_1^2 R^2(t_0)$. The rate is actually reduced for two reasons. Because of the red shift each quantum of energy at $t = t_1$ is reduced by a factor $1+z$ by the time it reaches the observer at $r = 0$. Further, the 'rate' of flow of energy has to be measured with respect to the observer's time t_0, not t_1. The two proper times are related by (3.16), which brings in another factor $1+z$. Thus the actual energy flow per unit area per unit time at $r = 0$ is

$$l = \frac{L}{4\pi r_1^2 R^2(t_0)(1+z)^2} = \frac{L}{4\pi D^2} \tag{3.18}$$

where

$$D = r_1 R(t_0)(1+z) \tag{3.19}$$

is known as the 'luminosity distance' of the source.

In practice the radiation measured is over a limited frequency range and so (3.18) is not immediately applicable. Suppose $I(\nu)\,d\nu$ denotes the fractional energy spectrum at the source, normalized so that

$$\int_0^\infty I(\nu)\,d\nu = 1. \tag{3.20}$$

The frequency ν at the source corresponds to a frequency $\nu_0 = \nu(1+z)^{-1}$ at the observer. The observed energy spectrum function is therefore given by $(1+z)\,I\{\nu_0(1+z)\}$. Thus if Δl denotes the energy of radiation in the frequency range $\Delta\nu_0$ crossing unit area per unit time at $r = 0$. then

$$\Delta l = \frac{LI\{\nu_0(1+z)\}}{4\pi r_1^2 R^2(t_0)(1+z)}\Delta\nu_0. \tag{3.21}$$

For radio astronomy $I(\nu)$ can be approximated by (constant) $\times\, \nu^{-x}$, over the relevant frequency range, where x is the 'spectral index'. In that case

$$\Delta l = \frac{LI(\nu_0)\,\Delta\nu_0}{4\pi r_1^2 R^2(t_0)(1+z)^{1+x}}. \tag{3.22}$$

If $\Delta\nu_0 = 1$, then Δl becomes the 'flux density' of a radio source, usually denoted by $S(\nu_0)$. From (3.22), we have

$$S(\nu_0) = \frac{LI(\nu_0)}{4\pi r_1^2 R^2(t_0)(1+z)^{1+x}}. \tag{3.23}$$

The corresponding expression in the case of line element (3.7) is

$$S(\nu_0) = \frac{LI(\nu_0)(1+\frac{1}{4}kr_1^2)^2}{4\pi r_1^2 R^2(t_0)(1+z)^{1+x}}. \tag{3.23}$$

In optical astronomy $I(\nu) \propto \nu^{-2}$ approximately, and

$$\Delta l = \frac{LI(\nu_0)\,\Delta\nu_0}{4\pi r_1^2 R^2(t_0)(1+z)^3}, \quad \Delta l = \frac{LI(\nu_0)\,\Delta\nu_0(1+\frac{1}{4}kr_1^2)^2}{4\pi r_1^2 R^2(t_0)(1+z)^3} \tag{3.24}$$

for line elements (3.6) and (3.7), respectively.

It is customary to express l, Δl, etc. in terms of apparent magnitudes. Thus (see footnote in § 2.1) we write the apparent bolometric magnitude as

$$m_{\mathrm{bol}} = -2{\cdot}5\log l_{\mathrm{bol}} + \text{constant}. \tag{3.25}$$

Similarly visual (m_v), photographic (m_{pg}) and other magnitudes are defined over relevant frequency ranges.

The expression (3.18) has been used to resolve Olbers' paradox. In a static, infinite universe with a uniform distribution of radiating objects, assumed capable of shining for ever, or at least until thermodynamic equilibrium is reached, the radiation density at each point would be the same as that at the effective surface of a typical bright object. If absorption by dust, etc. is small, the sky should therefore be very bright—whether the Sun is present or not. This paradox raised for a somewhat academic universe in 1828 was atrophied when the expansion of the Universe became known. As seen in (3.18), the expansion introduces a factor $(1+z)^2$ in the denominator and so greatly reduces the radiation received from a distant object, and leads to a background radiation density of the order of that observed. This question of background radiation will be considered again in § 6.

3.4. *Number counts and angular diameters*

In this section we derive two important results which are implied in the discussion of observational tests in § 6.

(i) Consider the number of galaxies with coordinates $r \leqslant r_1$. To evaluate this number we proceed as follows. Let $n(t)$ denote the number of galaxies per unit *proper* volume at time t. Then in the evolutionary cosmologies

$$n(t)\,R^3(t) = \text{constant} = n_0\,R^3(t_0). \tag{3.26}$$

Now, an element of proper volume is given by (3.17). Hence the required number is given by

$$\begin{aligned} N(r_1) &= \int_0^{r_1}\int_0^{\pi}\int_0^{2\pi} n(t)\,R^3(t)\frac{r^2}{(1-kr^2)^{\frac{1}{2}}}\sin\theta\,dr\,d\theta\,d\phi \\ &= 4\pi n_0\,R^3(t_0)\int_0^{r_1}\frac{r^2\,dr}{(1-kr^2)^{1/2}}. \end{aligned} \tag{3.27}$$

Note that if we had used the line element (3.7) instead, we should get

$$N(r_1) = 4\pi n_0\,R^3(t_0)\int_0^{r_1}\frac{r^2\,dr}{(1+\frac{1}{4}kr^2)^3}. \tag{3.28}$$

In the steady-state cosmology $n(t) = \text{constant} = n_0$ (say). Then we have

$$N(r_1) = 4\pi n_0\int_0^{r_1} r^2\,R^3(t)\,dr \tag{3.29}$$

where $R(t) = \mathrm{e}^{Ht}$ corresponds to the value of R at the time of emission of light from the galaxy. If the light is to reach $r = 0$, we must have, by (3.11),

$$r = c\int_t^{t_0}\mathrm{e}^{-Ht}\,dt = \frac{c}{H}(\mathrm{e}^{-Ht}-\mathrm{e}^{-Ht_0})$$

that is,

$$\mathrm{e}^{-Ht} = \left(\frac{rH}{c}+\mathrm{e}^{-Ht_0}\right).$$

Without loss of generality we can choose $e^{Ht_0} = 1$. This means at $r = 0$, the coordinate r measures radial *proper* distance. We then get

$$N(r_1) = 4\pi n_0 \int_0^{r_1} \frac{r^2\,dr}{(1+rH/c)^3}. \tag{3.30}$$

(ii) Consider a spherical source at (r_1, θ, ϕ) with a linear diameter d. Light from such a source emitted at $t = t_1$ reaches the observer at $r = 0$ at $t = t_0$. Let $\Delta\theta$ be the angular diameter of the source as observed at $r = 0$. Then we have, in terms of line element (3.6),

$$d = R(t_1)\, r_1 \Delta\theta. \tag{3.31}$$

Therefore

$$\Delta\theta = \frac{d}{R(t_1)\, r_1} = \frac{d(1+z)^2}{D}. \tag{3.32}$$

The presence of $1+z$ in the numerator suggests that angular diameters need not get progressively small as D increases. Indeed, in some cosmologies, they even increase with D for sufficiently large D (cf. Sandage 1961 a).

3.5. *Mach's principle*

It can be shown that by a suitable transformation $r' = r'(r,t)$, $t' = t'(r,t)$, $r' \to 0$ as $t' \to 0$, the line element (3.6) can be transformed to

$$ds^2 = c^2\,dt'^2\left(1 - r'^2\frac{\ddot{R}}{R}\right) - \frac{dr'^2}{1 - kr'^2/R^2 - r'^2\dot{R}^2/R^2} - r'^2(d\theta^2 + \sin^2\theta\, d\phi^2) \tag{3.33}$$

for a sufficiently small r'. This line element approaches that of special relativity as $r' \to 0$. Rotation relative to such a coordinate frame would produce inertial forces. In this frame the coordinates θ, ϕ of distant galaxies (moving geodesically) are constant. We can therefore say that the line element (3.6) brings out the observed result that the local inertial frame is the same as one in which the distant parts of the Universe are non-rotating. This observation was responsible for Mach's (1883) principle.

Mach argued that the validity of Newton's laws of motion in a reference frame determined by a cosmological observation suggests the presence of a long-range interaction between the local and distant parts of the Universe. Since the very concept of inertia depends on the laws of motion, this long-range interaction may be responsible for inertia.

Mach's principle has not been formulated in a unique way. Different physicists have taken different views. Some take it seriously; others dismiss the observation on which it is based as accident. Einstein belonged to the former group, and it was his hope that this principle could somehow be incorporated in general relativity. This hope has not been fully realized so far. In the case of the Robertson–Walker line element the observation regarding the inertial frame is explained. The function $R(t)$ and the parameter k are, in principle, determinate in terms of the distribution of matter and energy in the Universe. This is in agreement with Machian ideas. But the Robertson–Walker line element is not the only one permitted by general relativity. Something in the nature of a counter example was given by Gödel (1949) with the following line element:

$$ds^2 = c^2\,dt^2 + 2e^{x^1}\,dt\,dx^2 - (dx^1)^2 + \tfrac{1}{2}e^{2x^1}(dx^2)^2 - (dx^3)^2 \tag{3.34}$$

where the galaxies have their world lines given by x^μ = constant. This belongs to the class of line elements (3.4) and is homogeneous but anisotropic. In Gödel's model the matter in the Universe is found to be rotating, at least locally, relative to local inertial frames. That general relativity permits models like this may be an indication that Mach's principle is not fully incorporated in it. Gödel's model was the first of a series of models which have been discussed subsequently by others (Raychaudhuri 1955, Heckmann and Schücking 1958, etc.). These models exhibit shear and rotation; those involving rotation appear not to satisfy Mach's principle. These will be discussed in § 9.4.

As noted above, the observed coincidence responsible for Mach's principle can be accounted for by the Robertson–Walker line element. Further Machian implications of the fact that the Universe is apparently described by such a line element will be considered in § 5.

4. Evolutionary or relativistic cosmology

4.1. *Exploding-type universes of Robertson–Walker metric*

The space–time metric of the homogeneous isotropic world models which are based on the cosmological principle and Weyl's postulate is given in one form by equation (3.7). It is well known that the equations of general relativity, neglecting the cosmological constant Λ, yield the following ordinary differential equations when applied to this metric:

$$\frac{3\dot{R}^2}{R^2} + \frac{3kc^2}{R^2} = \kappa\rho \tag{4.1}$$

$$\frac{2\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{kc^2}{R^2} = -\kappa p \tag{4.2}$$

where the dot notation indicates differentiation with respect to t.

Here $\rho(t)$ is the mean or smoothed-out energy density in the Universe, depending on cosmic time t only in accordance with the assumed homogeneity of space. All forms of mean energy are included in ρ, material rest energy and kinetic energy in the form of galaxies, intergalactic gas (if any), intergalactic cosmic rays (if any), together with radiation (including a possible neutrino component) and magnetic energy. Similarly $p(t)$ is the mean cosmic pressure arising from all these sources at epoch t. The constant κ is the Einstein gravitational constant, equal to

$$1{\cdot}86 \times 10^{-27} \text{ c.g.s. units.}$$

The cosmological constant is taken to be zero in (4.1) and (4.2), since this assumption is usually made in current applications of relativistic cosmology. The constant must, of course, be very small since it would otherwise have an observed effect on planetary motion. There are, in fact, good arguments, such as an appeal to Mach's principle, for rejecting it completely. The Lemaitre and Eddington–Lemaitre universes will, in consequence, not be considered in this review. In passing it may be noted that the observational data on the red shift appear to rule out these models in any case, since the data indicate a decelerating expansion rather than an accelerating one (§ 6).

By differentiating (4.1) and combining with (4.2) we get a cosmic energy equation in the form

$$\frac{d}{dt}(\rho R^3)+p\frac{d}{dt}(R^3) = 0. \tag{4.3}$$

Since the function $R(t)$ is proportional to the distance apart at time t of any two cosmic objects co-moving with the coordinate mesh, it follows that this equation may be written

$$\frac{d}{dt}(\rho v)+p\frac{d}{dt}(v) = 0 \tag{4.4}$$

where $v(t)$ is any volume of space of co-moving boundaries. This equation expresses the principle of conservation of total mass–energy, allowing for adiabatic energy changes, which is fundamental to general relativity.

It follows that if we extrapolate the present expansion backwards in time then, according to (4.3) or (4.4), ρR^3 or ρv must steadily increase or remain constant if p remains non-negative. At present basic physics provides no encouragement for a belief that the pressure of matter at very high density can ever become negative. (On the other hand, an assumption equivalent to this possibility is necessary to account for the creation of matter in the steady-state theory on the basis of general relativity (McCrea 1951). In the Hoyle–Narlikar version (§ 5) of the steady-state theory the conventional interpretation of the energy tensor is retained but a creation field term is added to the Einstein field equations.) A more detailed analysis shows that according to the Einstein equations $\rho \to \infty$ as $R(t)\to 0$. Mathematically, therefore, a reversal of the present expansion leads to an indefinitely high density. This is the much discussed singularity of isotropic world models in relativistic cosmology.

It is easily seen that this singularity occurs at a finite time in the past. For evidently (4.1) and (4.2) combine to give

$$\frac{\ddot{R}}{R} = -\tfrac{1}{6}\kappa(\rho+3p) \tag{4.5}$$

which indicates that, for the conventional physical assumptions ($\rho, p \geqslant 0$), the expansion of the Universe undergoes *deceleration* in relativistic cosmology. Accordingly, if we put

$$\frac{\dot{R}_0}{R_0} = H = T^{-1} \tag{4.6}$$

where R_0 is the value of $R(t)$ for the present epoch t_0, then the constant H represents the present rate of expansion per unit distance (Hubble's 'constant'). Hence, since $\ddot{R}<0$, the lapse of time since the singularity must be less than $H^{-1} = T$. If we take $t = 0$ at the singularity then in relativistic cosmology the *age of the universe*, or at least the time since the singularity, is t_0 where

$$t_0 < H^{-1} \simeq 10^{10} \text{ year} \tag{4.7}$$

(cf. § 2).

Again, just as $\dot{R}/R$ is the rate of expansion per unit distance, $\ddot{R}/R$ is the rate of acceleration (negative) per unit distance. Moreover, since $\rho \to \infty$ as the singularity is approached, (4.5) shows that $\ddot{R}/R$ and hence $\dot{R}/R \to \infty$ as $R \to 0$. The interpretation then is that at $t = 0$ the contents of the model exploded with infinite velocity from a state of infinite density. How far this conclusion can be applied to the actual Universe, which is not perfectly homogeneous or isotropic, has lately received much discussion and this matter will be considered further in § 9.

4.2. *The behaviour of exploding models at large t*

It is convenient to divide the exploding models into three categories made possible by cosmic conditions at the present epoch. This depends on the fact that the cosmic pressure p, arising from the sources already mentioned, is evidently small at the present epoch compared with ρ. In fact, reference to equations (2.6), (2.17), (2.19), (2.21) and (2.22) indicates that the ratio is certainly less than 10^{-2} and more probably less than 10^{-5}. In this case equation (4.2), setting $p = 0$ as a first approximation, has the first integral

$$\dot{R}^2 = \frac{A}{R} - kc^2 \quad (A > 0). \tag{4.8}$$

The three categories of model at the stage of negligible pressure evidently depend on the curvature of space, i.e. on the value of k:

(i) First, let us consider the case where $k = 1$ (positive curvature, elliptic closed space). In this case $\dot{R}$ vanishes when

$$R_{\max} = A/c^2 \tag{4.9}$$

and the model oscillates regularly between a singularity and this maximum radius.

(ii) Next, consider the case where $k = 0$ (flat, Euclidean space). Here the complete solution for $R(t)$ is

$$R(t) = (\tfrac{3}{2}A^{1/2}\, t)^{2/3} \tag{4.10}$$

taking $t = 0$ when $R(t) = 0$, as before. This is the well-known Einstein–de Sitter universe. We note that $\dot{R} \to 0$ as $t \to \infty$ so that the expanding sub-stratum of matter ultimately comes to rest in the infinite future.

(iii) Finally, we shall consider the case where $k = -1$ (negative curvature, hyperbolic open space). In this case $\dot{R} > 0$ always and in fact

$$R(t) \to ct, \quad t \to \infty. \tag{4.11}$$

The ultimate state of the hyperbolic models is therefore that of the well-known Milne universe, the sub-stratum expanding with constant velocities.

The different possibilities for $R(t)$, assuming $p = 0$, are shown in figure 4.

For the present epoch t_0, the neglect of p provides from (4.1) and (4.2) the relations

$$3H^2 + \frac{3kc^2}{R_0^2} = \kappa\rho_0 \tag{4.12}$$

$$2q - 1 = \frac{kc^2}{H^2 R_0^2} \tag{4.13}$$

where

$$q = -\frac{\ddot{R}_0/R_0}{H^2} \tag{4.14}$$

is the dimensionless *deceleration parameter* of the model. Another relation is provided by combining (4.12) and (4.13), or directly from (4.5):

$$q = \kappa\rho_0/6H^2. \tag{4.15}$$

We note that, corresponding to models (i), (ii) and (iii), we have

$$\left.\begin{array}{l}\text{(i) } q > \tfrac{1}{2};\ \rho_0 > 3H^2/\kappa \\ \text{(ii) } q = \tfrac{1}{2};\ \rho_0 = 3H^2/\kappa \\ \text{(iii) } \tfrac{1}{2} > q > 0;\ \rho_0 < 3H^2/\kappa\end{array}\right\}. \tag{4.16}$$

These relations are useful in observational tests (§ 6).

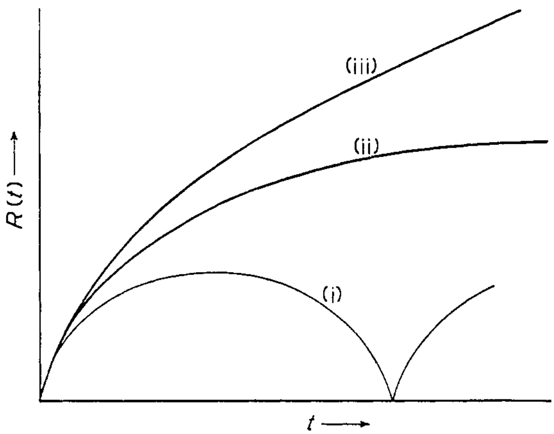

Figure 4. The graphs of the universal expansion in three categories of exploding models in relativistic cosmology are presented. The ordinate gives the expansion factor $R(t)$ as a function of cosmic time t in the cases of : (i) the oscillating model of positive spatial curvature and deceleration parameter $q > \frac{1}{2}$, (ii) the ever-expanding Einstein–de Sitter model of zero curvature and $q = \frac{1}{2}$ (here $\dot{R} \to 0$ as $t \to \infty$ so that the galaxies come to relative rest after an infinite time), and (iii) the ever-expanding hyperbolic model of negative curvature and $\frac{1}{2} > q > 0$ (here $\dot{R} \to$ constant as $t \to \infty$, so that galaxies ultimately recede with constant relative velocities).

For the case $p = 0$ it is possible to find the present age t_0 of the Universe in terms of H and q only. The value of the result depends on the assumption that the early period when p is not negligible will be of relatively short duration, owing to the high rate of expansion then obtained. For this purpose we note that the constant A of (4.8) and also R_0 can be expressed in terms of H and q only. Thus, expressing (4.8) for the epoch t_0 and dividing by R_0^2, yields

$$A = 2qH^2R_0^3. \tag{4.17}$$

Then R_0, and hence A, are related to q and H by means of (4.13). Consequently, when (4.8) is integrated for t as a function of R one finds that for the epoch t_0 the categories (i), (ii) and (iii) yield, respectively,

$$\text{(i)}\ t_0 = \left\{\frac{1}{1-2q}+\frac{2q}{(2q-1)^{3/2}}\sin^{-1}\left(\frac{2q-1}{2q}\right)^{1/2}\right\}H^{-1} \tag{4.18}$$

$$\text{(ii)}\ t_0 = \tfrac{2}{3}H^{-1} \tag{4.19}$$

$$\text{(iii)}\ t_0 = \left\{\frac{1}{1-2q}-\frac{q}{(1-2q)^{3/2}}\log\left(\frac{1-q}{q}+\frac{(1-2q)^{1/2}}{q}\right)\right\}H^{-1}. \tag{4.20}$$

4.3. *Behaviour of exploding models at small t, assuming that $p \to \frac{1}{3}\rho$ as $t \to 0$*

If the density of matter becomes very high it is reasonable to assume that there is also a very high temperature. In this case, according to nuclear physics, it is appropriate to postulate the presence of radiation with $p_r = \frac{1}{3}\rho_r$ as equation of state (suffix r referring to radiation). Indeed, as the singularity is approached on reversal of time we should expect material energy to be converted to radiation energy, or, in the forward direction of time,

$$\frac{d}{dt}(\rho_r R^3)+p_r\frac{d}{dt}(R^3) \leqslant 0 \tag{4.21}$$

and

$$\frac{d}{dt}(\rho_m R^3)+p_m\frac{d}{dt}(R^3) \geqslant 0 \tag{4.22}$$

where suffix m refers to material energy. Since the first relation can be written $(d/dt)(\rho_r R^4) \leqslant 0$, it follows that $\rho_r \to \infty$ at least like R^{-4} as $R \to 0$. On the other hand, $p_m < \frac{1}{3}\rho_m$ for material energy, so that equation (4.22) indicates that $\rho_m \to \infty$ at a rate definitely less than R^{-4}. Consequently, as pointed out by Gamow many years ago, radiation energy must dominate over material energy as $t \to 0$. That is, $\rho_r \to \rho \to \infty$ as $t \to 0$, although perhaps $\rho_m \neq 0$ for any t.

For sufficiently small t we can therefore put $p \simeq \frac{1}{3}\rho$ in (4.1), (4.2) to obtain

$$\frac{\ddot{R}}{R}+\frac{\dot{R}^2}{R^2}+\frac{kc^2}{R^2} = 0 \tag{4.23}$$

which has the first integral

$$\dot{R}^2 = \frac{A'}{R^2}-kc^2 \qquad (A'>0). \tag{4.24}$$

Evidently when R is small the curvature term is negligible and the model's expansion behaves like

$$R(t) = B't^{1/2} \quad (B' = \text{constant}) \tag{4.25}$$

for all values of k. The corresponding total energy density (predominantly radiation as we have assumed) is given by (4.1):

with

$$\left.\begin{aligned}\rho &= \frac{3}{4\kappa t^2}\\ p &= \frac{1}{4\kappa t^2}\end{aligned}\right\}. \tag{4.26}$$

For sufficiently large values of t, when the matter present in such a 'radiation universe' interacts negligibly with the radiation and has a pressure p_m small compared with ρ_m, we shall have, by (4.22), $(d/dt)(\rho_m R^3) = 0$ so that, by (4.25),

$$\rho_m = \frac{C}{t^{3/2}} \tag{4.27}$$

where C is a constant. This result is important in the theory of the origin of the elements according to Gamow *et al.* (§ 8).

4.4. *The background radiation in isotropic world models*

The measured intensity of the general background of radiation can have great importance for cosmology since it may reveal significant aspects of the past history of the Universe. Theoretically, the present intensity must depend on the rate of emission of radiation from galaxies during past ages, its rate of absorption by galaxies and possibly intergalactic gas, and, finally, on how much radiation existed in the Universe initially (cf. § 4.3).

In the past it has been thought necessary to calculate the background radiation density by integrating the contributions from all parts of space, allowing for red shift and time retardation effects (McVittie 1962, Whitrow and Yallop 1964, Bonnor 1964). However, the same result (Davidson 1962 b, Harrison 1964, Davidson 1965) is obtained if one merely time-integrates the local equation of total energy for radiation in the form

$$\frac{1}{R^3}\left\{\frac{d}{dt}(\rho_r R^3)+p_r\frac{d(R^3)}{dt}\right\} = E_r \tag{4.28}$$

where E_r is the excess of rate of emission of radiation over rate of absorption in unit volume. For the latter we can take the expression ρ_r/χ where $\chi(t)$ is a frequency-averaged mean free time between the emission and absorption of a photon. Hence, if $S(t)$ is the rate of emission from unit volume, due to galaxies and possible intergalactic sources, equation (4.28) becomes

$$\dot{\rho}_r + 4\dot{R}\rho_r = S - \frac{\rho_r}{\chi} \tag{4.29}$$

or

$$\frac{d}{dt}\left(\rho_r R^4 \exp\int^t \frac{dt}{\chi}\right) = SR^4 \exp\int^t \frac{dt}{\chi}.$$

If t_* is the local time on the observational horizon of the model, integrating this equation between $t = t_*$ and $t = t_0$, we find

$$\rho_r(t_0) = \frac{1}{R_0^4}\int_{t_*}^{t_0} SR^4 \exp\left(-\int_t^{t_0}\frac{dt}{\chi}\right)dt + \frac{\rho_r(t_*)\,R^4(t_*)}{R_0^4}\exp\left(-\int_{t_*}^{t_0}\frac{dt}{\chi}\right). \tag{4.30}$$

The first term is the contribution due to direct emission allowing for subsequent absorption. The second term gives the present background due to radiation already existing at epoch t_*. In the particular case of the *exploding models* we have to set $t_* = 0$ and, if we make Gamow's assumption, the second term arises from the primeval black-body radiation associated with the singular state. In this case

$R(t) = B't^{1/2}$ as $t \to 0$ and the corresponding density of radiation is given by (4.26). Thus we verify that the product $\rho(t_*)R^4(t_*)$ is finite in this case, so that this term may well have to be taken account of in observational investigations (§ 6).

The density of background radiation at a given frequency may be obtained in a similar manner. For the special case when the spectral variation of the background obeys a power law, viz. $\rho(\nu, t) \propto \nu^{-x}$ at a given t, where x is a constant and $\rho(\nu, t)$ is the background density per *unit* frequency at frequency ν, we can write (cf. Davidson 1962 c) for a general model

$$\rho_r(\nu_0, t_0) = \int_{t_*}^{t_0} S(\nu_0, t)\left(\frac{R}{R_0}\right)^{3+x} \exp\left(-\int_t^{t_0}\frac{dt}{\chi}\right) dt + \rho_r(\nu_0, t_*)\left\{\frac{R(t_*)}{R_0}\right\}^{3+x} \exp\left(-\int_{t_*}^{t_0}\frac{dt}{\chi}\right). \tag{4.31}$$

In this result $S(\nu_0, t)$ is the rate of emission per unit volume, per unit frequency at frequency ν_0 and at epoch t; $\chi(\nu_0, t)$ is the mean free time for photons of frequency ν_0 at epoch t.

The results (4.30) and (4.31) will evidently also give the energy density of *neutrinos*, regarding the neutrino energy from the point of view of frequency. The *number* of neutrinos in the background at a given energy level would be obtained by dividing (4.31) by $h\nu$ where ν is the associated frequency. To find the total number of neutrinos regardless of energy level one would proceed from an equation of the form (see also Weinberg (1962 a) who uses a somewhat different approach and appears to ignore the possibility of primeval neutrinos)

$$\frac{1}{R^3}\frac{d}{dt}(nR^3) = N - \frac{n}{\tau} \tag{4.32}$$

where $N(t)$ is the rate of emission of neutrinos in unit proper volume, $n(t)$ is the total number of background neutrinos in unit volume and τ is the mean free time between emission and absorption (extremely large compared with H^{-1} in the case of neutrinos). Integration of this equation between $t = t_*$ and $t = t_0$ yields

$$n(t_0) = \frac{1}{R_0^3}\int_{t_*}^{t_0} NR^3 \exp\left(-\int_t^{t_0}\frac{dt}{\tau}\right) dt + \frac{n(t_*)R^3(t_*)}{R_0^3}\exp\left(-\int_{t_*}^{t_0}\frac{dt}{\tau}\right) \tag{4.33}$$

for the present total number density of neutrinos in the Universe.

In conclusion it may be remarked that (4.30), (4.31) and (4.33) apply also to the *steady-state model* if we set $t_* = -\infty$, take $R(t) = e^{t/T}$ $(T = H^{-1})$ and regard all other quantities as independent of t. This means that in each result the contribution due to source-free radiation now vanishes, as is evidently necessary in an expanding sub-stratum of infinite age.

4.5. *Newtonian cosmology*

The ability of Newtonian theory to describe the large-scale motions of matter in the Universe, and in particular its ability to generate the equations (4.1) and (4.2) of evolutionary cosmology will not be dealt with in detail in this article. This matter has been adequately treated in earlier reviews (McCrea 1953, Bondi 1959).

Some reference to the results of Newtonian cosmology will be made in § 9. Here we content ourselves with a few brief remarks.

As is well known, Newtonian arguments can successfully reproduce the main results of relativistic cosmology, albeit they require certain ideas of the relativistic era to accomplish it (for example, the equivalence of mass and energy unless the pressure is zero). What is the reason for their success on the cosmological scale as well as on the local scale?

Essentially, it seems that the particular success of the Newtonian theory in cosmology lies in its local application where it yields the cosmic forms of Gauss's theorem and Poisson's equation. We may also include the local equation expressing the conservation of total energy. These locally applied Newtonian equations are relations connecting the local density and pressure, as functions of t, with the local mass motions. It is just at this point that general relativity can agree with Newtonian theory, for the invariants of general relativity are local observables and the relativistic laws approximate to the Newtonian laws in the local inertial frames (if v/c is small). Moreover, the term involving the space curvature in the general relativity approach enters only as a constant of integration at the Newtonian level.

In the practical reality the notion of absolute space in Newton's theory has to be expressed in terms of the inertial (Newtonian) frame of reference. Consequently, Newton's absolute space is realized in a sense in any local frame of reference moving with the mean motion of cosmic matter. This is clear at least in the case of cosmic isotropy. For in such a frame the recession of distant galaxies is isotropic, which defines this local frame uniquely. Indeed, such a frame may be defined *locally*. For by sufficiently accurate measurements a 'zero-point' level of (isotropic) radiation, which exists at any co-moving point in the Universe as discussed in § 4.4, may be discovered, whether it is related to the Universe at large or not. Moreover, in a unique frame of reference the local four-momentum density (tensor components T_{4i}) of this radiation will be $(0, \rho_0)$, whereas according to special relativity its four-momentum density in another frame will be

$$\frac{[-\frac{4}{3}\rho_0 \mathbf{V}, \rho_0(1+\frac{1}{3}V^2/c^2)]}{1-V^2/c^2}$$

where $\mathbf{V}$ is the relative velocity of the frame. Thus a unique reference frame is determined locally at each point of the Universe and, as regards this local experiment at least, all inertial frames in special relativity are not equivalent or indistinguishable. Unfortunately, Newtonian cosmology is itself quite unable to deal with the wide range of radiation phenomena in the Universe.

The relevance of Newtonian theory to at least some aspects of a universe in isotropic conditions, and also, as shown in recent years, even in universes containing shear and rotation (§ 9), is most encouraging and extremely useful. As long as the analysis is confined to cosmic and astrophysical densities of an order not too different from those presently observed, the Newtonian approach gives reliable results for most purposes, and of course is very much simpler than relativity theory. Even at ultra-high densities the Newtonian theory may well be a first approximation to reality. In short, while Newtonian cosmology (and astrophysics) clearly lacks the power and scope of general relativity, it is nevertheless of the greatest heuristic importance, invaluable in the exploration and exposition of cosmological (and astrophysical) principles.

5. Steady-state theory

5.1. *Introduction*

One of the most interesting questions which a cosmological theory should be able to answer is: what is the origin of matter that we see around in the Universe? It has been argued by some authors that this question is not satisfactorily answered by the evolutionary cosmologies described in the preceding section. In the models which start with a 'big-bang' the creation of primary matter is supposed to have taken place at the time of the big-bang ($t = 0$). However, the equations on which these models are based become singular at $t = 0$ and are not able to provide a satisfactory picture of the Universe for $t \leqslant 0$. It is possible, at this stage, to take the view that near $t = 0$ the state of the Universe is so different from the present state that it is unrealistic to apply the laws of physics as we know them and that the appearance of singularity in the equations is an indication of this. If this view is correct then we can maintain the picture of an evolving universe for $t \geqslant \epsilon$, where ϵ is some small number and argue that for $t < \epsilon$ the laws of physics may be too substantially modified to render the present description applicable. The interesting question of origin of primary matter then belongs to the phase $t < \epsilon$, and is unanswerable in present terms.

There is another way of resolving the difficulty at $t = 0$, and that is through the oscillating cosmologies, in which the Universe oscillates between a dense and a rare phase, and is infinitely old. This is an attractive possibility, but has also some difficulties which are not yet resolved. How does a contracting universe re-expand? Conventional treatment in terms of the Einstein equations appears to lead to a gravitational collapse into a singular state. Even assuming that some new law of physics is able to stop the collapse, nucleosynthesis presents another problem. As will be discussed in a later section, nucleosynthesis in stars tends to build up elements from hydrogen to the most stable nuclei of the iron group. If this process goes on indefinitely the Universe should now consist of predominantly the iron group nuclei, which is not in agreement with the observations. Hence thermodynamic conditions must be such as to reverse the trend of nuclear reactions in the contracting phase of the Universe. This and indeed the general problem of thermodynamics in an oscillating universe needs to be solved satisfactorily.

Many cosmologists believe that the picture of an evolving (big-bang or oscillating) universe is essentially correct and that the difficulties described above may be resolved by a modification of the existing laws of physics (e.g., general relativity, quantum theory, etc.) when the Universe is in a highly dense phase. Others have taken the view that these difficulties justify looking elsewhere than classical general relativity for more satisfactory models. Various cosmological models have therefore been put forward from time to time. One of the most important of these is the steady-state model, put forward in 1948, by Bondi and Gold (1948) and by Hoyle (1948). The approach of Bondi and Gold differs from that of Hoyle, although the ultimate picture of the Universe is essentially the same in both cases. The two points of view are described below.

5.2. *The perfect cosmological principle*

This has already been mentioned in § 3.1. Bondi and Gold argued that the ordinary cosmological principle is not sufficient as a starting point of cosmological investigations. In the notation of § 3.1, this principle requires that the spaces t = constant are homogeneous and isotropic. This ensures that physical conditions in our locality are similar to that of any other locality lying in the same sub-space t = constant. However, cosmological observations refer to the distant parts of the Universe and light takes long times in traversing such distances. Hence when we look at the distant parts of the Universe, we are looking at it as it was in the past. It would be easier to draw valid conclusions from these observations if we were sure that physical laws in the distant past were of the same type as those observed in our locality. If the laws of physics are also determined by the state of the Universe, then constancy of these laws implies an unchanging, i.e. a steady-state, Universe. Bondi and Gold assumed, therefore, that the Universe is not only uniform in space but is also uniform in time. That is, it presents the same large-scale view at all times. This is the perfect cosmological principle.

It follows that the expansion factor in the Robertson–Walker line element must be such that the Hubble constant is independent of the epoch, i.e.

$$H = \dot{R}/R = \text{constant}$$

i.e.

$$R = e^{Ht}; \tag{5.1}$$

the constant of integration has no physical significance and can be absorbed in the scale factor of R. The constancy of curvature of the spaces t = constant requires $k = 0$ and the line element becomes

$$ds^2 = c^2\,dt^2 - e^{2Ht}\{dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)\}. \tag{5.2}$$

This line element is free from singularity. Furthermore, a constant density in an expanding universe can be maintained only provided new matter is created *continually*. Thus creation of matter is not relegated to a distant past, but is made into a law of nature which can, in principle, be tested by observations. This has been considered an attractive feature of the theory.

At what rate should matter be created? What is the mean density of matter in the Universe? These questions are not answered in the Bondi–Gold version of the steady-state theory. The reason is that the theory is deductive in nature and relies entirely on the perfect cosmological principle. This is a powerful principle; but it does not give any quantitative results. Hoyle's version, which uses field equations, gives more information.

5.3. *The C field*

Hoyle adopted the view that creation, like other phenomena in physics, should be describable by physical laws and that the steady-state solution should emerge as a solution of Einstein's equations if extra terms describing creation of matter are included. Thus in equation (5.3), T_{ik} stands for the combination of matter and the creation terms:

$$T_{ik} = \underset{(\text{matter})}{T_{ik}} + \underset{(\text{creation})}{T_{ik}}. \tag{5.3}$$

The total is conserved, but the two terms on the right-hand side need not be separately conserved. An approach along a different line but within the framework of general relativity has also been suggested by McCrea (1951).

What are the creation terms? At first sight the choice might seem enormous. It can be narrowed down if one insists on simplicity. To describe creation Hoyle invented a scalar field C. A simple and elegant way of deriving the energy tensor from an action principle was given by M. H. L. Pryce (private communication). Hoyle and Narlikar (1963 a, 1964 a, b) have considered the properties of the C field in detail. They are summarized below.

The C field can be derived from an action integral

$$J = \frac{c^3}{16\pi G}\int R\sqrt{(-g)}\,d^4x - \Sigma\int mc\,ds + \tfrac{1}{2}f\int C_i C^i\sqrt{(-g)}\,d^4x - \Sigma\int mC_i\frac{dx^i}{ds}\,ds \quad (5.4)$$

where the first two terms give the conventional Einstein equations and the last two terms describe the C field. $C_i = \partial C/\partial x^i$, and f is a coupling constant. Variations of J with respect to g_{ik} and C give the equations

$$R^{ik} - \tfrac{1}{2}g^{ik}R = -\frac{8\pi G}{c^4}\left\{\underset{(\text{matter})}{T^{ik}} - f(C^iC^k - \tfrac{1}{2}g^{ik}C^lC_l)\right\} \quad (5.5)$$

$$fC^i{}_{;i} = n \quad (5.6)$$

where n = number of particles created in a unit four-dimensional proper volume. The particle world lines in this theory are broken, to allow for creation (and destruction). A variation of world lines shows that particle motions are not affected by the C field. The C field, however, does affect the creation of a particle. Thus a particle of momentum p_i will be created at a point provided the C field at that point satisfies the equation

$$C_i = p_i. \quad (5.7)$$

Thus momentum of created particle is balanced by C-field momentum. Divergence of (5.5) gives

$$\underset{(\text{matter})}{T^{ik}{}_{;k}} = fC^iC^k{}_{;k}. \quad (5.8)$$

Thus creation of matter (or destruction) is possible if $C^k{}_{;k} = 0$. Since f is taken to be positive, the C field behaves as a negative energy field. It is easy to verify that (5.2) is a solution of equations (5.5) and (5.6) provided

$$\dot{C} = m, \quad H = \left(\frac{4\pi Gfm^2}{3}\right)^{1/2}, \quad \rho = \frac{3H^2}{4\pi G} \quad (5.9)$$

where m is the mass of a typical particle created. $\dot{C} = m$ is just the condition (5.7). The Hubble constant is thus related to the coupling constant f and the mass created. With the observed value of $H \simeq 3\times 10^{-18}$ sec^{-1}, the mean density of matter in the Universe is about 3×10^{-29} g cm^{-3} and the creation rate about 3×10^{-46} g cm^{-3} sec^{-1}, i.e. about 10 hydrogen atoms per litre every million million years.

Although (5.2) is a solution of the equations (5.5) and (5.6) the question arises: is it the only solution? Clearly it is not. The significance of other solutions can be best understood in connection with Mach's principle.

5.4. *Mach's principle and the creation of matter*

It was pointed out earlier that the Robertson–Walker line element expresses the observed result about the inertial frames—the result which was responsible for the formulation of Mach's principle. General relativity also permits apparently non-Machian solutions such as Gödel's solutions. One can exclude such solutions in the evolutionary cosmologies by saying that the Universe started in a homogeneous isotropic form. Thus it may be that only by a special choice of initial conditions can such an observation be accounted for in classical general relativity.

The steady-state theory does away with the ideas of initial conditions, because according to the theory there was no single moment when the whole Universe started. Hoyle and Narlikar (1963 a) have argued that one would expect that if the creation goes on the Universe would tend asymptotically to a steady state, however far it was from such a state at any given moment. The reason for this expectation is as follows. As the expansion goes on, any inhomogeneities and rotations tend to die away. On the other hand, new matter which is brought in by the C field tends to make the Universe uniform and reduce any rotation as the expansion proceeds. The equation (5.7) shows that p_i, the momentum of the newly created particle, is a gradient of a scalar and has no vorticity.

In principle one could proceed as follows. Set up a system of surfaces $C =$ constant and define $t \propto C$. Suppose we are given the distribution of matter and its velocity on a particular member of the C surfaces. To permit calculation of the situation off the surface it is necessary to specify

$$g_{ik}, \quad \frac{\partial g_{ik}}{\partial x^{\mu}}, \quad \frac{\partial g_{ik}}{\partial t}, \dots \tag{5.10}$$

on the surface. Then, provided the creation goes on, i.e. $C^i_{;i} = 0$, the Universe should eventually tend to the steady-state solution over any given finite volume, for all possible choices of the metric tensor.

The problem as stated above is very complex. Hoyle and Narlikar have considered the case where the line element differs from (5.2) by quantities of the first order. They find that the Universe tends to the steady-state situation in times of the order H^{-1}.

It is essential that $C^i_{;i} = 0$. If $C^i_{;i} = 0$ the C field behaves in an entirely different way. Thus in the homogeneous isotropic case $C = \text{constant}/R^2$, not $\dot{C} = m$, and the steady state is never reached. This solution will be considered in § 9.

5.5. *The nature of creation*

Returning to the question raised in the beginning of the section, we find that the steady-state theory does provide a framework within which creation of matter can be discussed as any other physical process. In the Bondi–Gold approach, the creation process may have to be discussed in a phenomenological way. In the Hoyle version, the quantized theory of the C field should, in principle, yield information about the nature of particles created. This has not been considered so far. The attempts to come to grips with the problem, which will be described briefly here, have been generally of a phenomenological or classical nature.

A few years ago Burbidge and Hoyle (1956) considered the possibility that creation was symmetric between matter and antimatter. This would involve the creation of electron–positron or proton–antiproton pairs. While this possibility is attractive in that there is a theoretical symmetry between matter and antimatter, the observational evidence appears to be against it. Annihilation of matter–antimatter would produce a strong background of γ rays several magnitudes above that observed.

The creation in the form of hydrogen atoms—or electron–proton pairs—has also been considered. Bondi and Lyttleton (1959) pointed out that, if there is a slight difference in the magnitude of electric charges of electron and proton, there will be a net repulsion between hydrogen atoms and this could account for the expansion of the Universe. Laboratory experiments (Hillas and Cranshaw 1959) indicate, however, that the charge excess required by Bondi and Lyttleton (about 1 part in 10^{18}) is far outside the experimental errors (about 1 part in 10^{20}) around the result of zero charge excess. The alternative possibility suggested by Bondi and Lyttleton, that there may be a number excess, is not easy to verify.

An interesting possibility—that of neutron creation—was raised by Gold and Hoyle (1958) and Hoyle (1958). Neutrons decay into electrons, protons and anti-neutrinos, with an energy release of about 2×10^{17} erg per gramme of hydrogen. This energy corresponds to a kinetic temperature of approximately 10^9 °K. At such temperatures pressure gradients are likely to play an important role. Gold and Hoyle considered the possibility that condensations of the order of 30 Mparsec may be formed in the intergalactic medium as a result of pressure gradients. These condensations would further break up into small blobs of about 3 Mparsec radius or less. These were identified with clusters of galaxies. Although this approach had the merit that it brought galaxy formation within cosmological discussion, it was not easy to see how the primary condensations could form in the first place. Further, a temperature of approximately 10^9 °K would lead to an X-ray background twenty times above that observed, as pointed out by Gould and Burbidge (1963) (§§ 6.2, 6.4).

Last year Hoyle and Narlikar (1966 a) considered the problem from an altogether different standpoint. They argued that the problem of galaxy formation was made unnecessarily complicated by the assumption that the Universe is homogeneous. Observations indicate inhomogeneities on the scale of 3–30 Mparsec in the form of clusters and super-clusters of galaxies. These cannot be ignored in a universe with a distance scale of about 3000 Mparsec. Now, as seen before, in a homogeneous steady-state universe the creation rate of about 3×10^{-46} g cm^{-3} sec^{-1} is uniform. Hoyle and Narlikar applied the C field approach to a universe which is inhomogeneous and contains a uniform distribution of massive objects. They found that the creation rate is non-uniform—being large in the neighbourhood of a massive object and small away from it. Massive objects would therefore act as pockets of creation. A similar idea on qualitative grounds has been put forward by McCrea (1964).

Whenever matter is created, a new C field is also generated. As mentioned before, the coupling constant of the C field is such that it has negative energy and therefore in the modified Einstein equations (5.5), it has a repulsive effect. (In the homogeneous isotropic form the effect is similar to that of the Λ term sometimes included in the Einstein equations.) The pockets of creation therefore recede from each other. Also, newly created matter increases the mass of the pocket which grows in size until it fragments, giving rise to smaller pockets. In a steady state, creation and expansion are in balance. The pockets of creation can be identified with massive galaxies or clusters of galaxies. Hoyle and Narlikar (1966 b, c) have suggested that a variant of this model could account for high-energy phenomena in astrophysics such as strong radio sources, quasars, cosmic rays, etc., and could also lead to a picture of galaxy formation. But this means abandoning the simple steady-state picture presented in this section. This model will be briefly described in § 9.

6. Observational tests of cosmological theories

In this section we shall consider observational tests of world models based on the homogeneous isotropic metric.

6.1. *The* (m, z) *relation*

The relation between the luminosity distance D of a galaxy and its directly measured apparent magnitude m is given in equation (2.9), which incorporates the red-shift correction K. The quantity D is defined by equation (3.19). It follows that, making use of equations (3.11), (3.16) and (3.19), D may be expressed in terms of the red shift z in a relation containing the required parameters of the theoretical model. In addition, the K correction may be expressed in terms of the red shift, given the intrinsic emission of the source. If m and M are photographic magnitudes, say, then if we ignore the possibility of any evolution of the emission characteristics of the selected galaxies both K and M may be determined empirically from a local standard galaxy. Thus in principle equation (2.9) provides a theoretical relation between z and m, both of which are directly measured quantities (we ignore corrections due to the apparatus used, obscuration corrections, etc.).

Until recently the measured values of z have been sufficiently small safely to permit expansion of D in series. In this case it is easily seen that

$$\log D = \log\left(\frac{cz}{H}\right) + \frac{1\cdot086}{5}(1-q)\,z + \mathrm{O}(z^2) \tag{6.1}$$

where q and H have their previous meanings. Then, if distance is measured in parsecs, the required relation between m_c $(= m - K)$ and z is

$$m_c = M - 5 + 5\log\left(\frac{cz}{H}\right) + 1\cdot086(1-q)\,z + \mathrm{O}(z^2). \tag{6.2}$$

Neglect of the term of order z in (6.2) evidently gives the first-order Hubble relation $cz = HD$ which is equation (2.10). As stated in § 2 the data give $H^{-1} = T = (1 \pm 0\cdot3) \times 10^{10}$ year.

At the next approximation the coefficient of z provides the fundamental parameter q, which distinguishes the three categories of exploding models in accordance with equations (4.13) and (4.16). As already stated this result depends on the possibility of ignoring any average galactic evolution.

The analysis of the large number of red shifts (up to $z = 0\cdot2$) and corresponding apparent magnitudes measured by Humason *et al.* (1956) yielded the result

$q = 2 \pm 0{\cdot}8$. The observations by Baum (1957, 1961 a, b) using photoelectric photometry yield a calculation having less residual error and this gives, Baum claims, a q lying in the range $\frac{1}{2} < q < 1\frac{1}{2}$ (figure 5). According to the analysis of the exploding

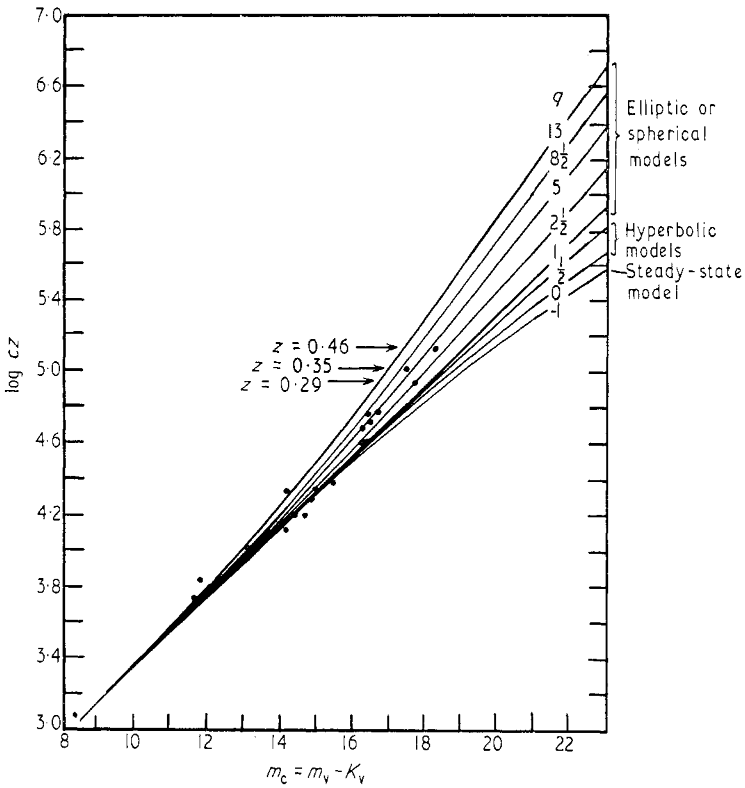

Figure 5. The mean red shifts and apparent magnitudes of 21 clusters of galaxies as given by Sandage (1961 a) and Baum (1961 a, b), together with the theoretical curves for exploding models determined by various values of q (Sandage 1961 a). The curve for the steady-state model ($q = -1$) is also shown. The abscissa gives the photovisual apparent magnitude with the K correction applied, and the ordinate is $\log cz$ when z is the red shift and c is the speed of light in km sec^{-1}. Baum has published (z, m) data, obtained photoelectrically, for eight clusters in all but only his three most distant clusters are shown here with the magnitudes changed to photovisual to correspond to Sandage's. Baum's complete data (1961 b) have considerably less scatter than Sandage's, giving very nearly a straight line on the $(\log cz, m)$ diagram. Baum concludes that $\frac{1}{2} < q < 1\frac{1}{2}$. However, no allowance has been made directly by Sandage or Baum for a plausible evolution of absolute magnitude of the 'standard' galaxy during the light time (but see Sandage 1961 b). Such an allowance might swing the higher points substantially to the right giving a smaller q (see also figure 6). (After Sandage 1961 a, Baum 1961 a, b.)

models in §4, both these results would indicate an oscillating universe of positive curvature and a relatively high density for the present epoch (exceeding $1{\cdot}87 \times 10^{-29}$ g cm^{-3}). Such a high density compared with the observed density of visible matter (equation (2.6)) would imply that more than 98% of the mass–energy of the Universe was in a non-luminous state, presumably in intergalactic space. Furthermore, if $q = 1$, say, the age of the Universe (cf. equation (4.18)) would be only $5{\cdot}7 \times 10^9$ year, taking $H^{-1} = 10^{10}$ year. This appears to be considerably less than the ages of some star clusters of the Galaxy (§ 6.6).

However, it is perhaps unreasonable to ignore evolutionary trends in the mean emission of galaxies since, as mentioned in § 2, a time lapse of $3{\cdot}5 \times 10^9$ year is involved in the case of the greatest red shift, $z = 0{\cdot}46$, observed by Baum. If this is accepted than it is important to notice that due allowance for an evolution of brightness must enter into the linear term in the (m_c, z) relation containing q. In fact to this order the (m_c, z) relation becomes (Davidson 1959, McVittie 1965)

$$m_c = M_0 - 5 + 5 \log\left(\frac{cz}{H}\right) + 1{\cdot}086(1 - q + \lambda^*)\, z + O(z^2) \tag{6.3}$$

in which M_0 is taken to be the absolute magnitude of the nearby standard source and λ^* is related to its rate of change *at the present epoch* by

$$\lambda^* = -0{\cdot}92\, \dot{M}_0/H. \tag{6.4}$$

Sandage (1961 b) and Wielen (1964) have estimated a likely value of $\dot{M}_0$ from the theory of evolution of stars in bright elliptic galaxies. It appears that it might easily be possible to have $\dot{M}_0 = 0{\cdot}09$ per 10^9 year, i.e. galaxies are getting dimmer at this rate at the present epoch. For a time lapse of $3{\cdot}5 \times 10^9$ year this would mean an increase of 0·32 magnitudes (see also Minkowski 1961). Considering the total range of absolute magnitude for all kinds of galaxy, viz. -15 to -22, such a change seems entirely plausible (cf. figure 6). By (6.4) this would mean $\lambda^* = -0{\cdot}83$, and this would convert a value of q which is equal to 1 without allowance for evolution into a value $q = 0{\cdot}17$. Such a model would fall well within category (iii) of § 4.2, i.e. in the range of ever-expanding hyperbolic models. It is interesting to note that for this value of q the present Riemannian curvature of space would be given, by equation (4.13), as $-1/R_0^2 = -7{\cdot}6 \times 10^{-57}\ \mathrm{cm}^{-2}$. Correspondingly equation (4.12) or (4.15) would give for the present mean mass density in the Universe the value $6{\cdot}4 \times 10^{-30}\ \mathrm{g\,cm}^{-3}$.

We note that the above value for the density still exceeds the observed value for luminous matter by a factor of 20, so that 95% of the mass must be non-luminous. It has perhaps not been fully appreciated, however, that the determination of q in this way is exceedingly sensitive to the value assigned to $\dot{M}_0$. Thus, if we take

$$\dot{M}_0 = 0{\cdot}108 \text{ per } 10^9 \text{ year} \tag{6.5}$$

instead of 0·09 as before, then this is an equally plausible value. But then a value of q, which is equal to 1 when no evolutionary allowance is made, now drops to the very low value

$$q = 0{\cdot}0064 \tag{6.6}$$

in the hyperbolic range of models. The curvature of space becomes slightly altered to

$$-1/R_0^2 = -1{\cdot}14 \times 10^{-56}\ \mathrm{cm}^{-2} \tag{6.7}$$

but the mean mass density is substantially changed to the value

$$\rho = 2{\cdot}4 \times 10^{-31}\ \mathrm{g\,cm}^{-3} \tag{6.8}$$

which is only slightly below the observed density of luminous matter.

There is an added attraction to a value of q such as (6.6) for then the present age of the Universe would be, by equation (4.20), very close to $H^{-1} \simeq 10^{10}$ year. This would be in more reasonable accord with the estimated age of the Galaxy (§ 6.6).

Evidently the precise determination of q in the future will depend on the selection of special types of source, the evolutionary characteristics of which must be determined accurately. Furthermore, if the red shifts of order unity recently measured are used, then exact relations between D, m and z will be required. The (m, z) relation in a more exact form must not only incorporate evolutionary effects, but must also take account of the perhaps quite different class of object being studied, for example, the quasi-stellar sources of which several already examined are 100 times brighter than normal galaxies quite apart from any evolutionary properties that they may possess (figure 6).

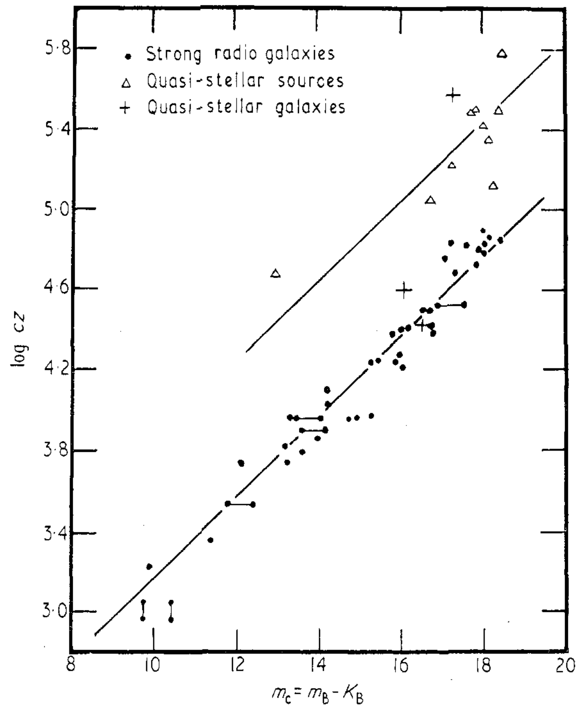

Figure 6. This demonstrates the red shifts and apparent magnitudes for identified radio galaxies, quasi-stellar sources and quasi-stellar galaxies as given by Sandage (1965), the red shifts being due mainly to Schmidt (1965 a, b). The ordinate gives log cz where c is in km sec^{-1}. The abscissa is the photoelectric blue apparent magnitude as measured by Sandage and corrected for the K effect. The joined points represent dumb-bell galaxies—one point representing the bright component, the other the two components combined. The radio galaxies lie on the straight line $m_c = 5 \log z + 22{\cdot}516$ approximately, which would be the first-order (m, z) relation appropriate to bright normal galaxies with $M = -20{\cdot}5$ (see equation (6.2)). The quasi-stellar sources, and at least one quasi-stellar galaxy, lie on the line given roughly by $m_c = 5 \log z + 18{\cdot}186$, showing that even the average quasi-stellar source is 4·33 magnitudes, or about fifty times brighter than the bright normal galaxies. *N.B.* Another 12 red shifts of quasi-stellar sources have been obtained very recently (§2.2). (From Sandage 1965.)

Already, however, it is clear that the (m, z) data are reasonably consistent with the predictions of the exploding models. On the other hand, it appears much more difficult to reconcile them with the steady-state cosmology in its original simple form, for this predicts $q = -1$ (since $R(t) = e^{t/T}$). However, the scatter on the (m, z) diagram is perhaps still too great and the data are inadequate to make this result conclusive.

6.2. *The counts of radio sources*

The first thoroughgoing analysis of the counts of radio sources was made by the Cambridge group (Ryle and Clarke 1961), with a view to determining whether they were consistent with the original simple form of the steady-state model. The results appeared to be conclusive against this possibility.

The radio sources are counted as a function $N(S)$ representing the number (sterad^{-1}) of sources whose flux density exceeds S at the given frequency ν_0. The Cambridge group have also defined a useful quantity $N_0(S)$, the number of sources of flux density exceeding S in a static Euclidean universe. Thus

$$N_0(S) = \tfrac{1}{3}\bar{n}_0 \bar{P}_0^{3/2} S^{-3/2} \tag{6.9}$$

where $\bar{n}_0$ is a weighted (dispersion-averaged) mean number density of the sources and $\bar{P}_0$ is their median power (w sterad^{-1} (c/s)$^{-1}$) at the frequency ν_0. These quantities are assumed to have the values observed at high flux density in the actual Universe. As pointed out in connection with equation (2.12), this means that if the Universe is homogeneous and isotropic at a given epoch then

$$N(S) \to N_0(S) \quad \text{as} \quad S \to \infty. \tag{6.10}$$

The coefficient $\bar{n}_0 \bar{P}_0^{3/2}$ is therefore given by the counts at high S so that $N_0(S)$ can be evaluated for any S.

Figure 7 shows the combined counts of the 3C and 4C surveys at 178 Mc/s (heavy lines) and also two limiting curves for the steady-state model between which all possible curves for that model must lie. The curve chosen depends on the median power of the sources and this was not known exactly in 1961. The abscissa gives the flux density S in flux units on a logarithmic scale. The ordinate gives $\log\{N(S)/N_0(S)\}$. The sources were counted individually down to $S = 2$ flux units and the extensions (i), (ii) and (iii) are possible interpretations of the counts for smaller S by statistical analysis of the recording chart.

The graphs show that even if the upper limiting curve for the steady-state model is taken, corresponding to a median power of the sources of 10^{24} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s, the actual count down to $S = 2$ flux units exceeds the predicted count in the steady-state model in the ratio 3 : 1. If the lower curve is taken, for which the median power $\bar{P}_0$ is 10^{26} w sterad^{-1} (c/s)$^{-1}$ the ratio is 9 : 1. As noted in § 2, the median power at high flux density has since been established to lie between 10^{25} and 10^{26} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s with roughly the same dispersion as that assumed by Ryle and Clarke. The Cambridge results, because of the relatively small range of observational and instrumental error (indicated in figure 7), and the known isotropy of distribution of the sources referred to in § 2, appear to rule out the simple steady-state model beyond all reasonable doubt.

Attempts have been made to reconcile the counts with a stationary model universe which contains a certain degree of inhomogeneity. Hanbury Brown (1962) supposed that there was an irregularity in the local distribution of sources which was associated with the Supergalaxy (the alleged concentration of local clusters of galaxies centred on the Virgo cluster). Apart from the fact that this model was based on a median power $\bar{P}_0$ of 10^{24} w sterad^{-1} (c/s)$^{-1}$, now known to be too low, no evidence has been found to associate radio sources with the Supergalaxy. In

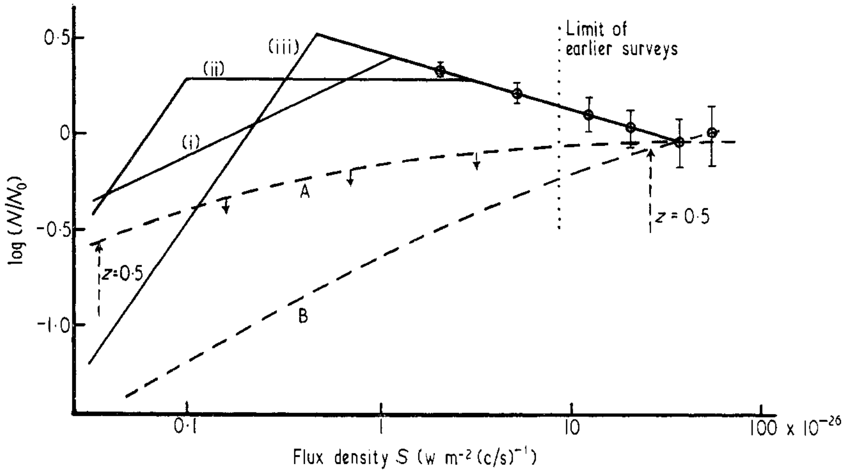

Figure 7. The counts of radio sources in the 4C survey at 178 Mc/s are expressed as a function of flux density S. The ordinate is $\log(N/N_0)$, where $N(S)$ is the number of sources (per steradian) of flux density exceeding S and $N_0(S)$ is the corresponding number expected in a static Euclidean universe. The sources are counted individually down to $S = 2$ flux units and thereafter there are extensions (i), (ii) and (iii) which are alternative statistical interpretations of the records. Two theoretical curves (broken curves) are shown for the steady-state model. In curve A the lowest possible median power $\bar{P}_0$ consistent with other observations (about 10^{24} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s) is chosen for the sources and all other curves must fall below this. In curve B the choice is $\bar{P}_0 = 10^{26}$ w sterad^{-1} (c/s)$^{-1}$ which is a little above the value since established as correct. Even in case A at $S = 2$ flux units the observed counts exceed the predicted counts in the ratio 3 : 1. (From Ryle and Clarke 1961.)

addition, the predicted curve gave a (too low) maximum at $S = 10$ flux units instead of near 1 flux unit as observed (Clarke *et al.* 1963). In another theory, Hoyle and Narlikar (1961, 1962 b) postulated inhomogeneities in the steady-state model on a scale of 500 Mparsec and added the hypothesis that in these expanding irregularities the galaxies of an age exceeding H^{-1} would have a sharp rise in probability of becoming a radio source. In this way they were able to obtain $\log N/N_0$ curves which at first rose as S decreased in the way observed. However, the physical origin of the inhomogeneities on such a large scale, based on an earlier theory of continuous creation by Gold and Hoyle (1958), required an intergalactic temperature of 10^9 °K, and conclusive evidence appears to have been produced against such a temperature existing in the intergalactic medium (Gould and Burbidge 1963, cf. § 6.4). The very recent theory of the steady-state model by Hoyle and Narlikar

(§ 5, § 9), predicting an appearance of the Universe closely resembling the present phase of an exploding universe of relativistic cosmology, would evidently require careful analysis to distinguish it from such a model. No attempt has been made to do this on the basis of the radio-source counts.

An effort to save the steady-state theory in its simple form has been made by Sciama (1963, 1964). In this theory he supposes that the greater proportion of the sources having $S < 10$ flux units lie within the Galaxy, an increasing proportion lying outside the Galaxy as S rises above this level. The rising $\log(N/N_0)$ curve is explained by supposing that the distribution of the (intrinsically weak) sources within the Galaxy has a hole of radius 17 parsec in our neighbourhood, i.e. there happens to be a relative dearth of sources near the Sun. The model requires that for $S > 9$ flux units about 57% of the sources should be in the Galaxy. On the other hand, the recent identifications with extragalactic objects of 160 out of the 328 sources of the revised 3C catalogue ($S > 9$ flux units) already accounts for 19 more than the permissible 43% that are extragalactic. Furthermore, at least another 80 of the more powerful sources having $S > 9$ flux units might be expected to lie beyond the reach of the 200 in. telescope if the median power lies between 10^{25} and 10^{26} w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s.

As Scott (1963) has pointed out, although in Sciama's model the galactic sources must have a power of only about 3×10^{10} w sterad^{-1} (c/s)$^{-1}$, they must nevertheless possess the same angular structure and spectra discovered for most sources having $S > 12$ flux units. Yet, 90% of 3C sources having $S > 40$ flux units and so far about 50% of those having $S > 9$ flux units are identified with extragalactic objects. Scott also claims that a lack of isotropy would be expected in a galactic distribution at $S \simeq 2$ flux units, yet no deviation from a random distribution has been found down to at least 0·25 flux units. It would seem, therefore, that Sciama's model, ingenious as it is, fails to predict the observed counts and requires several somewhat unlikely circumstances to sustain it in other respects.

Explanations and theories of the source counts in terms of evolutionary cosmology have been made by Oort (1961, unpublished), Davidson (1962 a, b), Davidson and Davies (1964 a, b). Oort has suggested that the observations may be explained by postulating that the fraction of galaxies that are strong radio sources increased by a factor of several hundred during the first 10^9 years of the life of the Universe. No details of how the counts can be exactly matched by such a hypothesis have been advanced but the conjecture is a plausible one which deserves further investigation.

Davidson (1962 a) has shown that for the $\log(N/N_0)$ curve to rise as S decreases from high values the condition

$$(\bar{n}_0 \bar{P}_0^{\,2})^{-1} \frac{d}{dt_0} (\bar{n}_0 \bar{P}_0^{\,2}) < -(5+2x) H \tag{6.11}$$

must be satisfied at the present epoch t_0. Here $\bar{n}_0$ is the weighted mean space density of the sources at this epoch (i.e. observed at high flux density), $\bar{P}_0$ is the median power of the sources so observed and x is their mean spectral index. This implies that either the mean power of the sources increases into the past at a present minimum rate, or the number density does, or there is a combination of both these effects.

A preliminary account (Davidson 1962 b) of how the counts could be fitted, in the case when the population of radio sources remains a constant fraction of that of the galaxies for all epochs, was followed by a detailed computation for a four-parameter family of evolutionary models. This family had the metric parameters $R(t) = t^n$ $(\frac{1}{2} < n < 1)$ and $k = 0$, while the power $P(\nu, t)$ of a radio source was assumed to depend on ν and t in accordance with power laws of the form

$$P(\nu, t) \propto \nu^{-x} t^{-m} \quad (m > 0)$$

in which x was taken to be 0·8. The dispersion of power of the sources, at the present epoch t_0 and at the frequency of observation ν_0, was taken to be Gaussian in $\log(P_0/\bar{P}_0)$ of a standard deviation σ. The quantities n, m, σ and $\bar{P}_0$ were therefore

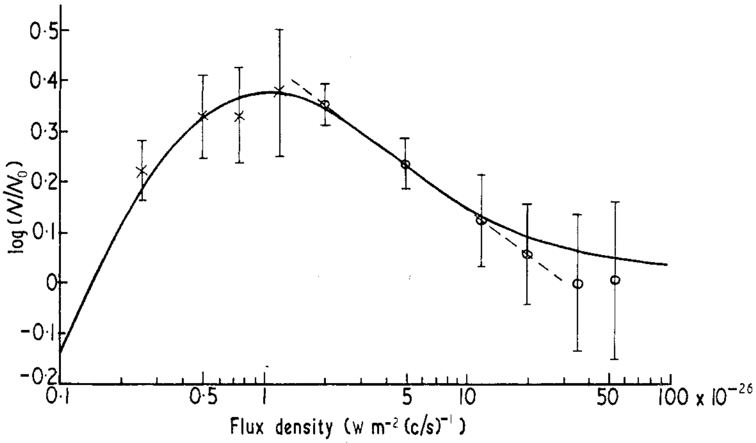

Figure 8. The earlier 4C counts at 178 Mc/s (see figure 7) are extended by the data of the North Polar counts using a more sensitive interferometer (Ryle and Neville 1962). Also shown is the best fitting curve for a four-parameter family of evolutionary models (Davidson and Davies 1964 a, b). These models are based on the supposition that the population of radio sources is a constant fraction of that of galaxies, and the possibility of fit depends on an increase of mean source power into the past up to a maximum at the earliest epoch of existence of radio sources. It is found to require, also, that the dispersion in power must have been substantially less in the past than that observed in the identified sources.

four disposable parameters to be determined as far as possible by the source count data. Radio sources were supposed to have been first formed at an epoch t^* (>0) suggested by the theory, these earliest sources having had the greatest power in accordance with the postulated power law.

The best fitting curve (Davidson and Davies 1964 b) is shown in figure 8 and corresponds to the parameters $\sigma = \frac{1}{2}$, $\{(1-n)^2/n^2\}\bar{P}_0 = 1{\cdot}25 \times 10^{25}$ w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s, and $\alpha = 1{\cdot}75$ where

$$\alpha = \frac{m - 3{\cdot}8n + 2}{2(1-n)}. \tag{6.12}$$

In particular, it was found that an Einstein–de Sitter type model $(n \simeq \frac{2}{3})$ would give a good fit and yield a median power

$$\bar{P}_0 = 5 \times 10^{25} \text{ w sterad}^{-1} (\text{c/s})^{-1}$$

at 178 Mc/s, close to that observed for the nearer sources (equation (2.13)). For this case (6.12) implies $m = 1{\cdot}7$, so that the power $P(\nu, t)$ at any frequency ν would vary as $t^{-1\cdot 7}$ for as long as radio sources have existed, viz. since the epoch t^*. For $n = \frac{2}{3}$ this epoch was found to be

$$t^* = 0{\cdot}053\,H^{-1} \simeq 5{\cdot}3 \times 10^8 \text{ year} \tag{6.14}$$

at which the red shift is

$$z^* = 4{\cdot}44. \tag{6.15}$$

With $m = 1{\cdot}7$ the mean initial power of the sources at $t = t^*$ would be

$$\bar{P}(\nu_0, t^*) = 4 \times 10^{27} \text{ w sterad}^{-1}\,(\text{c/s})^{-1}$$

at 178 Mc/s. This is the power of the Cygnus A source, somewhat less than the power of 3C 295, and considerably below that of the quasar 3C 9 which is 5×10^{28} at a red shift of $z = 2{\cdot}012$. It appears, therefore, that the above evolution of *mean* power is entirely compatible with the available data. It is to be borne in mind, however, that it may be possible to fit the counts with a smaller evolutionary power rate provided that a greater proportion of galaxies were radio sources in the past.

6.3. *The radiation background at radio and optical frequencies*

If we neglect the absorption of radiation at *radio* frequencies, which is likely to be small in space as a whole, and also neglect for the time being any black-body (primeval) component in the background, equation (4.31) reduces to

$$\rho_r(\nu_0, t_0) = \int_{t_*}^{t_0} S(\nu_0, t) \left(\frac{R}{R_0}\right)^{3+x} dt. \tag{6.16}$$

This is the background density at frequency ν_0 and epoch t_0 due to direct emission from matter in all space.

For the simple steady-state model we have to set $R(t) = e^{t/T}$ $(T = H^{-1})$, $t_* = -\infty$ and $S(\nu_0, t) = \text{constant} = S(\nu_0, t_0)$. Whereupon

$$\rho_r(\nu_0, t_0) = \frac{H^{-1}\,S(\nu_0, t_0)}{3+x} \quad \text{(steady-state model).} \tag{6.17}$$

For the exploding models we shall consider two cases as representative examples, the Einstein–de Sitter model for which $R(t) \propto t^{2/3}$, and the limiting case of the hyperbolic models, the Milne universe, for which $R(t) \propto t$. Any possible power evolution will be neglected here, plausible as it is; allowance for such an effect would increase the background by a factor of order unity (Davidson 1962 c). However, in relativistic cosmology it is appropriate to allow for the expansion of space, which diminishes the number of sources existing in unit volume of space at any instant. In the case of radio background sources we shall make the same assumption as for normal galaxies:

$$S(\nu_0, t) = S(\nu_0, t_0) \left(\frac{R_0}{R}\right)^3.$$

For the Einstein–de Sitter and Milne models we therefore derive, setting $t_* = 0$,

$$\rho_r(\nu_0, t_0) = \frac{2H^{-1}\,S(\nu_0, t_0)}{3+2x} \quad \text{(Einstein–de Sitter model)} \tag{6.18}$$

and
$$\rho_r(\nu_0, t_0) = \frac{H^{-1} S(\nu_0, t_0)}{1+x} \quad \text{(Milne model).} \tag{6.19}$$

In § 2 it was pointed out that, at frequencies of order 178 Mc/s, the contribution to the background from the strong radio sources was only 1/40 of that due to the normal galaxies. Considering the latter only, therefore, we take $x = 0{\cdot}6$ and $S(\nu_0, t_0) = 4\pi nP$, where $n = 0{\cdot}46$ galaxies/Mparsec^3 and $P = 10^{21}$ w sterad^{-1} (c/s)$^{-1}$ at 178 Mc/s. Thus $S(178, t_0) = 1{\cdot}94 \times 10^{-45}$ erg cm^{-3} sec^{-1} (c/s)$^{-1}$ and so

$$\rho_r(178, t_0) = \left.\begin{array}{ll} 1{\cdot}67 \times 10^{-28}\ \text{erg cm}^{-3}\,(\text{c/s})^{-1} & \text{(steady-state model)} \\ 2{\cdot}86 \times 10^{-28}\ \text{erg cm}^{-3}\,(\text{c/s})^{-1} & \text{(Einstein–de Sitter model)} \\ 3{\cdot}76 \times 10^{-28}\ \text{erg cm}^{-3}\,(\text{c/s})^{-1} & \text{(Milne model)} \end{array}\right\}. \tag{6.20}$$

Comparing with (2.15) we see that all these results are quite close to the upper limit for the extragalactic background detected at 178 Mc/s (cf. figure 9). Thus it appears that the radio background at this frequency, and presumably also in the very wide range of frequency where a mean spectral index of about 0·6 applies, can be satisfactorily attributed to the radio emission from normal galaxies. On the other hand, the proximity of the figures for the three models affords little room for distinguishing them observationally on this basis. It may be significant that the relativistic models are closer to the upper limit.

As remarked in § 2.5, it may be possible to explain the unexpectedly high result detected for the background at 4080 Mc/s by attributing it to residual black-body radiation in an exploding universe. Such an explanation is not available for a steady-state universe. Indeed, if one extrapolates from the observed upper limit at 178 Mc/s (80 °K) by a power law of the form $\rho_r(\nu_0, t_0) \propto \nu_0^{-x}$, then even for x as low as 0·5 one finds that, at 4080 Mc/s, $\rho_r(4080, t_0)$ would be no more than

$$6{\cdot}8 \times 10^{-29}\ \text{erg cm}^{-3}\,(\text{c/s})^{-1}.$$

This corresponds to a temperature of only 0·03 °K, whereas Penzias and Wilson find a temperature of 3·5 °K. Unless a satisfactory explanation in a steady-state universe was forthcoming this would provide further evidence against the *simple* steady-state model at least. Such an explanation might be possible by invoking a high temperature in intergalactic space ($> 10^4$ °K), which would mean that intergalactic hydrogen was ionized. According to Kaufman (1965) the Penzias and Wilson result can be accounted for by electron–proton bremsstrahlung if the hydrogen is at a temperature of about 10^6 °K and of density 10^{-30} g cm^{-3}. However, this density is a factor of 10 below the value usually assumed for the steady-state model. In the case of relativistic cosmology it would imply a hyperbolic model, but the ionized gas hypothesis is not at all plausible in relativistic cosmology (cf. §§ 6.4, 6.5). On the other hand, the following argument shows how such a high temperature might be explained in terms of primeval radiation in an exploding universe (see also Dicke *et al.* 1965).

By the Gamow argument in § 4 we have seen that the density of radiation for small t is

$$\rho_r = \frac{3}{4\kappa t^2}$$

$$= \frac{4{\cdot}48 \times 10^5}{t^2}\ \text{g cm}^{-3} \tag{6.21}$$

where t is in seconds. This means that the radiation temperature will be

$$T_r = \frac{1{\cdot}52\times 10^{10}}{t^{1/2}}\ {}^\circ\text{K}. \tag{6.22}$$

When T_r has dropped sufficiently for the interaction between matter and radiation (pair production, etc.) to have ceased, we shall have $\rho_m = Ct^{-3/2}$ as in equation (4.27). Now the results of the $\alpha\beta\gamma$ theory for the production of helium (§8) require that $\rho_m = 10^{-6}$ g cm^{-3} at $t = 670$ sec (Alpher and Herman 1950). Hence

$$\rho_m = \frac{1{\cdot}74\times 10^{-2}}{t^{3/2}}. \tag{6.23}$$

At much larger t when radiation is no longer the dominant influence in the Universe, equations (6.21)–(6.23) will no longer apply. On the other hand, ρ_r and ρ_m will continue to satisfy $\rho_r R^4$ = constant and $\rho_m R^3$ = constant, where $R(t)$ is the metric factor of space–time. It follows that $(T_r)^3/\rho_m$ = constant, so that (6.22) and (6.23) give

$$T_r^3 = 2{\cdot}02\times 10^{32}\,\rho_m.$$

Therefore, if ρ_m takes the value given by the present mean density of observed matter (equation (2.6)), then the present temperature of primeval radiation would be

$$T_r(t_0) = 4\ {}^\circ\text{K} \tag{6.24}$$

which is close to the value found by Penzias and Wilson (cf. figure 9).

It is to be noted that the present discussion implies a total mass–energy density in the Universe extremely close to the observed density in the form of matter. For by (6.24)

$$\rho_r(t_0) = 2{\cdot}2\times 10^{-33}\ \text{g cm}^{-3} \tag{6.25}$$

which is less than the assumed value for

$$\rho_m(t_0) = 3{\cdot}2\times 10^{-31}\ \text{g cm}^{-3} \tag{6.26}$$

by a factor of 150. In this case the argument of §6.1 would indicate that the deceleration parameter q is very small, of order 0·006, so that we should live in a hyperbolic universe with $R(t) \propto t$, approximately, at the present epoch.

Regarding the optical region of the background radiation, it is evident that the contribution of black-body radiation at a temperature 3·5 °K would be utterly negligible. Thus the result (4.30) for the *total* background due to galaxies, etc. reduces to

$$\rho_r(t_0) = \frac{1}{R_0^4}\int_{t_*}^{t_0} S(t)\,R^4 \exp\left(-\int_t^{t_0}\frac{dt}{\chi}\right)dt. \tag{6.27}$$

For the simple steady-state model $S(t)$ = constant = S_0 and χ = constant. Hence

$$\rho_r(t_0) = \frac{S_0}{4H+\chi^{-1}}.$$

Absorption by dust, etc. in galaxies, of density 0·46 galaxies/Mparsec^3 (equation (2.3)) and effective cross-sectional radius 15 000 light year (equivalent to an absorption of 20–25% of incident radiation) would yield $\chi \simeq 10^{11}$ year. This of course is probably a lower limit for χ; for dust-free galaxies χ might be as high as 10^{23} year owing to absorption by stars only. In any case we see that in the steady-state model, taking $H^{-1} \simeq 10^{10}$ year, absorption has a negligible effect on the background and

$$\rho_r(t_0) \simeq \tfrac{1}{4} S_0 H^{-1} \quad \text{(steady state).} \tag{6.28}$$

In the case of exploding models we shall again neglect possible evolutionary trends in the intrinsic luminosity of sources, but we assume that sources are conserved in number as the Universe expands. This means that

$$S(t) = S(t_0)\left(\frac{R_0}{R}\right)^3.$$

Hence in the Einstein–de Sitter and Milne models, for which $R(t) = t^n$ with $n = \frac{2}{3}$ and 1, respectively, we derive for $t_* = 0$, taking χ = constant as a first approximation,

$$\rho_r(t_0) = S_0 \int_0^{t_0} \left(\frac{t}{t_0}\right)^n \exp\left\{\frac{-(t_0 - t)}{\chi}\right\} dt.$$

Now $t_0 = nH^{-1}$, so that during the important part of the range of integration the exponential factor is very close to unity even if χ is as short as 10^{11} year. It follows that omitting a factor of order unity

$$\rho_r(t_0) \simeq \begin{cases} \frac{2}{5} S_0 H^{-1} & \text{(Einstein–de Sitter)} \\ \frac{1}{2} S_0 H^{-1} & \text{(Milne)} \end{cases}. \tag{6.29}$$

The three results again agree very closely, so that an observational distinction between them is probably impossible. According to (2.4),

$$S_0 = 3{\cdot}2 \times 10^{-32} \text{ erg sec}^{-1} \text{ cm}^{-3}$$

so that with $H^{-1} = 10^{10}$ year the results are all within a factor of 2 of

$$\rho_r(t_0) = 5 \times 10^{-15} \text{ erg cm}^{-3}. \tag{6.30}$$

Comparing with (2.17) and (2.18) this result appears to be in good agreement with the available data. Thus, as in the case of low radio frequencies, the observed optical background can be satisfactorily explained in terms of the emission of galaxies. It will be noted, however, that the *total density* of this directly emitted radiation would be less than that of the postulated black-body radiation (including neutrinos), as given by (6.25), by a factor of at least 400.

6.4. *The x-ray and γ-ray background*

It is clear that at X- and γ-ray frequencies a primeval black-body component at 4 °K would be even smaller than at optical frequencies and can be ignored.

If indeed the X-ray and γ-ray backgrounds have an extragalactic origin then examination of the spectra and intensities may yield important information about the evolution of the Universe (cf. figure 3).

Taking first the case of *x rays* the isotropy of the background appears to belie the possibility that its origin is galactic, if this depends on the more likely mechanisms. The synchrotron radiation of relativistic electrons (of cosmic rays) would depend on the variability of the ambient magnetic field of the Galaxy, and the interaction of cosmic rays with interstellar gas (by bremsstrahlung or inverse Compton effect of the electrons, or by π^0 meson decay) would be stronger in

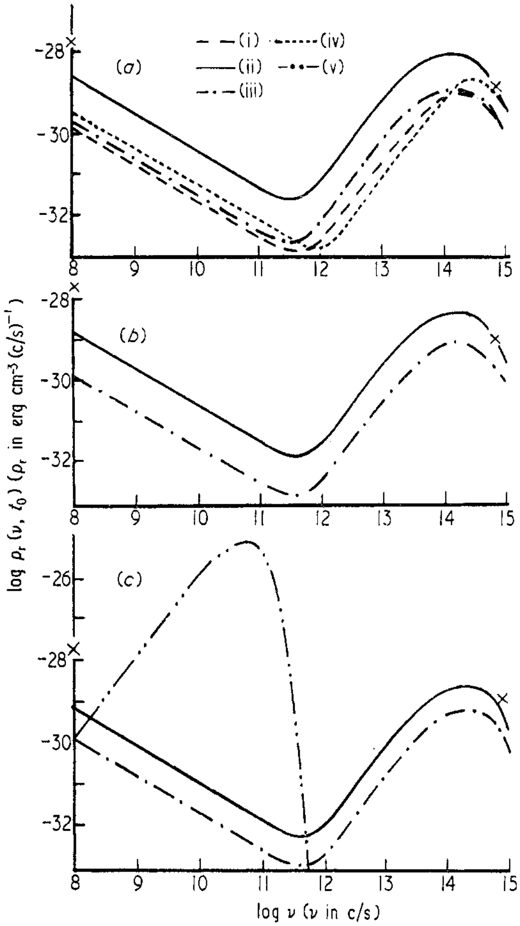

Figure 9. The densities of background radiation from radio to optical frequencies, as predicted for different evolutionary cosmologies and various equations of state, are shown. Two points (crosses) from observational data are also shown (cf. (2.15) and (2.16)).

Three theoretical models are considered: (*a*) in this case the total mean density in the Universe is taken to be that of luminous matter at zero pressure, viz.

$$\bar{\rho} = 5 \times 10^{-31}\,\mathrm{g\,cm^{-3}}, \quad \bar{p} = 0$$

(this might be assumed to correspond to a hyperbolic ever-expanding model of small q), (*b*) a universe approximating to the Einstein–de Sitter model in which

$$\bar{\rho} = 3H^2/\kappa = 1{\cdot}86 \times 10^{-29}\,\mathrm{g\,cm^{-3}}, \quad \bar{p} = 0$$

and (*c*) a universe in which radiation predominates due to, say, neutrinos with

$$\bar{\rho} = 3p = 1{\cdot}86 \times 10^{-29}\,\mathrm{g\,cm^{-3}}.$$

The computed curves (i)–(v) correspond to the following conditions: (i) Galaxies become luminous at time t when their spacing is 0·3 of what it is today, i.e.

$$R(t)/R(t_0) = 0{\cdot}3$$

the luminosity of galaxies being constant. (ii) Galaxies become luminous when $R(t)/R(t_0) = 0{\cdot}1$ and the total emission of type I stars evolves systematically.

directions of high gas concentration. In any case, as Gould and Burbidge (1963) have pointed out, the synchrotron origin in a magnetic field of 10^{-6} G by electrons of 10^{14} ev would require an unlikely intensity of electrons (primary or secondary). Also, the inverse Compton effect of stellar photons and relativistic electrons (Felten and Morrison 1963) would give a background 300 times too small. Finally, Ginzburg and Syrovatskii (1965) have shown that π^0 meson decay, electron bremsstrahlung and inverse Compton effect are all likely to give backgrounds several orders of magnitude below the observed level.

If we consider an extragalactic origin it is appropriate to postulate a mean space emission function $S_x(\nu, t)$ (erg sec^{-1} cm^{-3} (c/s)$^{-1}$) for X rays and calculate the present background density $\rho_x(\nu_0, t_0)$ (erg cm^{-3} (c/s)$^{-1}$) by means of (4.31), neglecting the black-body term. It is probably a good approximation to neglect also the absorption factor since X rays of energy exceeding 1 kev would be weakly absorbed by intergalactic hydrogen. Therefore, if we also ignored evolutionary trends in the mean emission, and assumed a spectral index, we would obtain values for $\rho_x(\nu_0, t_0)$ in the three models, the steady-state, Einstein–de Sitter and Milne models, completely analogous to (6.17), (6.18) and (6.19). These will all be close to the value

$$\rho_x(\nu_0, t_0) = \tfrac{1}{2}H^{-1}\, S_x(\nu_0, t_0).$$

Such a result, calculated for assumed mechanisms of emission, and possibly absorption, could be compared with the observed background spectrum and used to test the validity of these assumptions. Permissible mechanisms would give a pointer to the evolution of the Universe and acceptable cosmological models.

It is to be noted that an X-ray quantum, even if emitted from a region of red shift as high as $z = 4$, will have its energy reduced by a factor of 5 only, on reaching our neighbourhood. Consequently such radiation will still be mainly in the X-ray region when it is detected. As an alternative to the above analysis, therefore, it is appropriate to consider a total X-ray emission function, $S_x(t)$, and a corresponding background density $\rho_x(t_0)$ (erg cm^{-3}). For the same three cosmological models the values of this quantity will all be close to (cf. (6.29))

$$\rho_x(t_0) = \tfrac{1}{2}H^{-1}\, S_x(t_0). \tag{6.31}$$

Can the X-ray background be attributed to the integrated effect of normal galaxies containing sources of X rays as discovered in the Galaxy? The total emission of X rays from the Crab nebula, deduced from the received flux and its estimated distance, is about 10^{37} erg sec^{-1}. As stated in § 2, ten such sources have been detected. On this basis a normal galaxy might have an X-ray emission of

---

Figure 9. (*cont.*)

(iii) Galaxies become luminous when $R(t)/R(t_0) = 0{\cdot}1$ and their luminosity remains constant. In cases (i)–(iii) a plausible evolution of the total radio emission (non-thermal) of galaxies is assumed. (iv) Galaxies of constant luminosity are assumed to be at relative rest and to occupy a sphere of radius CH^{-1} in Euclidean space. (v) Equilibrium black-body radiation of present temperature 1 °K (cf. (6.24)).

The discrepancy between observation and calculation at $\nu = 10^8$ is presumably due to an adoption by the authors of too low a value for the space radio emission due to normal galaxies (cf. (6.20)). (From Doroshkevich and Novikov 1964.)

10^{38} erg sec^{-1}. (Clark (1965, unpublished) estimates a total emission of

$$3 \times 10^{37} \text{ erg sec}^{-1}$$

for the Galaxy.) If we are given the space density of galaxies (equation (2.3)), the present space emission of X rays due to normal galaxies would be

$$S_x(t_0) = 10^{-36} \text{ erg sec}^{-1} \text{ cm}^{-3}$$

so that the corresponding background density would be

$$\rho_x(t_0) = 1{\cdot}5 \times 10^{-19} \text{ erg cm}^{-3} \qquad (6.32)$$

(for normal galaxies). This value is about 15 times below the observed background density (equation (2.19)), but considering the approximations involved the result may be significant.

One of the more plausible explanations of the X-ray emission from the galactic sources is synchrotron emission of 10^{13} ev electrons in a magnetic field of about 10^{-3} G, provided a mechanism for sustaining the input of such high energy electrons can be found. With regard to the Crab nebula the lifetime must be at least 1000 year. Assuming a supernova explosion every 100 year this would give an average of ten sources observable at any one time, in agreement with the recent discoveries. However, once we accept a synchrotron contribution from the supernovae of normal galaxies, it becomes apparent that radio galaxies, quasars and other exploding galaxies might play a large part in maintaining the X-ray background. The average total ultra-violet emission of nine quasars whose red shifts have been given by Schmidt (1965 b) is evidently about 6×10^{45} erg sec^{-1}, and for the space density of all exploding objects we shall adopt a third of the value given in (2.14) for radio sources, which will be 10^{-7} Mparsec^{-3} or $3{\cdot}3 \times 10^{-81}$ cm^{-3}. Assuming that the X-ray emission is comparable with the ultra-violet emission we now find that

$$S_x(t_0) = 2 \times 10^{-35} \text{ erg cm}^{-3} \text{ sec}^{-1} \qquad (6.33)$$

and correspondingly by (6.31)

$$\rho_x(t_0) = 3{\cdot}1 \times 10^{-18} \text{ erg cm}^{-3} \qquad (6.34)$$

(for exploding objects). This result is so close to the observed background density that a possible connection between them must be seriously considered.

As stated in § 6.2, Gould and Burbidge (1963) have shown that the proton–electron bremsstrahlung in the 'hot' universe theory of Gold and Hoyle (1958) would lead to an X-ray background exceeding the observed value by a factor of at least 20 and possibly 100. This result depended on the assumption that the intergalactic medium had the steady-state density of about 2×10^{-29} g cm^{-3}, and was at a temperature of 10^9 °K owing to the decay of the neutrons into protons and electrons. Gould and Burbidge went on to show that at the assumed density the intergalactic temperature would have to be no more than 10^6–10^7 °K to give the required background of X rays (see also Field and Henry 1964). However, the calculation appears to be not very sensitive to the assumed density and, if an intergalactic temperature of 10^6–10^7 °K could be justified, it seems that a bremsstrahlung process might well account for the background with considerably lower intergalactic densities. However, such a high-energy intergalactic medium seems improbable in relativistic cosmology (cf. § 6.5).

The apparent isotropy of the γ-*ray* background also points strongly, for the same reasons as for the X-ray background, to an extragalactic origin. If cosmic rays have the same energy density in intergalactic space as in the Galaxy (10^{-12} erg cm^{-3}), and if the density of H atoms in space is at least 10^{-5} cm^{-3}, then it appears from calculations by Ginzburg and Syrovatskii (1965) that π^0 decay, bremsstrahlung and the inverse Compton effect could lead to a γ-ray background ($\nu > 100$ MeV) near the observed upper limit (equation (2.21)). However, extragalactic cosmic rays have low probability (§ 6.5) except possibly for cosmic rays of energy greater than 10^{15} ev, which even as observed in the Galaxy may well be of extragalactic origin. Such articles, however, would have insufficient flux for these mechanisms to be significant.

It does not seem possible that the integrated emission of normal galaxies could lead to the observed background. The above-mentioned authors have shown that the most potent mechanism, the inverse Compton effect between cosmic-ray electrons and thermal photons in normal galaxies, would lead to a typical galactic emission equivalent to 10^{38} erg sec^{-1}. This therefore gives a γ-ray background about equal to the X-ray background due to normal galaxies (6.32) and is therefore less than the observed order of magnitude for γ-rays by a factor of about 150. Again, however, objects such as quasars may emit γ rays as intense or even more so than their X-ray emission, so that the previous argument for the X-ray background due to exploding objects would lead in an analogous manner to a γ-ray density of nearly the right magnitude. Experiments of great potential importance for cosmology are currently being made in order to detect γ-ray photons from known quasars (Porter 1965, unpublished).

6.5. *Extragalactic cosmic rays and intergalactic matter*

It is clear that if extragalactic cosmic rays exist then this is not due to any residue of primeval high-energy particles. Thus in the $\alpha\beta\gamma$ theory the typical particle kinetic energy at $t = 670$ sec (equation (6.22)) corresponds to $kT \simeq 5 \times 10^4$ ev, so that the present energy density of such particles would be utterly negligible.

It is commonly believed that the cosmic rays of energy exceeding 10^{15} ev are unlikely to arise in the Galaxy. Not even supernovae seem to be able to produce particles of this energy at the observed intensity. At 10^{15} ev the spectrum of cosmic-ray energy steepens, suggesting that a different mechanism may be operating. Once again it is plausible to suggest that exploding galaxies and quasars may supply the necessary energy with the observed degree of isotropy. The existence of the γ-ray background supports this conclusion.

The fact that electrons constitute only about 0·1% of the particle number in cosmic rays of this energy indicates that the original electrons may have lost their energy in highly enervating processes which the protons have escaped. In the case of synchrotron radiation, bremsstrahlung and inverse Compton effect the electrons would lose 10^6 times more energy than the protons (proportional to the square of the mass ratio). Furthermore, cosmic rays appear to have a considerably higher proportion of heavy nuclei than other cosmic material, suggesting that they originate in processes sufficiently catastrophic to produce the necessary high densities and temperatures for the formation of such nuclei. Exploding galaxies would supply such processes for the higher energies.

On the other hand, it is evident, by an argument similar to that given in § 6.4 for the x-ray background, that the total cosmic-ray energy density in intergalactic space due to exploding objects can only be a small fraction of the value observed in the Galaxy. This depends on the fact that over a period H^{-1} the total nuclear energy emitted (mostly kinetic), regarded as the equivalent of electron energy radiated at a rate as high as 6×10^{46} erg sec^{-1} (or 10^{61} erg in 5×10^{6} year) from objects of space density $3{\cdot}3 \times 10^{-81}$ cm^{-3}, can lead to a background density of only 10^{-18} erg cm^{-3} at the most, compared with the galactic total density of 10^{-12} erg cm^{-3}.

Thus, in relativistic cosmology at least, it seems very unlikely that extragalactic cosmic rays are present except at the the highest energies, and these would have a low total density. In a steady-state universe, in which there is continuous creation, different criteria will of course apply; but here also, if the creation of matter takes place in objects such as quasars, the above argument would again lead to a low intergalactic cosmic-ray density. Further analysis of cosmic rays on these lines may therefore have cosmological significance.

Several tests for evidence of *intergalactic matter* in the form of neutral hydrogen have been made in recent years. Field (1962), Davies and Jennison (1964) and Davies (1964) have examined the spectrum of radio emission from Cygnus A to see if there is any absorption at the 21 cm line by intergalactic neutral hydrogen. This requires scrutiny of the spectrum between 1420 Mc/s and 1345 Mc/s, the latter being the red-shifted frequency of the 1420 Mc/s emission line at the distance of Cygnus A. These authors concluded that no detectable absorption was in evidence, and in fact obtained an upper value for the density n_H (atoms cm^{-3}) of neutral hydrogen in intergalactic space, divided by its spin temperature T_s, in the form

$$\frac{n_H}{T_s} \leqslant 3{\cdot}9 \times 10^{-8}. \tag{6.35}$$

By postulating specific conditions according to given cosmological models Davies (1964) was able to go further. In the case of relativistic cosmology, if the conditions were such that the present intergalactic gas temperature could be restricted to a few degrees absolute (cf. equation (6.24)) and the gas were totally neutral, then the observations implied that

$$n_H \leqslant 9 \times 10^{-8}\,\text{cm}^{-3}. \tag{6.36}$$

This means that the present intergalactic density of matter $\rho_m(\text{ig})$ is subject to

$$\rho_m(\text{ig}) \leqslant 1{\cdot}4 \times 10^{-31}\,\text{g cm}^{-3}. \tag{6.37}$$

Since this value is rather less than the mean density of luminous matter it follows that the present total mass–energy density in an exploding model would be

$$\rho = 3{\cdot}2 \times 10^{-31}\,\text{g cm}^{-3} \tag{6.38}$$

approximately (equation (2.6)). It is interesting to note that this would imply an ever-expanding hyperbolic model, in agreement with earlier indications in the present section (§§ 6.1 and 6.3).

The results (6.36) and (6.37) would also be implied in a steady-state model which postulated continuous creation of cold hydrogen in intergalactic space. Such a model predicts $n_H \simeq 10^{-5}$ cm^{-3} (Hoyle 1948, McCrea 1951) and is therefore eliminated by the above results. On the other hand, the upper limit (6.35), applied to a steady-state model (or a relativistic model) in which the intergalactic gas was highly ionized, would produce no useful information since T_s would then be very high.

A different approach was made by Goldstein (1963) who tuned his receiver to a frequency band below 1420 Mc/s and therefore should have received red-shifted 21 cm radiation from a spherical shell of hydrogen in space, if such existed. The radius d of the shell would evidently be given by

$$d = \frac{c(\nu_1 - \nu_0)}{\nu_0 H} \tag{6.39}$$

where ν_1 is 1420 Mc/s and ν_0 is the tuning frequency of the receiver. In this way Goldstein derived a limit $n_H < 2{\cdot}1 \times 10^{-5}$ cm^{-3} for intergalactic hydrogen.

A somewhat related procedure has recently been suggested by Scheuer (1965) for the Lyman α radiation from quasars, and results on this basis have recently been announced by Schmidt (1965 c, unpublished) and by Gunn and Peterson (1965). In this case the emission to be detected is from the ultra-violet continuum beyond the Lyman α line in the quasar spectrum. During propagation through intergalactic space this radiation will be red-shifted through the Lyman α line and absorbed or scattered by any neutral hydrogen that exists. Scheuer points out that even a density $n_H = 10^{-9}$ cm^{-3} would very effectively scatter the radiation, so that none should be detected from the quasar if n_H equals or exceeds this value. In the case of the quasar 3C 9, whose red shift is $z = 2{\cdot}012$, the Lyman α line is red-shifted from its emission position at 1215 Å to 3660 Å in the blue. One therefore has to look for a continuum in the violet side of 3660 Å in the received radiation. Absorption up to any given observed frequency ν_1 (greater than ν_0, where ν_0 corresponds to 3660 Å) in the received spectrum would take place within a sphere in space centred at the source and of radius d given by (6.39). Here we neglect any variations in H during the light time.

As a result of this observation of the spectrum of 3C 9, which in fact shows the continuum referred to, Schmidt, and Gunn and Peterson believe that n_H may be as low as

$$n_H = 2 \times 10^{-12}\,\text{cm}^{-3} \tag{6.40}$$

corresponding to an intergalactic gas density of

$$\rho_m(\text{ig}) = 4 \times 10^{-36}\,\text{g cm}^{-3}. \tag{6.41}$$

Once again, of course, this interpretation depends on the assumption that intergalactic hydrogen is not ionized, at a temperature exceeding 10^4 °K. From foregoing evidence this possibility is unlikely in relativistic cosmology, but to obtain a firm decision on this matter which affects so many issues must be one of the most important aims of future observational programmes.

6.6. *Remarks on observational tests of steady-state cosmology*

When the steady-state theory was first put forward, it was emphasized that the theory made definite predictions and was therefore very vulnerable to observational tests. Whatever the ultimate fate of the theory, it cannot be denied that attempts to test its predictions have led to important work in observational astronomy, much of which has been described in the present section. The theory also prompted important investigations in theoretical astrophysics such as the work on nucleosynthesis described in § 8. How does the theory show up in the light of present observations ? With regard to this question a few general remarks may be in order.

Hoyle (1961) has pointed out that cosmological observations can be broadly classified into two categories. There are observations that refer to the part of the Universe lying *on* the observer's past light cone. The red-shift–magnitude relation, counting of radio sources, etc. are tests of this kind. The other class of observations is confined to the part of the Universe *inside* the past light cone of the observer. Such observations do not involve the very distant parts of the Universe, although they are of cosmological significance. For example, the nature of the intergalactic medium, ages of galaxies, etc. are observations of this type. The two types of observations will be described below in that order:

(i) *Observations on the past light cone.* As has been pointed out previously, the observations of red shift and magnitude indicate that at face value the deceleration parameter q is around 1 (figure 5). For the steady-state theory

$$q = -\left(\frac{R\ddot{R}}{\dot{R}^2}\right) = -1. \tag{6.42}$$

The test therefore goes against the steady-state theory. The steady-state theorists point out, however, that the present scatter in the red-shift–magnitude diagram, and the possible errors arising from a selection effect in favour of bright galaxies, could very well mask the real effect. This test could be made more reliable if the observations were confined to a sub-class of galaxies which do not show a large scatter in their absolute brightness. Sandage has suggested that massive ellipticals dominating clusters of galaxies provide such a class of galaxies (cf. Baum 1961 b).

The counting of radio sources provides another test, which goes against the steady-state theory. The Cambridge survey (Ryle and Clarke 1961) shows that the $(\log N, \log S)$ curve has a slope $-1{\cdot}8$ down to about 2 flux units. The steady-state model predicts a slope which starts at $-1{\cdot}5$ and gets progressively less in magnitude. At the time of writing there is some doubt as to whether all the sources counted are extragalactic. If they turn out to be so the test must probably be considered decisive. However, Hoyle and Narlikar (1961, 1962 b) have pointed out that care must be taken in interpreting the radio data. Radio sources are rare phenomena and their theoretical count curve would show statistical fluctuations. These fluctuations would become enhanced if the radio source property is age-correlated. (Hoyle and Narlikar assumed that the probability of a galaxy being a radio source increased with age.) By carrying out a large number of hypothetical source counts on an IBM 7090 computer, Hoyle and Narlikar have shown that with varying assumptions the slope of the $(\log N, \log S)$ curve could fluctuate between the slopes of $-1{\cdot}2$ to $-2{\cdot}5$.

It may well be that the assumptions in the above counter example turn out to be wrong or inadequate. The example, however, illustrates an important point that is sometimes ignored in interpreting cosmological observations of the first type. How strictly should we take the homogeneity and isotropy implied by the cosmological principle? This question cannot be answered independently of the nature of the observations. If the observations refer to a certain property then the cosmological principle should be applied to regions large enough to make the property frequent. Thus the regions are larger in the case of radio sources than in the case of ordinary galaxies. This introduces fluctuations on a larger scale.

Recently Veron (1965, Conf. on Observational Aspects of Cosmology, Miami Beach, Florida) has examined the $(\log N, \log S)$ curve in detail for the optically identified sources from the revised 3C catalogue. He finds that the slope of the curve is $-1{\cdot}55$ for radio galaxies and $-2{\cdot}2$ for quasi-stellar radio sources. The radio galaxies are nearby objects with red shifts not exceeding 0·5. Hence the observed slope is not inconsistent with any cosmological model discussed here. If the large red shits of the quasi-stellars arise from the expansion of the Universe, so that they are really distant objects, the slope $-2{\cdot}2$ would disprove the simple steady-state model. On the other hand, if the quasi-stellars turn out to be comparatively nearby objects, this test loses its cosmological significance. This possibility cannot be excluded at present (cf. § 9.1).

(ii) *Observations inside the past light cone.* Considerations of nucleosynthesis described in § 8 show how the chemical composition of stars changes with age. This enables us to estimate the age of the stars in our Galaxy. The Galaxy should be at least as old as the oldest stars in it. Estimates of the age of our Galaxy (Fowler and Hoyle 1960) range from $1\text{–}1{\cdot}5 \times 10^{10}$ years. Lower estimates are also available (Dicke 1962). A figure of the above magnitude may prove embarrassing to most big-bang cosmologies. These cosmologies predict an age of the Universe (i.e. time elapsed since the big-bang) less than $H^{-1} \sim 10^{10}$ years. Only a revision in the nuclear dating process or in the observed value of H (or both) can resolve the discrepancy. In the steady-state model galaxies of arbitrarily large ages can exist. However, over a large region the age distribution of galaxies follows the law

$$Q(\tau)\,d\tau = (\text{constant}) \times \exp(-3H\tau)\,d\tau \tag{6.43}$$

where $Q(\tau)\,d\tau$ denotes the number of galaxies with age between $(\tau, \tau + d\tau)$. The average age of galaxies in the Universe is therefore only $\frac{1}{3}H^{-1}$, with the young galaxies being much more numerous than the old ones. Present observations indicate that old galaxies exist in much more numerous proportions than indicated by (6.43).

Hoyle and Narlikar (1962 a) have pointed out that the average age should not be taken as typical of the age of the galaxies we observe around us. If galaxies form in large groups together, large age-correlated regions will exist. We may be living in such a region where the typical age may be $\frac{4}{3}H^{-1}$ rather than $\frac{1}{3}H^{-1}$. Further, galaxies may evolve through a sequence irregular $\rightarrow$ spiral $\rightarrow$ elliptical and this would affect the simple distribution (6.43) for different galaxies. It is clear nevertheless that the present age distribution cannot be reconciled to (6.43) in the simple steady-state theory. A proper assessment of the data must wait until a good theory of galaxy formation is available.

The nature of the intergalactic medium is likely to throw important light on the cosmological problem. As indicated in the last section the nature of creation of matter in the steady-state theory affects the intergalactic medium. In the big-bang cosmologies the violent activities of the bang would by now have quietened down and we would expect very few disturbances in the intergalactic medium. As indicated in §§ 6.3, 6.4 and 6.5 the x-ray, γ-ray and cosmic-ray astronomy will prove to be important in this field.

7. Local effects of cosmological significance

7.1. *Introduction*

It is sometimes thought that cosmology is concerned with the distant parts of the Universe and therefore has very little to say about anything of 'local' importance. This is indeed not so. The question 'why is the sky dark at night?' has cosmological significance, and led to the well-known Olbers paradox. A satisfactory solution of the paradox, as was mentioned in § 3, came only when an account was taken of the expansion of the Universe. The 'coincidence' that the local inertial frame is the same as one in which the distant parts of the Universe are non-rotating, which led to Mach's principle, is another example. In this section we shall describe effects of a similar nature—effects which are local in their observational character but whose explanation takes us into the considerations of cosmology. In particular, we shall be concerned with questions of universal neutrino degeneracy and the arrow of time.

7.2. *Universal neutrino degeneracy*

The Olbers paradox refers to the flux of photons received from all over the Universe. Photons are particles of zero rest mass travelling with the speed of light. There are other particles with similar properties to be found in the Universe, the most important of them being the neutrinos. Is there an Olbers paradox for neutrinos? This question was raised by Weinberg (1962 a). He pointed out that neutrinos differ from photons in that they obey Fermi rather than Bose statistics. The number of neutrinos that can be packed up to any energy level is therefore limited and the extent of neutrino degeneracy will affect the reactions involving them, especially the β-decay reactions. Thus it is the number rather than the intensity of neutrinos that is of importance.

Let $N'(t)$ denote the numbers of neutrinos emitted in unit time per unit *coordinate* volume in a Robertson–Walker universe. Then a simple calculation, making use of (3.11)–(3.13) or (4.33) gives the number of neutrinos per unit volume at time $t = t_0$. If we neglect absorption and any primeval component, (4.33) yields

$$n(t_0) = R^{-3}(t_0)\int_{-\infty}^{t_0} N'(t)\,dt. \tag{7.1}$$

If we neglect evolution of emission, in the big-bang cosmologies

$$N'(t) = \text{constant} = N_0' \text{ (say)}$$

for $t > 0$, and $N'(t) = 0$ for $t < 0$. Hence $n(t_0)$ is finite. In the steady-state cosmology because of continuous creation of matter $N'(t) = N_0' e^{3Ht}$. In this case also the integral (7.1) converges. The integral diverges in oscillatory cosmologies unless there is an absorption of neutrinos at some stage of the oscillation.

Thus Olbers' paradox for neutrinos does not arise in the big-bang and steady-state cosmologies and can be avoided in the oscillatory cosmologies if there is an adequate absorption of neutrinos.

Are the neutrinos degenerate? If so, to what extent? This has been considered by Weinberg (1962 b) in the following way. Let $\Omega(W, t)$ denote the rate of emission of neutrinos of energy W and $\Lambda(W, t)$ their absorption rate at time t. Let $X(E, t)$ denote the degeneracy factor, i.e. the proportion of levels filled up to energy E. Then Weinberg gets the following general result

$$X(E, t_0) = \int_{-\infty}^{t_0} \Omega\left(E\frac{R(t_0)}{R(t)}, t\right) \exp\left[-\int_{t}^{t_0}\left\{\Lambda\left(E\frac{R(t_0)}{R(t')}, t'\right) + \Omega\left(E\frac{R(t_0)}{R(t')}, t'\right)\right\} dt'\right] dt. \tag{7.2}$$

From this he finds, quite generally, that at very low energy ($E \to 0$) exactly one-half of all levels of neutrinos and antineutrinos is filled in the steady-state and oscillating cosmologies. The proportion between neutrinos and antineutrinos would depend on the actual processes going on in the Universe which emit (and absorb) these particles.

What is the Fermi energy E_F in different cosmological models? Weinberg's calculations show that E_F has ridiculously low values for the steady-state and big-bang cosmologies. In the former $E_F \sim \exp(-10^{36})$ Mev while in the latter $E_F \sim 10^{-36}$ Mev or $\sim 10^{-24}$ Mev depending on whether at the time of the big-bang the energy in the Universe was mostly in the form of radiation or matter.

In the case of oscillating cosmologies the value of E_F turns out to be not so low:

$$E_F \simeq 5\left(\frac{R_{min}}{R_0}\right) \text{Mev}$$

where R_{min} is the minimum radius during the oscillations and R_0 the present radius.

Thus from an observational point of view the steady-state and the big-bang cosmologies do not provide any significant result. In the oscillating models, a sufficiently high E_F would significantly affect the local β-decay experiments done in the laboratory. If such an effect is noticed, it would indicate, according to Weinberg, an oscillating universe. The absence of an effect would put an upper limit on R_{min}. Such an effect would show up in the form of an apparent violation of energy conservation in β decay.

7.3. *The direction of time*

The theory of relativity tells us that at each point in space–time we can draw two light cones, one into the 'past' and one into the 'future'. However, as far as relativity is concerned, the distinction between past and future is purely conventional. Neither relativity nor any other basic law of physics known today tells us how to make this distinction; these laws are all symmetric with respect to time. Yet our own experience shows that we can influence (by interactions, signals, etc.) only one-half of the light cone, which we call the 'future' light cone. Can this observation be accounted for *within* the framework of physics as we know it today?

The distinction between the past and the future is made by the so-called irreversible phenomena. These can be found in thermodynamics, electrodynamics and cosmology. The increase of entropy in thermodynamics, radiation by electric charges in electrodynamics, and the expansion of the Universe in cosmology provide an *observational* means of distinguishing between the past and the future light cones, and lead to the 'arrow' of time. Can there be any connection between these different time-asymmetric phenomena? Many physicists believe that the connection, if at all, is very weak. While such a point of view may be correct, it is not of interest from the point of view of the present discussion. The other point of view tries to establish a *strong* correction, with the aim of explaining all of them in terms of cosmology. The work of Wheeler and Feynman (1945, 1949), Hogarth (1962), Hoyle and Narlikar (1963 b) and Narlikar (1962)) makes steps in that direction. The general ideas underlying this point of view are described below.

Consider the phenomenon of the radiating charge. The electromagnetic field surrounding the charge is described by a particular solution of Maxwell's equations, the so-called retarded solution. This solution denotes propagation into the future half of the light cone. The corresponding solution propagating into the past light cone—the advanced solution—is rejected because it does not agree with experience. Thus, although physical laws (in this case, Maxwell's equations) are time symmetric, Nature selects a time-asymmetric solution and thus brings in the arrow of time. Why?

Wheeler and Feynman considered this question in the following way. They began by reformulating electromagnetic theory in terms of direct interactions between charged particles. Such a theory does away with the notion of independent fields—electromagnetic fields arise solely because there are sources (the electric charges) present. The interaction between two particles A and B will be along light rays joining points on their world lines. It is inevitable, therefore, that interactions travel both backwards and forwards with respect to the time coordinate. A retarded interaction from A to B will have an advanced reaction from B to A. How are we to get rid of the advanced solutions which necessarily arise on an equal footing with the retarded solutions? If in the conventional Maxwell theory $F^{(a)}_{\text{ret}}$ denotes the retarded solution due to particle A, the corresponding solution in the Wheeler–Feynman theory can be written as

$$F^{(a)} = \tfrac{1}{2}(F^{(a)}_{\text{ret}} + F^{(a)}_{\text{adv}}). \tag{7.3}$$

However, Wheeler and Feynman demonstrated that whenever an electric charge moves, its retarded solution excites other charges and their motion contributes an *advanced* reaction. Thus the whole Universe must be taken into account, and it becomes essential to know about the nature of the cosmological model. For a static Minkowski universe, with a uniform distribution of charges, the reaction is

$$R^{(a)} = \tfrac{1}{2}(F^{(a)}_{\text{ret}} - F^{(a)}_{\text{adv}}). \tag{7.4}$$

Thus the total electromagnetic disturbance near A is

$$F^{(a)}_{\text{total}} = F^{(a)} + R^{(a)} = F^{(a)}_{\text{ret}} \tag{7.5}$$

as observed. Equation (7.4) is precisely the radiation reaction obtained by Dirac (1938).

Thus, at first sight, it appears that because of the interference from the Universe a time asymmetry arises in a basically time-symmetric theory. However, if, throughout the above argument, we reverse the sign of the time coordinate, we are back to a consistent *advanced* solution. Wheeler and Feynman recognized this difficulty and argued that the distinction between the two self-consistent solutions is made by thermodynamics. On the hypothesis of equal *a priori* probabilities one could reject the advanced solution as being extremely 'unlikely'. Wheeler and Feynman thus connected the electrodynamic arrow of time to the thermodynamic one.

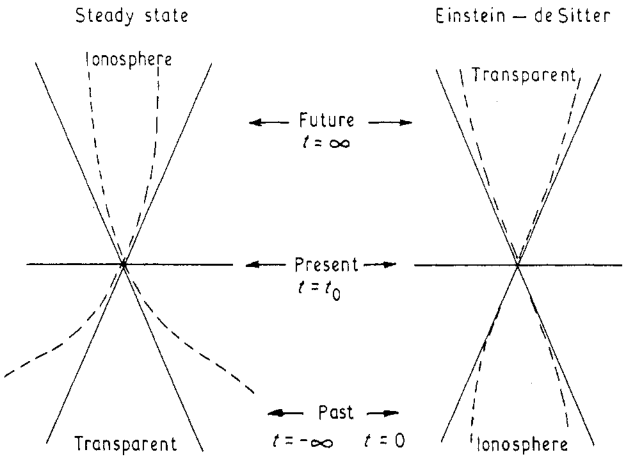

Figure 10. This represents the Wheeler–Feynman theory: in the steady-state cosmology the future acts as a reflecting ionosphere and is able to provide the correct radiation reaction to retarded signals. The past is essentially transparent to advanced signals. The retarded signals are consistent, not the advanced ones. The situation is reversed in Einstein–de Sitter cosmology.

This point of view was criticized by Hogarth, who pointed out that an appeal to thermodynamics was unnecessary. The asymmetry between the advanced and the retarded solution could be linked to the cosmological arrow of time which was ignored in Wheeler and Feynman's calculations in a static universe. In the static universe a change of sign of the t coordinate keeps the Universe invariant, but changes retarded solutions to advanced ones, or vice versa. An expanding universe, on the other hand, is *not* invariant under this transformation.

Calculations on similar lines as above have been done by Hogarth and by Hoyle and Narlikar on the various cosmological models described in the preceding sections. The results are striking in nature. It turns out that in the steady-state model retarded and not advanced solutions are self-consistent. In the Einstein–de Sitter and some other big-bang models, the advanced solutions are self-consistent and the retarded solutions are not. The reason for this can be seen as follows. For retarded solutions to be consistent, there must be enough matter per unit *proper* volume on the future light cone to send back the adequate reaction given by (7.4). This condition is easily satisfied in the steady-state theory but not in the evolutionary cosmologies. The opposite is true in the case of advanced solutions (see figure 10).

To summarize, if the Wheeler–Feynman approach is adopted, the laboratory observation of 'retarded' solutions is sufficient to disprove all evolutionary cosmologies which predict 'advanced' solutions. The test is in agreement with the predictions of the steady-state theory. It is possible to argue, however, that the test is dependent on the *assumptions* of the Wheeler–Feynman electrodynamics. Such a test would not be possible if the conventional Maxwell electrodynamics is adopted.

The above method links electrodynamics and cosmological time arrows. Can the two together imply the thermodynamic arrow? Hoyle and Narlikar have expressed the hope that this may be possible—although no explicit demonstration has been given. An expanding universe acts as a continuous 'sink' and local thermodynamics proceeds in such a way as to fill the sink. Retarded interactions ensure that the sink will tend to fill.

Once the three arrows of time are correlated the basic question about why time has an arrow disappears. For, one can say that we grow older because the Universe expands. An equivalent statement (mathematically) would be that we grow younger because the Universe contracts.

8. Origin of the elements

8.1. *Formation of the elements in the early Universe*

A clue to the origin of the chemical elements first became apparent when it was realized, owing to the work of Bethe, von Weizsäcker and Critchfield in 1938, that the source of the tremendous radiation from stars lay in the conversion of hydrogen to helium in their interiors. It was recognized that in the stars, at central temperatures of order 10^7 °K, the high-velocity collisions that must take place between nucleons would overcome their mutual repulsion and lead to the formation of more complex nuclei. However, at that time it was thought that stellar temperatures were insufficient to generate, starting from hydrogen, elements more massive than helium. Indeed the formation of helium in stars was not thought to add substantially to the helium already existing. Rather there was a general belief that the apparent singularity in the history of the Universe, suggested by the Hubble expansion and general relativity theory, was associated with the creation of the main bulk of the elements. Here uniquely, it was thought, were to be found the conditions of high temperature and density necessary for the assembly of the elements from primitive matter.

The $\alpha\beta\gamma$ theory, associated with the names of Alpher, Bethe and Gamow and others, was first outlined by Gamow (1946 a), who suggested that the formation of the elements was essentially a disequilibrium process which took place under the rapidly changing conditions in the first few minutes of the life of the Universe. As shown in §4.3, Gamow's tenet, that the radiation energy would have predominated over matter energy for sufficiently small t, leads to a time dependence of radiation energy given by (4.26), or (6.21), and of radiation temperature given by (6.22). We see, for example, that at $t = \frac{1}{2}$, one half-second from the singularity, the radiation density is $1{\cdot}8 \times 10^6$ g cm^{-3}. At $t = 1\frac{1}{2}$ the density has fallen by a factor of nearly 10. Thus, Gamow concluded, the conditions necessary for rapid nuclear reactions existed for only a very short time, perhaps half an hour. Therefore, the element abundances could not depend on equilibrium rates of formation as assumed in earlier theories. Rather the process was one governed by the rapidly decreasing density and temperature, which led to the created abundances being 'frozen in'.

According to nuclear physics, complex nuclei could not hold together at the extreme temperatures obtaining in the first few moments of the life of the Universe. By (6.22) we see that the temperature was $1{\cdot}5 \times 10^{12}$ °K when $t = 10^{-4}$ sec, and by (6.21) the total density was then of nuclear order, that is about $10^{13}\,\mathrm{g\,cm^{-3}}$. Individual particles would have had an energy of about 100 Mev. Alpher *et al.* (1948) and Gamow (1948) therefore considered the Universe from the stage at about $t = 20$ sec when the temperature had dropped to about 3×10^{9} °K. (The start of nucleogenesis is slow so that the final abundances are not very sensitive to the choice of initial time, and this was varied in later refinements of the theory.) At this temperature the total density was about $10^{3}\,\mathrm{g\,cm^{-3}}$ and the energy of particles about 4×10^{5} ev. As shown in a later paper (Alpher *et al.* 1953), when this temperature is reached, equilibrium processes involving pair formation, proton–neutron reactions with neutrino emission and absorption, meson formation and annihilation would have died out and only neutrons, protons and radiation would remain. The neutrino component of radiation would have become frozen in, but matter would still be in thermodynamic equilibrium (kinetically) with photons.

The essence of the $\alpha\beta\gamma$ theory is that, equilibrium conditions having terminated, the neutrons now begin to decay into protons and electrons. The increased numbers of protons would then capture some of the remaining neutrons and form more complex nuclei, deuterons, helium and so on. At each stage nuclei would capture more neutrons at a rate dependent on their cross sections for radiative capture (which increase with atomic weight A up to $A = 100$ and afterwards remain approximately constant). Nuclei would stabilize by β emission where necessary. In this way the nuclear abundances would be built up until all neutrons were captured or had decayed. This stage was reached after a total time of about 30 min (the free neutron half-life being about 12 min).

The only disposable parameter in the $\alpha\beta\gamma$ theory was the density of matter at the start of free neutron decay. This was adjusted to give the correct abundances. Too high a density leads to too rapid capture of neutrons by protons leaving very little hydrogen when the process is finished. Too low a density leaves too much hydrogen and very little helium. The choice for ρ_{m} by Alpher and Herman (1950) is given by (6.23).

The cosmic abundances, as determined in the crust of the Earth, meteorites, Sun, stars, cosmic rays, interstellar gas and dust are roughly (by weight) 63% hydrogen, 36% helium and 1% heavier atoms. A good fit of the abundance curve (as given by Brown (1949) and more recently by Suess and Urey (1956)) was originally claimed by the authors of the $\alpha\beta\gamma$ theory (figure 11). However, there now appears a serious difficulty, acknowledged by the authors, in getting the build-up process past the atomic weights 5 and 8 (^{5}He and ^{8}Be). This is because the decay of these unstable isotopes is too fast to be accommodated by the neutron capture process. It seems therefore that unless some way can be found to cross these gaps then only the nuclei up to atomic weight 4 could be accounted for in this way.

A further shortcoming of the original $\alpha\beta\gamma$ calculations was that they ignored the proton content resulting from equilibrium processes at temperatures above

3×10^9 °K, assuming free neutron decay as the only source of protons. This gave a proton–neutron ratio $p : n \simeq 1 : 7$ at the start of nucleogenesis. In their paper already mentioned, Alpher *et al.* (1953) included the equilibrium contribution, following Hayashi (1950), and found a substantially different value, viz.

$$p : n \simeq 4{\cdot}5 : 1.$$

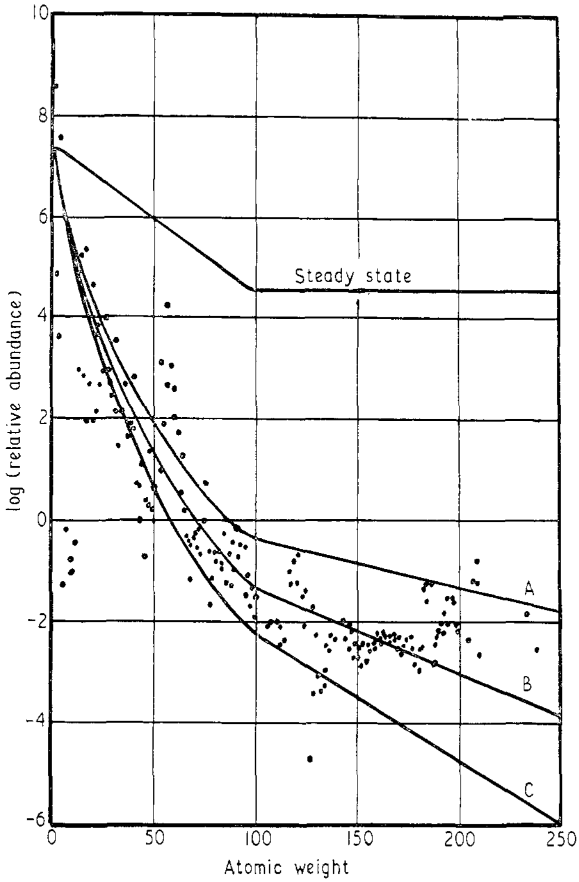

Figure 11. The fit of the neutron capture theory to the observed abundances of the elements (as given by Alpher and Herman (1950), following Gamow (1948)) is shown. The points for the element abundances indicate the ratios relative to silicon taken as 10 000 (logarithmic value 4). Ignoring the difficulty of passing $A = 5$ the theoretical curve B is the best fit, while A and C show the sensitivity of fit. Curve A has a smaller initial concentration of neutrons than curve B in the ratio 5 : 8. Curve C has a greater in the ratio 13 : 8. The curve marked 'steady state' indicates the equilibrium ratios that would result from a very high initial concentration of neutrons or an indefinite prolongation of the neutron capture process, i.e. ignoring cut-off due to neutron decay and expansion of the Universe. (From Alpher and Herman 1950.)

These authors went on to calculate the ultimate helium–hydrogen number ratio if all these neutrons were used to make helium only. This would evidently be

$$\frac{\mathrm{He}}{\mathrm{H}} = \frac{n}{2(p-n)} \simeq 0{\cdot}14. \tag{8.1}$$

It follows that the percentages *by weight* of helium and hydrogen present in cosmic matter today, neglecting helium manufactured in stars, would be

$$\left.\begin{array}{ll} \mathrm{He} & 36\% \\ \mathrm{H} & 63\% \end{array}\right\} \qquad (8.2)$$

approximately, in good agreement with observation.

It is a notable fact that it appears impossible to account for the present helium–hydrogen ratio by manufacture of helium in stars. As recently pointed out by Hoyle and Tayler (1964), the latter process falls short of the required production by a factor of about 14, i.e. it yields a number ratio He : H $\simeq$ 0·01. Hoyle and Tayler therefore argue that either a state of ultra-high density and temperature existed universally in the past, or else such densities and temperatures have been frequently realized in especially massive stars, such as have been postulated by Hoyle *et al.* (1964) to account for quasars.

Thus it appears that the explanation of the helium–hydrogen abundance is of the greatest importance for cosmological theories. Quite apart from this question, it seems that the origin of the heavier elements must necessarily be accounted for in various kinds of star, and this stellar synthesis of the elements will now be described.

8.2. *Nucleosynthesis in stars*

8.2.1. *The Burbidge, Burbidge, Fowler and Hoyle theory.* The single-event theory of nucleosynthesis described above encountered one great difficulty arising from the fact that no stable nucleus of atomic weight 5 exists. Gamow recognized this difficulty and admitted that "the lion's share of the heavy elements may well have been formed later in the hot interior of stars". Thus, for synthesizing all elements heavier than helium, it seems that the big-bang explosion is not of any help.

Nuclear processes in stars have been considered by Bethe (1939), Salpeter (1957), Cameron (1955) and others. A systematic and comprehensive account of nucleosynthesis in stars was, however, first given by Burbidge, Burbidge, Fowler and Hoyle (1957, to be referred to as BBFH). Their work was partly motivated by the steady-state cosmology. There is no superdense phase in steady-state theory. Creation of matter is always going on, essentially in a simple form such as neutrons, or electrons, or electrons and protons. The building up of complex nuclei from these simple ones cannot take place all over the Universe. Conditions suitable for such a process could, however, be found inside stars—Nature's own thermonuclear reactors. BBFH examined this problem. We shall describe their work in the rest of this section.

8.2.2. *Onset of nucleosynthesis.* Nucleosynthesis begins with the formation of a star. As stellar material, predominantly hydrogen, begins to contract under its own gravitation, it begins to heat up. A stage is reached when the temperature is high enough to trigger exothermic nuclear reactions, beginning with the conversion of hydrogen into helium. The heat generated by the reactions produces pressures which support the 'weight' of the star. As a result the star is in a relatively stable state for a long time—and is said to be on the 'main sequence' in the Hertzsprung–Russell (HR) diagram†. The actual nuclear processes can be described as follows.

† The Hertzsprung–Russell diagram plots the apparent magnitude (luminosity) of stars against their colour index (surface temperature).

Nuclear reactions begin in a star like the Sun when the central temperature reaches about 10^7 °K and the central density is as high as about 10 g cm^{-3}. The hydrogen → helium conversion takes place through the so-called proton–proton chain (p–p chain) which may operate in a number of ways described below:

$$^1\mathrm{H}(\mathrm{p}, \beta^+\nu)\,^2\mathrm{D}\,(\mathrm{p}, \gamma)\,^3\mathrm{He}\,(^3\mathrm{He}, 2\mathrm{p})\,^4\mathrm{He}$$

$$\hookrightarrow (^4\mathrm{He}, \gamma)\,^7\mathrm{Be}\,(\mathrm{e}^-, \nu)\,^7\mathrm{Li}\,(\mathrm{p}, \alpha)\,^4\mathrm{He}$$

$$\hookrightarrow (\mathrm{p}, \gamma)\,^8\mathrm{B}\,(\beta^+\nu)\,^8\mathrm{Be}^*\,(\alpha)\,^4\mathrm{He}$$

where $A(B, C)\,D \equiv A + B \rightarrow C + D$, $A^*(\alpha)\,B \equiv A^* \rightarrow \alpha + B$, etc.

The overall result is $4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He}$ with the emission of $4{\cdot}3 \times 10^{-5}$ erg of energy. Neutrinos of various energies are emitted during the chain. The process is believed to have been going on inside the Sun for about $4{\cdot}5 \times 10^9$ years, and may well go on as much longer. The time scales depend on the masses of stars, but are generally of this order. The helium formed in the central regions does not get mixed with the outer envelope of hydrogen. So as the star evolves the helium core grows in size. The temperature, about 10^7 °K, is not high enough to overcome the Coulomb barrier between two helium atoms, and hence to fuse them. As no more hydrogen is left in the core, the lack of nuclear reactions makes it contract under gravitation. This leads to generation of heat and a consequent rise in the temperature of the envelope. The result is that the envelope expands. This means a larger surface area of the star is available to radiate its energy, and hence the surface temperature is lowered. The star departs from the 'main sequence' and takes up the so-called 'red giant' form.

8.2.3. *Helium burning*. When the core reaches temperatures of the order of 10^8 °K and densities of the order of 10^5 g cm^{-3}, the Coulomb barrier between helium atoms can be surmounted. Two nuclei might be expected to form ^{8}Be. However, ^{8}Be is unstable and breaks up into two α particles. This presented a difficulty for a long time. It was resolved when it was realized that ^{8}Be could combine with another ^{4}He to form ^{12}C. Because of the low equilibrium concentration of ^{8}Be ($\sim 10^{-10}$ at 10^8 °K), Hoyle (1954) pointed out that only a 'resonant' reaction could save the situation. Laboratory experiments (Cook *et al.* 1957) have confirmed this prediction. An excited state of ^{12}C does exist with the properties predicted by Hoyle. The triple α process, as it is called, can therefore work inside red giants and make carbon, bypassing Li, Be and B. The overall process is described by

$$3\,^4\mathrm{He} \rightleftharpoons {}^{12}\mathrm{C}^* \rightarrow {}^{12}\mathrm{C}$$

where ^{12}C* denotes the excited state, and ^{12}C the ground state of the carbon nucleus. The ^{12}C often captures a helium nucleus and forms ^{16}O. Thus in stars there is no difficulty at the atomic weights 5 or 8.

Because of the small (0·07%) release of energy in the above process the red giant stage does not last long. Astronomical evidence suggests that the evolutionary track of the star on the HR diagram becomes reversed. Some stars may become unstable at this point and eject some of the carbon and oxygen into the interstellar regions. Others may remain stable to continue the nuclear process.

8.2.4. *The carbon–nitrogen cycle and the α process.* The ^{12}C and ^{16}O ejected by unstable stars will mix with the interstellar matter which may condense into a second generation star. In the new star the primary reaction $4\,^{1}H \rightarrow {}^{4}He$ can take place through the carbon–nitrogen cycle—now better known as the CNO bi-cycle. This is described by:

$$\rightarrow {}^{12}C\,(p,\gamma)\,{}^{13}N\,(\beta^{+}\nu)\,{}^{13}C\,(p,\gamma)\,{}^{14}N\,(p,\gamma)\,{}^{15}O\,(\beta^{+}\nu)\,{}^{15}N\,(p,\alpha)\,{}^{12}C \urcorner$$

or

$$\rightarrow {}^{14}N\,(p,\gamma)\,{}^{15}O\,(\beta^{+}\nu)\,{}^{15}N\,(p,\gamma)\,{}^{16}O\,(p,\gamma)\,{}^{17}F\,(\beta^{+}\nu)\,{}^{17}O\,(p,\alpha)\,{}^{14}N \urcorner .$$

When helium burning is over, the core contracts again and carbon and oxygen begin to burn. This burning results in the production of ^{20}Ne, ^{24}Mg, ^{28}Si, ^{32}S, etc., with the masses increasing by 4. For Z in the range 10–16, Coulomb repulsion is very strong and simple fusion is not possible. Instead, one nucleus photodisintegrates into a flux of α particles, which combine with another nucleus to form a heavier nucleus. Thus one ^{28}Si breaks into 7 α particles which combine with another ^{28}Si to form ^{56}Ni. ^{56}Ni decays (radioactively) into ^{56}Co and ^{56}Fe. This synthesis is called the α process. Thus nuclei up to the iron group can be built up by nucleosynthesis in hot stars.

8.2.5. *The e process.* The iron group nuclei are the most stable, in the sense that they have the maximum binding energies. The temperatures ($\sim 10^{9}$ °K) and densities ($\sim 10^{6}$ g cm^{-3}) inside the star at this stage are so high that all manner of nuclear reactions can occur. This allows one to deal with the mixture in a statistical way. BBFH found a good agreement between the observed and theoretical abundance values of the iron group elements on the assumption that their manufacture resulted from statistical equilibrium at a temperature about $3{\cdot}8 \times 10^{9}$ °K and density about 3×10^{6} g cm^{-3}. The equilibrium is between nuclei, free protons and free neutrons. The proton to neutron ratio is a free parameter which has the value 500:1 in the 'best fit' case. The agreement between observed and calculated abundances of iron group elements is very good, as shown by figure 12. Recently Tayler and Clifford (1965) have repeated these calculations in a more thorough manner with the help of an IBM 7090 computer.

In these calculations the reaction $e^{+} + e^{-} \rightarrow \nu + \bar{\nu}$ plays a critical role. This has not been observed in the laboratory. The astrophysical evidence strongly implies that the theory of weak interaction used in the above calculations must be substantially correct.

8.2.6. *The r, s and p processes.* The synthesis of elements heavier than the iron group is accomplished by means of the mechanism of neutron capture. Neutrons interact rapidly with the heavy elements at the low energies prevailing ($kT \sim 10$–100 MeV, $T \sim 10^{8}$–10^{9} °K), whereas charged particle reactions are inhibited by the Coulomb barrier. Neutron capture can take place in various ways. In the s process (slow process) the reactions are slow compared with the intervening β decays and produce nuclei mainly at the bottom of the valley of mass stability. The r process (rapid process), on the other hand, proceeds fast compared with the β decays and generally leads to the synthesis of a neutron-rich side of the mass valley. Apart from the r and s process there is a comparatively rare process which produces the proton-rich isotopes by exposure of the r-, s-process material to a high energy flux of photons or protons. This is the p process.

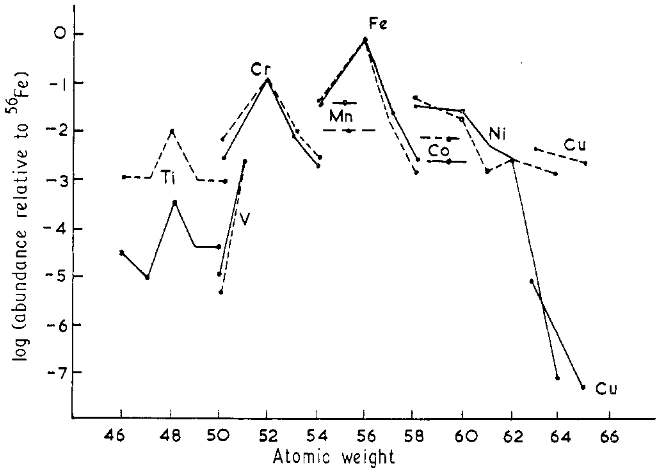

Figure 12. The diagram shows the iron-group abundances relative to ^{56}Fe nuclei produced in the e process (adapted from BBFH). Broken curve, solar abundances; full curve, calculated abundances.

To summarize, the element synthesis begins with the conversion of hydrogen to helium in the main sequence stars. The conversion of helium to carbon and oxygen takes place in the red giants. As the central core contracts and heats up, the α process builds up nuclei up to the iron group, provided the star is stable enough to go on. At any stage it may explode and eject all its material into the interstellar space. This is the supernova explosion. Prior to the explosion the e process has time to operate and determines the abundance of the stable iron group elements. The synthesis of heavier elements proceeds via the processes of neutron capture—the r and s processes—and rather infrequently by the p process. The s process generally takes place inside red giants which have been formed from galactic material containing H, C, O, Ne, Mg and the intermediate iron group elements. The r process takes place in the explosive envelopes or cores of supernovae.

Thus for element synthesis to work it is important that stars occasionally eject their material into the interstellar space and new stars are formed from the 'contaminated' interstellar material. This process could, in principle, work in any cosmology, provided sufficient time is available.

9. Some important observational and theoretical developments

9.1. *The recent discovery of quasi-stellar galaxies*

Extragalactic objects of large red shifts are of cosmological importance. The larger the red shift the further is the object according to the cosmological theories described here, and observation of distant objects provides us with information about the past history of the Universe. The recent discovery of quasi-stellar galaxies by Sandage (1965) has caused considerable excitement among cosmologists, because according to Sandage the quasi-stellar galaxies are very distant objects, some of which may have red shifts as high as 5!

Before considering Sandage's observations it is worth noting again the differences between the quasi-stellar objects (quasi-stellar sources and quasi-stellar galaxies) and ordinary galaxies. The quasi-stellar objects are several times brighter than galaxies and can be observed out to far greater distances. Moreover, the spectra of galaxies are of the absorption type, while those of quasi-stellar objects are of the emission type. For this reason the quasi-stellar objects, if they are really distant, are likely to prove more valuable than galaxies in testing the red-shift–magnitude relation described in § 6.

A plot of the first few quasi-stellar radio sources on a $(U-B, B-V)$ diagram indicated that these objects possessed a strong ultra-violet excess. (U, B, V here stand for the apparent magnitudes in the ultra-violet, blue and visual systems.) This provided a criterion for finding more (Ryle and Sandage 1964). However, it became clear that not all such objects could be identified with radio sources. Sandage called such objects interlopers, and later came to the conclusion that the blue objects catalogued by Haro and Luyten (1962) are of the same type. He gave various reasons for concluding that the Haro–Luyten objects are superluminous galaxies with large red shifts, and called them quasi-stellar galaxies. The quasi-stellar galaxies, unlike the quasi-stellar sources, are radio quiet down to flux level $S \sim 10^{-25}\,\mathrm{w\,m^{-2}\,(c/s)^{-1}}$ at 178 Mc/s. Their average absolute blue magnitude $\langle M_B \rangle$ is about -25 ± 2 and Sandage calculates that the average red shift z at $B = 22$ is about 5 ($H = 75\,\mathrm{km\,sec^{-1}\,Mparsec^{-1}}$).

Sandage arrived at these conclusions on the basis of (i) the colour diagram, (ii) the number count and (iii) the red-shift measurement. As mentioned above, the $(U-B, B-V)$ plot of Haro–Luyten objects with apparent visual magnitude $V > 14{\cdot}5$ resembles that of the quasi-stellar sources with known red shift. The plot for $V < 14{\cdot}5$ is different and resembles more that of stars. The plot of $\log N(m)$ against m_{pg}, (where $N(m)$ stands for the number of objects brighter than apparent magnitude m) has a change of slope near $m_{pg} \simeq 15$. For brighter objects the slope is 0·070 ($m_{pg} < 13$) and for fainter objects ($m_{pg} > 16$) it is 0·40 (see figure 13). The slope for extragalactic objects of negligible red shift is 0·60. Sandage therefore argued that the fainter objects are very distant and that the slope is reduced because of large red shift. He found that the slope would be of this order if quasi-stellar galaxies had the same $\langle M_B \rangle$ as the quasi-stellar sources and if we were living in a universe with the deceleration parameter $q \simeq 1$ (cf. § 4). Examination of the spectra of a few of the blue objects yielded large red shifts; in one case $z = 1{\cdot}241$ (figure 6). In other cases the results are still uncertain.

If Sandage's arguments are accepted the quasi-stellar galaxies would form an important constituent of the Universe and because of their large red shifts they must throw light on the cosmological problem.

At the time of writing the observational situation is not clear. Evidence is being put forward (Dent 1965, Kinman 1965) to suggest that the quasi-stellar objects may not all be distant, as imagined so far, and that the proportion of quasi-stellar galaxies may have been grossly overestimated. Burbidge and Hoyle (1966) have suggested that the possibility of the quasi-stellars' being local cannot be entirely excluded by present observations. Their large red shifts may be due to their being shot out of our own Galaxy or the nearby NGC 5128 at explosive speeds. According to their picture, if NGC 5128 is a source of these objects, some might have small blue shifts.

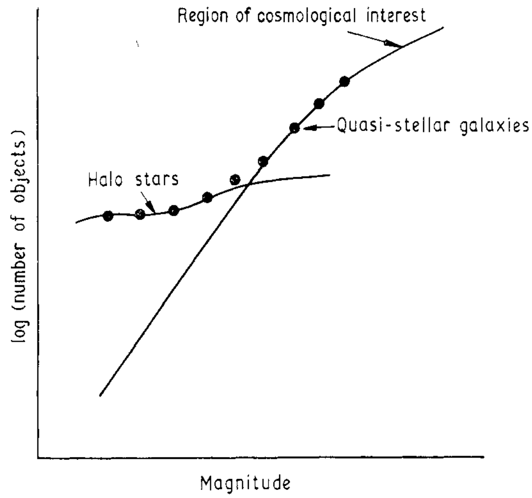

Figure 13. The number–magnitude relation for Haro–Luyten objects is shown in the diagram. The change of slope in the number–magnitude curve around $m_{pg} = 15$ is also shown. Sandage interprets this as a result of the mixing of halo stars with the distant quasi-stellar galaxies. (From Sandage 1965.)

Observations are in progress to explore in detail the various proposals, which are at present very tentative. From a theoretical point of view it is premature to draw any firm conclusions yet. As the following subsection will show the nature of quasi-stellar objects is very much an unsolved problem.

9.2. *Theories of quasi-stellar objects*

The very unusual appearance and behaviour of the quasi-stellar objects has prompted novel theories to account for them. At the time of writing there are at least half a dozen theories in this field with perhaps many more to come. In an article of this nature it is not possible to do justice to all theories in detail. We shall limit ourselves to describing the main streams of ideas on this subject.

A few months before the discovery of the first quasi-stellar sources, Fowler and Hoyle (1963 a) put forward the idea that masses of the order of 10^6–10^8 solar masses may accumulate at the centres of galaxies or in the intergalactic space. By following the evolution of such objects Fowler and Hoyle hoped to account for the energy production in radio galaxies. It soon became clear, however, that in objects of such large mass gravitational energies become far more important than the nuclear energies. This is because the former varies as M^2 and the latter as M, M being the mass of the object.

Early observations of 3C 48 and 3C 273 (Schmidt and Greenstein 1964) indicated that masses of the order of a million times the solar mass may be concentrated in volumes of the order of a cubic parsec in these objects. The massive object theory of Fowler and Hoyle was therefore a possible explanation of the quasi-stellar sources. Theoretically, as the object contracts under its own gravitation, energy of the order of the rest-mass energy of the object is available in the form of gravitation. The actual mechanisms considered (Fowler and Hoyle 1963 b, Gell-Mann, private communication, Fowler 1964) could, however, extract only a few per cent of this energy. Some of the mechanisms included gravitational radiation. By comparison, nuclear processes release only about $10^{-3}\,Mc^2$ energy.

Recently, Fowler and Hoyle (1965) have suggested that the light variation of the source 3C 273 could be understood in terms of a massive star of mass $M = 10^5$–$10^6\,M_\odot$. Such a star pulsates. As it contracts the nuclear reactions go faster and provide enough energy to make it bounce. (This is not possible in a more massive star because gravitation is much stronger.) The time period of pulsation is of the order of the light variation time observed by Smith (1963).

It has been argued that the massive object theories do not explain how the object was formed in the first place. Alternative theories have been put forward. Gold *et al.* (1963), Ulam and Walden (1964) and others have suggested that the quasi-stellars may arise out of collisions of a large number of ordinary stars (about 10^{10}) in a small volume (~ 1 parsec^3). The light variations are difficult to understand in such a model, although it does not have the difficulty of accounting for the formation of a single coherent massive object. Field (1964) has suggested that the quasi-stellar objects may be new galaxies in the process of formation. If the system has angular momentum large enough to prevent the formation of a single massive object, but small enough to prevent the formation of a disk, it will condense into a large number of stars on the main sequence with masses up to $100\,M_\odot$. Light variation may be accounted for by supernovae. This model accounts for the source 3C 48 but not for 3C 273 which is perhaps too young for the supernovae to have formed.

The discovery of quasi-stellars has aroused interest in another subject, the gravitational collapse. Assuming that a massive object has been formed with no appreciable angular momentum, will it continue to contract under its own gravitation ? Indications are that a sufficiently massive object will implode into a space–time singularity in classical general relativity. The nature of the singularity is similar to the singularity of relativistic cosmology. Is this singularity unavoidable ? Wheeler (1963) has argued that before a singularity is reached the matter becomes so dense that quantum theory will play an important role, and may cause the system to bounce. However, this has not been explicitly demonstrated. The singularity may be avoided by the use of the C field described in § 5 in connection with the steady-state theory (Hoyle and Narlikar 1964 a) (see also § 9.4).

If the sign of the time coordinate is reversed, an exploding object is obtained from an imploding one. Néeman (1965) has suggested that the quasi-stellar sources may be explained in terms of an object exploding through its Schwarzschild radius.

A similar explanation could apply to the object oscillating with the help of a C field, although such an object does not go inside its Schwarzschild radius. Novikov has suggested that the quasi-stellar sources may represent examples of delayed big-bangs in a universe which explodes from a singularity. Hoyle and Narlikar have tried to account for the energy production in the radio sources and the quasi-stellar sources in terms of explosive creation of matter in strong gravitational fields.

The nature of the quasi-stellar sources also depends on the question raised in § 9.1. Are they local or distant? The possibility that their large red shifts may be gravitational is more or less ruled out (Schmidt and Greenstein 1964). The short-term variations in the light and radio output of some of the sources are more difficult to account for in terms of the distant rather than the local theory. Recent observations indicate changes in the optical brightness of the source 3C 345 by about 0·4 magnitude in twenty days, accompanied by shifts in the spectral lines (Burbidge and Burbidge, private communication). Variations of such short time scale put a severe limit on the size of the source and therefore on its distance. An unambiguous solution of this question will take us a long way towards understanding the nature of the quasi-stellar objects.

9.3. *A new cosmological theory*

Towards the end of § 5 we described briefly Hoyle and Narlikar's (1966 a) picture of the steady-state universe in which creation of matter occurs in pockets of strong gravitational field. The overall creation rate is about 3×10^{-46} g cm^{-3} sec^{-1}, the value given by the homogeneous steady-state theory. It was hoped that the theory would also account for the energy of strong radio sources and of the quasi-stellar objects and for the production of high-energy particles observed in the cosmic rays. It turned out that high-energy particles could be created in very strong gravitational fields near a highly collapsed massive object. The calculated energy spectrum of particles was also in general agreement with that of cosmic rays. In one respect the theory fails altogether. The theoretical creation rate near mass M cannot exceed

$$Q = 16\pi f m_p{}^2 c^{-1} G^2 M^2 = 10^{15}\left(\frac{M}{M_\odot}\right)^2 \text{ erg sec}^{-1}. \tag{9.1}$$

In the above relation f is the coupling constant of the C field described in § 5 and m_p is the mass of a proton. For the Crab nebula $M \sim M_\odot$ and hence

$$Q \sim 10^{15} \text{ erg sec}^{-1}$$

—too low by a factor of about 10^{20}. For a radio galaxy, $M \sim 10^6$–$10^8 M_\odot$; again the total output falls short by the factor of about 10^{20}.

Hoyle and Narlikar (1966 b) suggested that the creation rate (9.1) could be brought in line with observations, if the coupling constant f were raised by a factor 10^{20}. This would explain all high-energy phenomena in astrophysics in terms of creation of matter. However, a glance at (5.9) shows that the cosmological aspects are radically changed. The effective Hubble constant is raised by a factor of about 10^{10} so that $H^{-1} \sim 1$ year. The mean density of matter in the Universe is raised to approximately 3×10^{-9} g cm^{-3}. Clearly we do not live in such a universe!

A closer look at the picture, however, shows that in such a universe where creation is going on at a rapid rate in pockets, the masses in the pockets would grow large enough to become unstable. Fragmentation into smaller masses would occur. This is an unstable situation. For, creation rate is brought down (because of the M^2 dependence) by fragmentation. This reduces the strength of the C field and this in turn reduces the creation rate even more. Thus in a particular region (of finite extent) creation would stop altogether. The region in question then expands as an evolutionary universe, without creation. The difference is that the starting conditions for the expansion are set by the steady-state universe with mean density of about 3×10^{-9} g cm^{-3}. The expansion is similar to that of a bubble in a dense medium. Hoyle and Narlikar argued that we live in such a bubble.

Such bubbles are likely to develop in the Universe in a number of places but not synchronously everywhere at once, since the pattern of inhomogeneities will not be the same everywhere. Depending on the starting conditions, an expanding bubble may either continue to expand indefinitely (as in the Einstein–de Sitter universe) or it may contract eventually (as in the cosmology with $k = +1$). In the latter case as it contracts the C-field strength rises and a stage may come when conditions are suitable for creation. In that case the bubble merges into the Universe. Another possibility is that the bubble bounces back and re-expands. This is possible because of the negative energy density of the C field. In the equations (5.5) and (5.6) if we put $C^i{}_{;i} = 0$ we get

$$\dot{C} = A/R^3, \quad A = \text{constant} \tag{9.2}$$

and the function R satisfies (in the case $k = +1$) the equation (cf. equation (4.8))

$$R^4 \dot{R}^2 = \frac{8\pi G}{3}(BR^3 - \tfrac{1}{2}fA^2) - R^4 \tag{9.3}$$

where $\rho = B/R^3$, $B = \text{constant}$. Thus as R is reduced $\dot{R}$ approaches zero. The bubble oscillates between two finite radii. (This is not possible in the conventional evolutionary cosmology because there is no term like the negative term $-\frac{4}{3}\pi GfA^2$ in (9.3) which causes the Universe to bounce.)

One of the attractive features of the theory is that it explains the formation of elliptical galaxies. With $H^{-1} \sim 1$ year, the mass of the 'observable' Universe is of the order of $10^{13} M_\odot$. Hoyle and Narlikar (1966 c) suggest that if an inhomogeneity is present in the form of a massive object, this restrains the expansion of the matter in a bubble, which falls back in the form of a galaxy. A mass of the order of $10^9 M_\odot$ will be able to restrain a total mass of $10^{12} M_\odot$ from expanding beyond the normal galactic dimensions. The figure of about $10^{13} M_\odot$ places an upper limit to the masses of the galaxies. A spherically symmetric expansion round the central object would produce a spherical galaxy. A deviation from spherical symmetry, produced by shear, can give rise to ellipsoidal shapes. The important result that emerges is that rotation plays a secondary role, and on this hypothesis the elliptical galaxies are not rotating. This should be checked by observations. The theory also predicts a light intensity distribution in elliptical galaxies which agrees with the observations.

According to the theory, spirals arise as a result of condensation whereas the ellipticals arise as a result of expansion. It may be possible to understand the variety of galactic forms as a result of the mixing of the two processes.

9.4. The singularity in relativistic cosmology

In § 4.1 it was shown that, when the cosmological constant Λ vanishes, those world models of general relativity which are spatially homogeneous and isotropic have a physical singularity in finite past time. These are the Robertson–Walker exploding models. More generally, as Robertson showed many years ago (Robertson 1933), if Λ is retained the homogeneous isotropic models (of metric equation (3.6) or (3.7)) have either (i) a singularity in the finite past and a finite upper radius, or (ii) finite lower and upper radii, or (iii) a singular or finite lower radius from which the model expands to infinite diffusion during an infinite time. There is also, of course, the unstable static Einstein universe, which with a slight disturbance would come under category (i) or (iii). As demonstrated by Tolman (1934) the finitely oscillating models of category (ii) can be eliminated for thermodynamic reasons. Also the models of category (iii) which have a finite lower radius are contrary to the observed (m, z) relation (§ 6.1) since they predict an accelerating expansion $(\Lambda > 0)$. The Λ term is unimportant, in any case, if a high-density is approached. In relativistic cosmology, therefore, by assuming a model suggested by the apparent large-scale homogeneity and isotropy of the Universe, we have to reckon with a cosmic singularity, a finite time $(\leqslant H^{-1})$ ago.

However, if the oldest star systems of the Galaxy have ages exceeding H^{-1} as some astronomers have calculated (§ 6.6), then an obvious contradiction exists since stars could not survive an ultra-high-density phase. Although at this stage the stellar age calculations could easily be wrong, the difficulty has led many theorists to look for ways to avoid a singularity in relativistic cosmology.

One possibility that has been pursued by several authors is to assume that the Universe is spatially homogeneous but not isotropic. Such a model may admit shear and rotation as well as expansion. The idea of a rotating universe occurred to Gamow (1946 b), and Gödel (1949) discovered a static universe with local spin of matter as a solution of the Einstein equations. Gödel (1950) has also discussed solutions that are spatially homogeneous and combine rotation with expansion of incoherent matter $(p = 0)$. By taking the x_4 lines to be the world lines of matter the space–time metric of a homogeneous but anisotropic model will be given by (3.4). In this case, however, we do not assume Weyl's postulate but measure x_4 or t, as the proper time along the world lines of matter drawn from some initial three-surface. The existence of a local spin (relative to local inertial frames) then implies that $g_{\mu 4} \neq 0$. Indeed, as Gödel showed, it means that a system of three-surfaces orthogonal to the world lines of matter does not exist. In such models there is therefore no measure of local time that is in a sense absolute as in the isotropic models. In particular, there is no non-trivial transformation that will transform the three-spaces $x_4 =$ constant into themselves. A peculiar consequence is that an observer may apparently reach events in his own past by following a suitable time-like path. However, this result may depend on the assumed connectivity of space.

It has been argued that since general relativity admits models with rotation the theory does not automatically incorporate Mach's principle, since it suggests that inertial frames are not determined only by the influence of world matter. However, this point is perhaps still obscure since all that appears to be established in the rotating models is that matter rotates locally relative to the local compass of inertia. On the other hand, a satisfactory theory of Mach's principle would account for inertia in terms of the Universe as a whole. The question that still remains unanswered is: what is the significance of inertial frames in general relativity, in particular, in rotating universes?

Following Gödel, spatially homogeneous solutions of the field equations have been given by Heckmann and Schücking (1958, 1962), and Shepley (1964). These models have shear, or rotation with shear. By motion with shear we mean an asymmetry of the field of flow relative to matter at a given point. According to the field equations, expanding rotating matter will also have shear. The above authors were able to show that some of the models have a singularity but the general behaviour of the class was not clear. No models were found without singularity.

Heckmann and Schücking (1955, 1956) also studied the problem of the singularity on the basis of Newtonian cosmology. Rotation of matter was superimposed on expansion in homogeneous Newtonian models of uniform density and pressure (functions of t only). A 'cosmological constant' Λ, giving repulsion or attraction according to sign, was also admitted. This cosmology permits rotation without shear, and Heckmann and Schücking claimed that models exist which with sufficient initial spin can expand from a finite density to infinite diffusion. These cases exist even if $\Lambda = 0$, and in this event there is a decelerating expansion and ultimate isotropy. This claim has been resisted by Narlikar (1963) who finds that if $\Lambda = 0$ the homogeneous spinning models, with or without shear, all have an infinite density in the past. Further investigation of these contradictory results is clearly called for.

An important and illuminating analysis of the motion of world matter in general relativity has been provided by Raychaudhuri (1955). He assumes zero world pressure (dust) and obtains a differential equation for the time variation of the scalar of expansion relative to local matter in an arbitrary world model. Expressed in local inertial coordinates the result is the following equation for $R(t)$, the scale function giving the mean local separation of matter at time t:

$$\ddot{R} - \tfrac{1}{3}R(\Lambda - q_{\mu\nu}\,q_{\mu\nu}) - \tfrac{2}{3}R\omega_\mu\,\omega_\mu + \frac{GM}{R^2} = 0 \tag{9.4}$$

where $3\dot{R}/R = \Theta$ is the local scalar of expansion, $q_{\mu\nu}$ $(\mu, \nu = 1, 2, 3)$ is the shear tensor in local three-space, ω_μ is the local spin vector, and $M = \frac{4}{3}\pi\rho R^3$ is a constant of the motion, ρ being the density. The equation shows how the expansion of the matter is affected by shear and rotation, gravity and cosmological term. In the absence of shear and rotation the effect of gravity (the last term) is to increase $|\Theta|$ indefinitely as R decreases, while that of Λ is relatively small if $R \to 0$. This indicates the singularity already discussed for the isotropic Friedmann models. If, however, the shear and spin terms are also present a new situation arises. These terms may get large when $R \to 0$, and whereas the shear terms evidently support gravitational collapse that of rotation has the opposite effect. Judged from Raychaudhuri's analysis the general existence of a singularity is therefore open to question.

We remark, incidentally, that the equation (9.4) has an exact counterpart in Newtonian cosmology. This fact, coupled with the existence of the Heckmann–Schücking Newtonian models, extends the validity of Newtonian cosmology (for dust) from the isotropic models through the homogeneous anisotropic models to an arbitrary inhomogeneous model (cf. § 4.5).

In a series of papers Lifshitz (1946), Lifshitz *et al.* (1961) and Lifshitz and Khalatnikov (1961 a, b, 1963, 1964) of the Russian school have attempted to show that the singularity in relativistic cosmology is present only if restricting assumptions, for example, isotropy, homogeneity, etc., are assumed. It will not feature, they suggest, in a completely general solution not subject to special initial conditions of symmetry. The method of Lifshitz and Khalatnikov (1964) has been to assume a singularity to occur universally in space at $t = 0$. Two wide classes of anisotropic solutions are then considered in its neighbourhood, that is, the principal terms only are retained. These classes contain as special cases all the known solutions with a singularity, and indeed the authors believe that the classes exhaust the models with singularity. As is always possible, a 'synchronous' frame of reference is adopted, in which $g_{44} = 1$, $g_{4\alpha} = 0$. Of course, such a frame will not in general have the world lines of matter as $x^\alpha =$ constant, i.e. it will not be co-moving.

In every case examined the authors find that the number of physically arbitrary functions of the spatial coordinates in the solution is less than eight, which they show is the number expected in a general solution of the field equations (excluding the three arbitrary functions associated with possible transformations of the spatial coordinates in the chosen synchronous reference frame). On the other hand, it is demonstrated by Lifshitz and Khalatnikov that any synchronous frame will always have a coordinate singularity in a finite time, owing to the inevitable intersections of the geodesic coordinate mesh ($x^\alpha =$ constant). By a suitable choice of the coordinates these intersections can be made to occur at a single instant of time t. This coordinate singularity coincides with a physical singularity when the coordinates are co-moving with matter, as in the Friedmann models. But such a coincidence would not obtain in homogeneous anisotropic models with local spin, or in general inhomogeneous models. In fact the authors show that, in a contracting universe, any perturbations from initial homogeneity and isotropy would grow with time. This confirms their belief that a singularity would not occur *in future time* in a general solution, unrestricted by built-in relations to secure special symmetries and the stability of these symmetries.

Invaluable and impressive as this analysis is, its logical approach has evidently a negative character. It does not prove conclusively that the general solution has no physical singularity. Since a family of solutions with singularity was found with as many as seven arbitrary functions of the spatial coordinates, one is left with some doubt as to the validity of the authors' conclusions.

Very recently a fresh approach to the singularity has been made by relativists in England. Hawking and Ellis (1965) have concentrated on the problem of finding how weak initial conditions can be made for the field equations to lead to a singularity. Their methods have been partly inspired by Penrose (1965) in his treatment of a collapsing star. Penrose sets up the following conditions for gravitational collapse of a star: minimum and plausible restrictions on the behaviour of matter at all times, viz. $p < \frac{1}{3}\rho$, $\rho > 0$; there exists a future and past direction of time everywhere; there exists a non-compact Cauchy surface intersecting every time-like and null line once and once only; and finally, a two-dimensional 'closed trapped surface' exists —that is, one that encloses a region from which light is prevented gravitationally from escaping. If these conditions are satisfied then the region must implode to a singularity, regardless of any inhomogeneity or rotation that may develop.

Hawking and Ellis apply similar criteria to a collapsing homogeneous universe. In this case the existence of a closed trapped surface does not have to be assumed initially. The homogeneity is specified by the existence, in the four-dimensional manifold of space–time, of a group of motions G_r $(r > 3)$ which is transitive on at least one space-like hypersurface (tentatively the three-dimensional cosmic space observed today). Space–time is assumed not to be stationary. With these data the authors apply Raychaudhuri's equation (for $\Lambda = 0$) to an irrotational geodesic congruence orthogonal to the surfaces of homogeneity, and argue that these surfaces will inevitably end up as a collapsing two-surface. The world lines of matter will converge on the two-surface in a finite time. There will therefore be a singularity according to Hawking and Ellis if only the Universe is spatially homogeneous, so that anisotropy with shear and rotation could exist. If this theorem is correct the suggestions of Heckmann and Schücking denying the necessity of a singularity in rotating models would be controverted. However, the part played in the theorem by Raychaudhuri's equation for the world lines of matter, applied to a geodesic congruence which is in general not the world lines of matter, is not entirely clear. Further clarification of this important result would be welcome.

Hawking (1965) has also advanced a proof of the existence of the singularity for an *open* model of the Universe, which is homogeneous overall but has local inhomogeneities. By postulating the first three of the Penrose conditions, Hawking shows that there must exist a closed trapped spherical surface for sufficient radius, and hence a physical singularity would occur in a contracting universe or occurred in the past in an expanding universe. There is still an assumption of symmetry in this theorem, of course, since the Universe must remain homogeneous overall during its contraction or expansion.

The question of the singularity in relativistic cosmology is evidently a crucial one for further understanding of the Universe. If singular conditions did not occur in the actual Universe then the question arises: how can the Universe expand from or contract to a finite radius? If the Universe oscillates in this way, why is there still so much hydrogen not converted to heavier elements—unless the density and temperature get so high ($T > 10^{10}$ °K) as to disrupt the heavier elements at maximum contraction? But in this case the term 'singularity' would be fairly appropriate, except possibly if high-density conditions were confined to discrete massive objects as recently suggested by Hoyle and Tayler (1964). If the elements are not disrupted then there must be creation of matter, and this might be connected with the reversal of expansion without singularity. The zero-point stress introduced into general relativity by McCrea (1951), the C-field theory of Hoyle and Narlikar, and further discoveries of nuclear physics and fundamental particle theory may all throw light on this contingency.

If a time singularity (i.e. universal) did occur in the finite past, then, as we have seen, this is consistent with many particular solutions of the Einstein field equations. As Lifshitz and Khalatnikov (1964) have indeed remarked, it is precisely such a solution that may have the necessary exceptional character that should describe the real world. In fact the initial conditions in the real Universe may not be 'arbitrary', and a 'general' world solution of the field equations may be forever irrelevant.

10. Possible directions for future development of observational and theoretical programmes

In the course of this review several gaps in our present knowledge have been indicated. In this final section it may be useful to summarize these and other possible directions of future research in cosmology.

10.1. *Observational programme*

(i) One of the most valuable sources of cosmological information will continue to be the measurements of red shift and apparent magnitude of distant galaxies. At the present time there is a need to reduce the scatter in the (m, z) diagram and to obtain higher values of z. If the brilliant quasi-stellar sources and quasi-stellar galaxies are used for this purpose, it will be necessary to determine with considerable accuracy the intrinsic brightness of these objects and if possible their evolutionary characteristics. It is probable that this will be acquired only through the amassing of sufficient statistical information and new identifications. However, it is possible that strong clues may come from a success in explaining the objects theoretically.

For the higher values of z likely to be reached it will be necessary to use an exact (m, z) relation.

(ii) The counts of radio sources have already proved to be of great cosmological significance. It will be necessary, of course, to make completely certain that the majority of the weaker sources and the quasi-stellar sources now observed are in fact at cosmological distances. Confirmation of the counts from wider surveys at low flux densities and an extension to even lower values of S would have immediate relevance to cosmological theories. Further investigations into the degree of source clustering and a search for anisotropy at low values of S would provide vital information.

(iii) Further observations designed to discover the nature of the intergalactic medium appear to have high priority. The existence and density of intergalactic hydrogen and of intergalactic cosmic rays, x rays and γ rays are questions of great significance. We have seen that the interpretation of several observations already made hinge on whether intergalactic hydrogen is ionized. It is also particularly important to establish the intensity of isotropic background radiation over the widest possible range of frequencies. Such information will, of course, help to sift out the more likely theories of emission at these frequencies. Detection of anisotropy in the background due to localized sources of x rays and γ rays will be of immense value. The level of the neutrino background must also be found.

(iv) More accurate determinations of the helium–hydrogen and other cosmic abundance ratios may be of crucial importance for theories of the origin of the elements, and therefore perhaps for cosmology.

(v) The nature of the central regions of galaxies is still largely obscure. Further observations directed to revealing the size, space density and motions of stars and gas clouds in these regions must react on a wide range of astrophysical and cosmological issues. Investigations of exploding galaxies and a search for evidence of past explosions in normal galaxies will be particularly important.

(vi) Systematic measurements of the angular diameter of radio sources, including quasars, down to lower limits of flux density are certain to be fruitful. Attempts at fitting the number counts to exploding universes have shown that the dispersion in the intrinsic power of the sources must have been considerably less in the past. Consequently, any information that can help to verify this indication will be most valuable. Furthermore, if radio sources get more homogeneous in character as the distance increases it may be possible to show a minimum of angular diameter of the sources as predicted for some evolutionary cosmologies.

10.2. *Theoretical programme*

(i) An obvious priority is to discover the true nature of the quasi-stellar sources and quasi-stellar galaxies and therefore to devise relevant theories to describe them which not only explain the data but can be checked *by their predictions of further aspects* of these objects. Is a collapsing quasar a model for a collapsing universe?

(ii) Related to the above question is the whole subject of the cosmogony of galaxies and stars and their evolution. This subject is very difficult but evidently of vital importance for the whole of cosmology. The problem of the formation of galaxies from a primordial medium has received little or no attention in this review since few positive results are available. Several authors have come to the tentative conclusion that in relativistic cosmology galaxies cannot form in the time available if they arise by small perturbations of the primitive gas (Lifshitz 1946, Gamow 1952, Bonnor 1956, 1957). However, recent work has presented a fresh approach (Layzer 1963, Irvine 1965, Peebles 1965), which although preliminary offers some interesting possibilities.

(iii) Of absorbing interest is the question of the 'singularity' in relativistic cosmology. Can non-singular solutions be found that are consistent with the available observational data and lie within the domain of known physics and classical general relativity? Can a singularity be prevented by the introduction of a new fundamental field? If so, what experiments can be suggested to detect this field in physics? Would the logical consequence of such a field be an oscillating universe?

If an ultra-high-density state of the Universe turns out to be inevitable it will be important to find an analysis that will reach as far as possible towards it, as attempted in the $\alpha\beta\gamma$ theory. In such an attempt to retrace the history of the Universe it will be of paramount interest to find at what points the theory has a bearing on present observations.

(iv) The whole question of cosmic thermodynamics, arrow of time, time reversal and its possible connection with fundamental physics deserves continued study. It is interesting to enquire how thermodynamics would behave in a finitely oscillating universe. Would nucleosynthesis go the opposite way in the contracting phase? Would electromagnetic signals be retarded or advanced?

(v) The thorny question of Mach's principle and its role in cosmological theories needs further examination. What is the significance of inertial frames in general relativity, if not in accordance with Mach's principle?

10.3. *Concluding remarks (summary)*

In the last decade the momentum of observational and theoretical research in cosmology has increased rapidly and the subject is being tackled with fruitful results on a very wide front. Observational results obtained by radio and optical astronomers have shown unequivocably that cosmological theories related to the observations can in turn be subjected to further decisive tests in the classical scientific manner.

With regard to the two theories most seriously considered in recent years, the steady-state theory and relativistic cosmology, a heavy weight of evidence at its face value now stands against the former theory in its original simple form. The red-shift and apparent-magnitude measurements, the counts of radio sources, the background radiation at certain wavelengths and the observed He/H abundance ratio all support a Universe that was denser in the past and whose expansion is now decelerating. In addition, recent tests suggest that intergalactic space is extremely vacuous, the density there being far below that of luminous matter.

By the same token relativistic cosmology finds very substantial observational support, since its equations possess many solutions which predict an ultra-high-density state and which are representative of the world as observed today. On the other hand, there is a certain difficulty in so far as the present estimates of the Hubble age ($\simeq 10^{10}$ year) indicate that the period since the high-density state is less than the evolutionary ages calculated for the oldest stars $(1{\cdot}2\text{–}2) \times 10^{10}$ year. The simple steady-state theory is also in disagreement with the calculations of stellar ages, and also with the apparent age distribution of galaxies. However, the steady-state cosmology yields consistent retarded solutions for electromagnetic waves in the Wheeler–Feynman theory, which does not seem to be possible in the exploding cosmologies. Very recent versions of the steady-state theory can also provide solutions consistent with a high-density state in the past. This possibility depends on the introduction into physics of a negative energy creation field and a non-uniform creation of matter.

So far no non-singular model of the Universe has been obtained within the framework of Einstein's original equations alone. A universe oscillating between finite radii is mathematically possible, but physically unreal or directly contrary to observation. Strenuous arguments have been asserted for and against the necessary existence of a singularity in the relativity equations. In the latter case the singularity may have to be regarded as a unique aspect of the real world. However, the difficulty may possibly be resolved by a modification of general relativity, or as indicated above by the introduction of new fields in physics at high density and temperature.

If the cosmological red shifts of quasars should be confirmed then these objects will be established as amazing new features of the Universe. In this case there will be strong observational clues ready to hand which will surely have the greatest relevance to cosmology and high-energy physics.

References

Abell, G. O., 1958, *Astrophys. J.* (*Suppl.* 3), 211.
Adgie, R. L., *et al.*, 1965, *Nature, Lond.*, **208**, 275.
Alexander, J. K., and Stone, R. G., 1965, *Astrophys. J.*, **142**, 1327.
Allen, L. R., *et al.*, 1962, *Mon. Not. R. Astr. Soc.*, **124**, 477.
Alpher, R. A., Bethe, H., and Gamow, G., 1948, *Phys. Rev.*, **73**, 803.
Alpher, R. A., Follin, J. W., and Herman, R. C., 1953, *Phys. Rev.*, **92**, 1347.

ALPHER, R. A., and HERMAN, R. C., 1950, *Rev. Mod. Phys.*, **22**, 153.
AMBARTSUMIAN, V. A., 1958, *11th Solvay Conference* (*La Structure et l'Evolution de l'Univers*), *Brussels*, Ed. R. Stoops (Brussels: Inst. Internat. de Phys. Solvay), p. 241.
BAADE, W., and SWOPE, H. H., 1955, *Astr. J.*, **60**, 151.
BAUM, W. A., 1957, *Astr. J.*, **62**, 6.
—— 1961a, *Observatory*, **81**, 114.
—— 1961 b, *Symp. Int. Astr. Un.*, *University of California*, Ed. G. C. McVittie (New York: Macmillan), p. 390.
BENNETT, A. S., 1962, *Mem. R. Astr. Soc.*, **68**, 163.
VAN DEN BERGH, S., 1960 a, *Astrophys. J.*, **131**, 558.
—— 1960 b, *Mon. Not. R. Astr. Soc.*, **121**, 387.
—— 1961 a, *Z. Astrophys.*, **53**, 219.
—— 1961 b, *Publ. Astr. Soc. Pacif.*, **73**, 46.
BETHE, H. A., 1939, *Phys. Rev.*, **55**, 103, 434.
BOLTON, J. G., CLARKE, M. E., SANDAGE, A. R., and VERON, P., 1965, *Astrophys. J.*, **142**, 1289.
BONDI, H., 1959, *Rep. Progr. Phys.*, **22**, 97 (London: Physical Society).
BONDI, H., and GOLD, T., 1948, *Mon. Not. R. Astr. Soc.*, **108**, 252.
BONDI, H., and LYTTLETON, R. A., 1959, *Proc. Roy. Soc.* A, **252**, 313.
BONNOR, W. B., 1956, *Z. Astrophys.*, **39**, 143.
—— 1957, *Mon. Not. R. Astr. Soc.*, **117**, 104.
—— 1964, *Mon. Not. R. Astr. Soc.*, **128**, 33.
BROWN, H., 1949, *Rev. Mod. Phys.*, **21**, 625.
BURBIDGE, E. M., 1965 a, *Astrophys. J.*, **142**, 1291.
—— 1965 b, *Astrophys. J.*, **142**, 1674.
—— 1966, *Astrophys. J.*, **143**, 612.
BURBIDGE, G. R., BURBIDGE, E. M., FOWLER, W. A., and HOYLE, F., 1957, *Rev. Mod. Phys.*, **29**, 547.
BURBIDGE, G. R., BURBIDGE, E. M., and SANDAGE, A. R., 1963, *Rev. Mod. Phys.*, **35**, 947.
BURBIDGE, G. R., and HOYLE, F., 1956, *Nuovo Cim.*, [10], **4**, 558.
—— 1966, *Astrophys. J.*, in the press.
CAMERON, A. G. W., 1955, *Astrophys. J.*, **121**, 144.
CLARKE, R. W., SCOTT, P. F., and SMITH, F. G., 1963, *Mon. Not. R. Astr. Soc.*, **125**, 195.
CONWAY, R. G., KELLERMANN, K. I., and LONG, R. J., 1963, *Mon. Not. R. Astr. Soc.*, **125**, 261.
COOK, C. W., FOWLER, W. A., LAURITSEN, C. C., and LAURITSEN, T., 1957, *Phys. Rev.*, **107**, 508.
DAVIDSON, W., 1959, *Mon. Not. R. Astr. Soc.*, **119**, 54.
—— 1962 a, *Mon. Not. R. Astr. Soc.*, **123**, 425.
—— 1962 b, *Mon. Not. R. Astr. Soc.*, **124**, 79.
—— 1962 c, *Nature, Lond.*, **196**, 155.
—— 1965, *Nature, Lond.*, **206**, 249.
DAVIDSON, W., and DAVIES, M., 1964 a, *Mon. Not. R. Astr. Soc.*, **127**, 241.
—— 1964 b, *Mon. Not. R. Astr. Soc.*, **128**, 363.
DAVIES, R. D., 1964, *Mon. Not. R. Astr. Soc.*, **128**, 133.
DAVIES, R. D., and JENNISON, R. C., 1964, *Mon. Not. R. Astr. Soc.*, **128**, 123.
DENT, W. A., 1965, *Science*, **148**, 1458.
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.* A, **167**, 148.
DOROSHKEVICH, A. G., and NOVIKOV, I. D., 1964, *Soviet Physics–Doklady*, **9**, 111.
EDGE, D. O., *et al.*, 1959, *Mem. R. Astr. Soc.*, **68**, 37.
FELTEN, J. E., and MORRISON, P., 1963, *Phys. Rev. Letters*, **10**, 453.
FIELD, G. B., 1962, *Astrophys. J.*, **135**, 684.
—— 1964, *Astrophys. J.*, **140**, 1434.
FIELD, G. B., and HENRY, R. C., 1964, *Astrophys. J.*, **140**, 1002.
FOWLER, W. A., 1964, *Rev. Mod. Phys.*, **36**, 545.
FOWLER, W. A., and HOYLE, F., 1960, *Ann. Phys., N.Y.*, **10**, 280.
—— 1963 a, *Mon. Not. R. Astr. Soc.*, **125**, 169.

Fowler, W. A., and Hoyle, F., 1963 b, *Nature, Lond.*, **197**, 533.
—— 1965, *Quasi-Stellar Sources and Gravitational Collapse* (Chicago: University of Chicago Press), p. 97.
Friedmann, A., 1924, *Z. Phys.*, **21**, 326.
Gamow, G., 1946 a, *Phys. Rev.*, **70**, 572.
—— 1946 b, *Nature, Lond.*, **158**, 549.
—— 1948, *Nature, Lond.*, **162**, 680.
—— 1952, *Phys. Rev.*, **86**, 251.
Giacconi, R., Gursky, H., Paolini, F. R., and Rossi, B., 1962, *Phys. Rev. Letters*, **9**, 439.
Giacconi, R., Gursky, H., and Waters, J. R., 1965, *Nature, Lond.*, **207**, 572.
Ginzburg, V. L., and Syrovatskii, S. I., 1965, *Soviet Physics–Uspekhi*, **7**, 696.
Gödel, K., 1949, *Rev. Mod. Phys.*, **21**, 447.
—— 1950, *Proc. Int. Congress of Mathematicians, Cambridge, Massachusetts*, 1952 (Providence: American Mathematical Society), p. 175.
Gold, T., Axford, W. I., and Ray, E. C., 1963, *1st Texas Symposium on Relativistic Astrophysics, Dallas* (Chicago: University of Chicago Press), p. 93.
Gold, T., and Hoyle, F., 1958, *I.A.U. Symposium on Radio Astronomy, Paris* (Stanford: University of Stanford Press), p. 583.
Goldstein, S. J., 1963, *Astrophys. J.*, **138**, 978.
Gould, R. J., and Burbidge, G. R., 1963, *Astrophys. J.*, **138**, 969.
Gunn, J. E., and Peterson, B. A., 1965, *Astrophys. J.*, **142**, 1633.
Hanbury Brown, R., 1962, *Mon. Not. R. Astr. Soc.*, **124**, 35.
Hanbury Brown, R., and Hazard, C., 1961 a, *Mon. Not. R. Astr. Soc.*, **122**, 479.
—— 1961 b, *Mon. Not. R. Astr. Soc.*, **123**, 279.
Haro, G., and Luyten, W., 1962, *Boln Observs Tonantzintla Tacubaya*, **3**, 37.
Harrison, E. R., 1964, *Nature, Lond.*, **204**, 271.
Hawking, S. W., 1965, *Phys. Rev. Letters*, **15**, 689.
Hawking, S. W., and Ellis, G. F. R., 1965, *Phys. Letters*, **17**, 246.
Hayashi, C., *Progr. Theor. Phys., Japan*, 1950, **5**, 224.
Heckmann, O., and Schücking, E., 1955, *Z. Astrophys.*, **38**, 95.
—— 1956, *Z. Astrophys.*, **40**, 81.
—— 1958, *11th Solvay Conference (La Structure et l'Evolution de l'Univers), Brussels*, Ed. R. Stoops (Brussels: Inst. Internat. de Phys. Solvay), p. 149.
—— 1962, *Gravitation*, Ed. L. Witten (New York: John Wiley), p. 438.
Hillas, A. M., and Cranshaw, T. E., 1959, *Nature, Lond.*, **184**, 892.
Hogarth, J. E., 1962, *Proc. Roy. Soc.* A, **267**, 365.
Hoyle, F., 1948, *Mon. Not. R. Astr. Soc.*, **108**, 372.
—— 1954, *Astrophys. J.* (*Suppl.* 1), 121.
—— 1958, *11th Solvay Conference (La Structure et l'Evolution de l'Univers), Brussels*, Ed. R. Stoops (Brussels: Inst. Internat. de Phys. Solvay), p. 53.
—— 1961, *International School of Physics, 'Enrico Fermi' Course XX, Varenna* (New York, London: Academic Press), p. 141.
Hoyle, F., Fowler, W. A., Burbidge, G. R., and Burbidge, E. M., 1964, *Astrophys. J.*, **139**, 909.
Hoyle, F., and Narlikar, J. V., 1961, *Mon. Not. R. Astr. Soc.*, **123**, 133.
—— 1962 a, *Observatory*, **82**, 13.
—— 1962 b, *Mon. Not. R. Astr. Soc.*, **125**, 13.
—— 1963 a, *Proc. Roy. Soc.* A, **273**, 1.
—— 1963 b, *Proc. Roy. Soc.* A, **277**, 1.
—— 1964 a, *Proc. Roy. Soc.* A, **278**, 465.
—— 1964 b, *Proc. Roy. Soc.* A, **282**, 178.
—— 1966 a, *Proc. Roy. Soc.* A, **290**, 143.
—— 1966 b, *Proc. Roy. Soc.* A, **290**, 162.
—— 1966 c, *Proc. Roy. Soc.* A, **290**, 177.
Hoyle, F., and Tayler, R. J., 1964, *Nature, Lond.*, **203**, 1108.
Humason, M. L., Mayall, N. U., and Sandage, A. R., 1956, *Astr. J.*, **61**, 97.
Irvine, W. M., 1965, *Ann. Phys., N.Y.*, **32**, 322.

Kaufman, M., 1965, *Nature, Lond.*, **207**, 736.
Kenderdine, S., 1963, *Mon. Not. R. Astr. Soc.*, **126**, 41.
Kiang, T., 1961, *Mon. Not. R. Astr. Soc.*, **122**, 263.
Kinman, T. D., 1965, *Astrophys. J.*, **142**, 1241.
Kraushaar, W. L., and Clark, G. W., 1962, *Phys. Rev. Letters*, **8**, 106.
Kraushaar, W. L., *et al.*, 1965, *Astrophys. J.*, **141**, 845.
Layzer, D., 1963, *Astrophys. J.*, **138**, 174.
Leslie, P. R. R., 1961, *Mon. Not. R. Astr. Soc.*, **122**, 371.
Lifshitz, E. M., 1946, *J. Phys., Moscow*, **10**, 116.
Lifshitz, E. M., and Khalatnikov, I. M., 1961 a, *Soviet Physics-JETP*, **12**, 108.
—— 1961 b, *Soviet Physics-JETP*, **12**, 558.
—— 1963, *Advanc. Phys.*, **12**, 185.
—— 1964, *Soviet Physics–Uspekhi*, **6**, 495.
Lifshitz, E. M., Sudakov, V. V., and Khalatnikov, I. M., 1961, *Soviet Physics–JETP*, **13**, 1298.
Longair, M. S., and Scott, P. F., 1965, *Mon. Not. R. Astr. Soc.*, **130**, 379.
Lynds, C. R., Stockton, A. N., and Livingston, W., 1965, *Astrophys. J.*, **142**, 1667.
McCrea, W. H., 1951, *Proc. Roy. Soc.* A, **206**, 562.
—— 1953, *Rep. Progr. Phys.*, **16**, 321 (London: Physical Society).
—— 1964, *Mon. Not. R. Astr. Soc.*, **128**, 335.
Mach, E., 1883, *The Science of Mechanics* (La Salle, Illinois: Open Court Publishing Company), p. 229.
McVittie, G. C., 1962, *Phys. Rev.*, **128**, 2871.
—— 1965, *General Relativity and Cosmology* (London: Chapman and Hall), p. 189.
Mills, B. Y., Slee, O. B., and Hill, E. R., 1958, *Aust. J. Phys.*, **11**, 360.
—— 1960, *Aust. J. Phys.*, **13**, 676.
Minkowski, R., 1960, *Astrophys. J.*, **132**, 908.
—— 1961, *Symp. Int. Astr. Un., University of California*, Ed. G. C. McVittie (New York: Macmillan), p. 379.
Moffett, A. T., and Maltby, P., 1962, *Astrophys. J.* (*Suppl.* 7), 141.
Narlikar, J. V., 1962, *Proc. Roy. Soc.* A, **270**, 553.
—— 1963, *Mon. Not. R. Astr. Soc.*, **126**, 203.
Néeman, Y., 1965, *Astrophys. J.*, **142**, 1241.
Oke, J. B., 1965, *Astrophys. J.*, **142**, 810.
Oort, J. H., 1958, 11*th Solvay Conference* (*La Structure et l'Evolution de l'Univers*), *Brussels*, Ed. R. Stoops (Brussels: Inst. Internat. de Phys. Solvay), p. 163.
Page, T., 1959, *Astr. J.*, **64**, 53.
Pauliny-Toth, I. I. K., and Shakeshaft, J. R., 1962, *Mon. Not. R. Astr. Soc.*, **124**, 61.
Peebles, P. J. E., 1965, *Astrophys. J.*, **142**, 1317.
Penrose, R., 1965, *Phys. Rev. Letters*, **14**, 57.
Penzias, A. A., and Wilson, R. W., 1965, *Astrophys. J.*, **142**, 419.
Raychaudhuri, A. K., 1955, *Phys. Rev.*, **98**, 1123.
Robertson, H. P., 1933, *Rev. Mod. Phys.*, **5**, 62.
—— 1935, *Astrophys. J.*, **82**, 284.
—— 1936, *Astrophys. J.*, **83**, 187, 257.
Rowson, B., 1963, *Mon. Not. R. Astr. Soc.*, **125**, 177.
Ryle, M., and Clarke, R. W., 1961, *Mon. Not. R. Astr. Soc.*, **122**, 349.
Ryle, M., and Neville, A. C., 1962, *Mon. Not. R. Astr. Soc.*, **125**, 39.
Ryle, M., and Sandage, A. R., 1964, *Astrophys. J.*, **139**, 419.
Salpeter, E. E., 1957, *Phys. Rev.*, **107**, 516.
Sandage, A. R., 1958, *Astrophys. J.*, **127**, 513.
—— 1961 a, *Astrophys. J.*, **133**, 355.
—— 1961 b, *Astrophys. J.*, **134**, 916.
—— 1964, *Sci. Amer.*, **211**, 38.
—— 1965, *Astrophys. J.*, **141**, 1560.
Sandage, A. R., Veron, P., and Wyndham, J. B., 1965, *Astrophys. J.*, **142**, 1307.
Scheuer, P. A. G., 1965, *Nature, Lond.*, **207**, 963.

SCHMIDT, M., 1963, *Nature, Lond.*, **197**, 1040.
SCHMIDT, M., 1965 a, *Astrophys. J.*, **141**, 1.
—— 1965 b, *Astrophys. J.*, **141**, 1295.
SCHMIDT, M., and GREENSTEIN, J. L., 1964, *Astrophys. J.*, **140**, 1.
SCIAMA, D. W., 1963, *Mon. Not. R. Astr. Soc.*, **126**, 195.
—— 1964, *Mon. Not. R. Astr. Soc.*, **128**, 49.
—— 1965, *Nature, Lond.*, **204**, 767.
SCOTT, P. F., 1963, *Mon. Not. R. Astr. Soc.*, **127**, 37.
SCOTT, P. F., and RYLE, M., 1961, *Mon. Not. R. Astr. Soc.*, **122**, 389.
SHAKESHAFT, J. R., *et al.*, 1955, *Mem. R. Astr. Soc.*, **67**, 106.
SHEPLEY, L., 1964, *Proc. Nat. Acad. Sci., Wash.*, **52**, 1403.
DE SITTER, W., 1917, *Mon. Not. R. Astr. Soc.*, **78**, 3.
SMITH, F. G., 1964, *Symp. Int. Astr. Un.*, No. 23, *Liège*, p. 305 (*Ann. Astrophys.*, 1964, No. 6; 1965, Nos. 1, 2, 3, 4).
SMITH, H., 1963, *1st Texas Symposium on Relativistic Astrophysics, Dallas*, (Chicago: University of Chicago Press), p. 221.
SUESS, H. E., and UREY, H. C., 1956, *Rev. Mod. Phys.*, **28**, 53.
TAYLER, R. J., and CLIFFORD, F., 1965, *Mem. R. Astr. Soc.*, **69**, 21.
TOLMAN, R. C., 1934, *Relativity, Thermodynamics and Cosmology* (Oxford: Clarendon Press), p. 429.
TURTLE, A. J., and BALDWIN, J. E., 1962, *Mon. Not. R. Astr. Soc.*, **124**, 459.
TURTLE, A. J., PUGH, J. F., KENDERDINE, S., and PAULINY-TOTH, I. I. K., 1962, *Mon. Not. R. Astr. Soc.*, **124**, 297.
ULAM, S. M., and WALDEN, W. E., 1964, *Nature, Lond.*, **201**, 1202.
DE VAUCOULEURS, G., 1949, *Ann. Astrophys.*, **12**, 162.
WALKER, A. G., 1936, *Proc. Lond. Math. Soc.*, [2], **42**, 90.
WEINBERG, S., 1962 a, *Nuovo Cim.*, **25**, 15.
—— 1962 b, *Phys. Rev.*, **128**, 1457.
WEYL, H., 1923, *Phys. Z.*, **24**, 230.
WHEELER, J. A., 1963, *1st Texas Symposium on Relativistic Astrophysics, Dallas* (Chicago: University of Chicago Press), p. 000.
WHEELER, J. A., and FEYNMAN, R. P., 1945, *Rev. Mod. Phys.*, **17**, 157.
—— 1949, *Rev. Mod. Phys.*, **21**, 425.
WHITROW, G. J., and YALLOP, B. D., 1964, *Mon. Not. R. Astr. Soc.*, **127**, 301.
—— 1965, *Mon. Not. R. Astr. Soc.*, **130**, 31.
WIELEBINSKI, R., and YATES, K. W., 1965, *Nature, Lond.*, **205**, 58.
WIELEN, R., 1964, *Z. Astrophys.*, **59**, 129.