PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Material Content of the Universe: Introductory Survey

R. J. Tayler

Phil. Trans. R. Soc. Lond. A 1986 **320**, 435-446
doi: 10.1098/rsta.1986.0127

Phil. Trans. R. Soc. Lond. A **320**, 435–446 (1986)
Printed in Great Britain

Material content of the Universe: introductory survey

By R. J. Tayler
Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, U.K.

Matter in the Universe can be detected either by the radiation it emits or by its gravitational influence. There is a strong suggestion that the Universe contains substantial hidden matter, mass without corresponding light. There are also arguments from elementary particle physics that the Universe should have closure density, which would also imply hidden mass. Observations of the chemical composition of the Universe interpreted in terms of the hot Big Bang cosmological theory suggest that this hidden matter cannot all be of baryonic form but must consist of weakly interacting elementary particles. A combination of observations and theoretical ideas about the origin of large-scale structure may demand that these particles are of a type which is not yet definitely known to exist.

1. Hidden matter

Although this symposium is entitled Material content of the Universe, it will rapidly become clear that it is mainly concerned with matter whose presence is only deduced indirectly, or hidden matter, rather than with the obvious luminous matter, stars and gas clouds. There are two key questions: does the Universe contain a significant fraction of its mass in the form of hidden matter and what is the nature of the hidden matter? These questions will be discussed in detail in later papers; I shall introduce the general ideas of the subject.

The presence of matter can be deduced either from the radiation that it emits (or absorbs) or from its gravitational influence on other objects. If the matter is observed from its radiation it is usually possible to obtain some clear idea of its nature and of its quantity and the conclusion is that this *visible matter* is ordinary baryonic matter similar to that of which the Earth is composed; when I refer to baryonic matter I include with it those leptons (electrons) that provide charge neutrality but which make a negligible contribution to the mass. In contrast, if the matter is detected only by its gravitational influence, it is *hidden matter* and the only thing that can be immediately deduced from the observations is that there is something present which interacts through the force of gravitation. Sometimes theoretical considerations will favour the choice of one particular form of hidden matter such as a neutron star or a black hole in a close binary stellar system.

Even in what I have written so far there is an important assumption, which should be made explicit. I am assuming that we understand the laws of physics and that these laws apply throughout space and time. This assumption has been used in identifying the luminous matter and in detecting the hidden matter. As I shall mention, there are possible ways of solving some of the problems relating to hidden mass by making modifications to the generally accepted laws of physics on length scales on which they have not been directly tested.

An early introduction of the concept of hidden mass was made by Oort (1932). He used the observed motions of stars perpendicular to the galactic plane in the solar neighbourhood to estimate the surface density of gravitating matter in the galactic disc. The figure that he obtained was significantly higher than the known mass of stars. At that time astronomy was essentially optical astronomy and very little indeed was known about the interstellar gas, most of which does not show up very well at optical wavelengths. The extension of the wavelength range available to astronomers, and in particular the prediction and discovery of the 0.21 m radiation from cold hydrogen, has meant that the interstellar gas in the solar neighbourhood can be added to the known matter. A significant discrepancy still exists between the current value of the Oort limit and the observed surface density; I shall mention this further in §2 and it will be discussed by Bahcall (this symposium).

Another reason for believing that there must be hidden mass in flattened galaxies, such as our own, arises from a study of their rotation curves. In such a galaxy, random motions of stars and gas clouds are small compared with the ordered motion of galactic rotation and to a good approximation

$$v_\phi^2/\varpi = -\partial\Phi/\partial\varpi \tag{1.1}$$

must apply in cylindrical polar coordinates ϖ, ϕ, z, where v_ϕ is the rotation speed and Φ the gravitational potential defined so that the force per unit mass is $+\nabla\Phi$. There are two ways in which (1.1) can be used to demonstrate the presence of hidden mass. The simpler one arises because in many galaxies v_ϕ is found to be essentially independent of ϖ at distances from the axis of rotation where there is little light from the galaxy (see, for example, Rubin *et al.* 1985). If the mass distribution were to follow the light distribution at this value of ϖ, v_ϕ should be approaching the Keplerian law $v_\phi \propto 1/\varpi^{\frac{1}{2}}$, which must apply at large distances from a gravitating mass. The fact that it does not have this form implies that the mass:light ratio must vary with ϖ. Thus, if all the mass at some radii is visible, there must be hidden mass at large radii. The second way in which hidden mass may be deduced is by calculating the right-hand side of (1.1) for the visible mass and by realizing that this will not account for the rotation velocity even at small radii. Because rotation curves are flat as far out as they can be measured there is uncertainty concerning the total extent and mass of galaxies.

The next place in which the presence of hidden matter is indicated is in groups and clusters of galaxies. If it is assumed that a cluster of galaxies is in equilibrium, so that it is neither expanding nor contracting as a whole, or, for example, that a pair of galaxies is in stable orbit about its centre of mass, there is a simple relation between the gravitational energy, V, and the kinetic energy, T, of the system relative to its centre of mass

$$2T + V = 0, \tag{1.2}$$

the *virial theorem*. Only one component of velocity of a galaxy can be observed but an estimate of the total kinetic energy can be obtained if it is assumed that the overall galaxy motions in a group are isotropic. When estimates of T and V are compared, it is usually found that $T > V$, with $T \gg V$ in clusters of galaxies (see, for example, Faber & Gallagher 1979). This result can be reconciled with (1.2) if there is a large amount of hidden mass because, even if this mass has the same velocity dispersion as the visible mass, V increases quadratically with the total mass whereas T increases linearly. This virial mass discrepancy can be very large indeed implying much more hidden mass than that in individual galaxies.

In §2 I shall comment further on these indications of hidden mass. They do not depend on any fine details of the cosmological model which is being used. I now turn to some further ideas related to the particular cosmological model, which is being adopted by most, if not all,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

subsequent authors in this symposium. Present observations of the large-scale structure of the Universe are certainly at least in good qualitative agreement with the standard model of the hot Big Bang cosmological theory, in which the Universe started very hot and very dense and was very close to a state of homogeneity and isotropy, and in which we can hope to understand the formation of galaxies and clusters out of the small deviations from uniformity which did exist at early epochs. We shall be concerned with several different ways in which the amount of matter in the Universe and the form of that matter interact with the theory.

As far as the previous observations are concerned, a natural first assumption is that all of the hidden matter is ordinary baryonic matter. It does not consist of luminous stars, and there are also forms of gas which can be ruled out. There is, however, no immediately obvious observational reason why the hidden mass should not be dead remnants of stars or more massive objects, subluminous low mass stars or even lower mass solid objects, although some such solutions may be made implausible by detailed studies of galactic or cluster evolution. We shall, however, see that cosmological arguments suggest strongly that, if all the suggested hidden mass really is present, much of it must be non-baryonic.

The simplest way in which the total mass density of the Universe enters cosmology is by the rate of slowing down of the expansion of the Universe and the relation of this to its estimated age. The expansion of the homogeneous, isotropic Universe can be described in terms of a function $R(t)$, which measures the separation, caused by the expansion, of two objects. To a first approximation the velocity of recession of a distant galaxy, v, is related to its distance, r, by

$$v = H_0 r, \tag{1.3}$$

where

$$H_0 = [(\mathrm{d}R/\mathrm{d}t)/R]_0 \tag{1.4}$$

is known as Hubble's constant, the suffix zero referring to the present epoch. At large distances there should be a deviation from (1.3) owing to deceleration of the expansion caused by the matter in the Universe. There is no good observational value for the deceleration parameter,

$$q_0 = -[R\mathrm{d}^2R/\mathrm{d}t^2/(\mathrm{d}R/\mathrm{d}t)^2]_0, \tag{1.5}$$

but, in the simplest case in which there is no cosmological constant in the general theory of relativity,

$$2q_0 = \rho_0/\rho_{\mathrm{crit}} \equiv \Omega_0, \tag{1.6}$$

where

$$\rho_{\mathrm{crit}} = 3H_0^2/8\pi G \tag{1.7}$$

is the density which just closes the Universe. The present age of the Universe is related to the value of Ω_0 being $1/H_0$ for $\Omega_0 = 0$ and $2/3H_0$ for $\Omega_0 = 1$ and less than that for $\Omega_0 > 1$. If the Big Bang theory is to be correct, the age of the Universe must be greater than that of galactic objects. I will discuss this further in §3.

There is a particular interest in the value $\Omega_0 = 1$. In general, Ω is not constant as the Universe evolves but, in the special case, $\Omega_0 = 1$, $\Omega \equiv \Omega_0$. Present observations are certainly inconsistent with Ω_0 lying outside the range

$$10^{-2} < \Omega_0 < 10, \tag{1.8}$$

where the limits in (1.8) are very generous. Now one of Einstein's equations can be written in the form

$$HR^2(\Omega - 1) = kc^2, \tag{1.9}$$

where the value of k is $+1$, 0 or -1. Solution of Einstein's equations shows that at earlier epochs R^2H^2 must have been very much larger than it is now so that $|\Omega - 1|$ was correspondingly smaller. Thus, whatever is the value of Ω today, it must have been equal to unity to many decimal places at earlier epochs. It may be aesthetically more satisfying to have $\Omega \equiv 1$. The second argument for $\Omega_0 = 1$ comes from the inflationary Universe scenario (see, for example, Guth 1982), which accounts for various problems with the Big Bang theory and which requires $\Omega_0 \approx 1$.

If it is accepted that Ω_0 may be *ca.* 1, there is then a very strong argument that the matter cannot all be baryonic. This arises from observations of the abundances of the light isotopes, particularly ^{2}H, ^{3}He and ^{4}He. These isotopes are believed to have been produced from protons and neutrons in the first few minutes of the expanding Universe. Calculations indicate that, although the fraction of ^{4}He is only slightly dependent on the density, the abundances of ^{2}H and ^{3}He are very strongly dependent and their deduced primeval abundances are not consistent with $\Omega_0 = 1$ if all of the matter is baryonic; high density does also make agreement of theory and observation for ^{4}He more difficult (see, for example, Schramm 1982). This problem may not arise if much of the matter is non-baryonic.

If a large amount of the mass in the Universe is not baryonic, what is it and how does it come to be present? It is believed that the Universe is electrically neutral and that the positive charges of atomic nuclei are balanced by negative charges of electrons. What we are therefore looking for are neutral stable particles with finite rest mass or possibly very long-lived unstable particles. For any particle there is a time that is early enough in the expansion for it to be in equilibrium with radiation; if it is stable or very long lived it will survive in some quantity today. The most obvious candidate is the neutrino. We know that there are at least three types of neutrino, the theory predicts that *microwave* neutrinos as well as microwave photons should be present in the Universe and a typical neutrino mass of a few tens of eV/c^2 would give $\Omega_0 \approx 1$. The neutrino is the only candidate for non-baryonic matter which rests at present on hard experimental fact. Particle theories, however, predict that various other particles might exist and have masses which could dominate the Universe. This will be discussed further below, where I shall also comment on the manner in which non-baryonic hidden matter might affect the formation of galaxies and clusters of galaxies.

There is one argument against believing that the Universe is dominated by stable neutral particles before their discovery. As has been stressed very frequently, it is perhaps rather surprising that there should exist such a particle, whose contribution to the density of the Universe probably differs from that of the baryons by only about one order of magnitude. It might be considered more plausible for either the baryons or the other particles to have clear domination. On the other hand, it has to be recognized that the existence of non-baryonic matter might give a natural explanation of the variation of the mass:light ratio with position in a galaxy, as this matter might settle differently in the gravitational field. In contrast, if all of the matter in a galaxy is baryonic, we need to understand why star formation has occurred quite differently at increasing distances from the galactic centre so as to produce the very different mass:light ratios. However, ultimately it must be admitted that the Universe is the way it is because the laws of physics, which include the spectrum of particles and their masses, and the initial conditions are what they are; if they combine to cause two types of matter to be dynamically important, that is what the Universe is like.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

2. Observational evidence

In §1, arguments have been given for the existence of hidden mass. How strong are these arguments? Inside disc galaxies the evidence is quite clear. If the mass distribution of a typical spiral galaxy followed the light distribution, the rotation curve could not have its observed form. There must be some invisible matter and the amount of gas which is observed by radio astronomers, for example, is far from sufficient to account for this. The precise quantity of unseen mass is a different matter both because, as mentioned earlier, the observed rotation curves do not show where the galaxy ends and because the rotation curve in the plane of a disc galaxy does not uniquely determine the mass distribution in the galaxy. A simple example demonstrates this. It is frequently stated that the flat rotation curves are evidence for massive halos and indeed a spherical mass distribution with $\rho \propto 1/r^2$ does produce a flat rotation curve. The same curve could, however, also be produced by an infinitely flattened spheroidal distribution whose mass is $2/\pi$ times that of the halo. Although this mass uncertainty is small it is not negligible.

The uncertainty in the masses of individual galaxies is not the most important thing about which we have to worry, because of the larger discrepancies between the visible mass and the estimated total mass of clusters of galaxies. It does, however, affect views about how much hidden mass may reside in galaxies and how much between galaxies. In clusters we have the largest mass discrepancies and the largest uncertainties and because of this it is worthwhile to list some of the problems that arise in obtaining an accurate estimate of the amount of hidden mass that is present. These include the following.

(*a*) Membership of a cluster. If one is to use a galaxy in a virial theorem calculation, it is necessary to have its redshift and this ensures that the most obvious sources of contamination by foreground and background objects are eliminated. Even so, contamination may be important (Turner *et al.* 1979; Fernley & Bhavsar 1984) and it may be difficult to distinguish a single cluster from two close and possibly merging clusters (Lucey 1983).

(*b*) Determination of the centre-of-mass radial velocity of a cluster. This is needed so as to determine the random speeds that must be used in the virial theorem. Here relative if not absolute masses of galaxies of different types are required, together with an accurate idea of the membership of the cluster, particularly inasmuch as the more massive galaxies are concerned.

(*c*) Isotropy of cluster velocity dispersion. This is usually assumed in order to convert line of sight velocities to three-dimensional velocities. It is most plausible when a cluster has a circular appearance.

(*d*) The assumption that the cluster actually is in virial equilibrium and the determination of the total mass and its comparison with the *known* mass in galaxies.

(*e*) The relation of the density distribution of the dark matter to that of the luminous galaxies. Bailey (1982), for example, showed that the mass of the Coma cluster was uncertain by a factor of at least four as a result of different possible arrangements of dark matter.

This subject is discussed in detail by Faber & Gallagher (1979).

The crucial step is the assumption of virial equilibrium. There is no direct way of telling whether this assumption is correct if we are discarding the *obvious* observational result that $2T+V \neq 0$. The first indirect approach is to ask whether or not the time it takes a galaxy to cross a cluster, t_c, is small compared with H_0^{-1}; if that is not the case it is unreasonable to suppose that the group has been able to reach virial equilibrium. Even if t_c is significantly less than H_0^{-1}, the group may not have had time to virialize. In the case of small groups of galaxies this problem was studied by Gott & Turner (1977), who decided that most of them had a small crossing time but that some would not yet have virialized. The position is less clear for large clusters and here it is necessary to appeal to numerical simulations of the origin of clusters in the Universe. These simulations are regarded as successful if they produce the observed pattern of galaxy clustering, although as Frenk explains (this symposium) the best results at present are obtained by assuming that galaxies only form at the highest peaks in the matter distribution; biased galaxy formation (Rees 1985; Schaeffer & Silk 1985). In principle the simulation which gives the observed clustering can be checked to see whether the clusters are in virial equilibrium and therefore whether it is consistent with the virial mass discrepancy initially deduced from the observations.

There is one further way in which the mass of our local cluster, the Virgo cluster, can be estimated. The Galaxy is not at rest in the local inertial frame defined by the microwave background radiation. A significant part of its motion can be interpreted as being caused by the gravitational attraction of the Virgo cluster. There is also an apparent infall of galaxies in the Virgo southern extension towards the central region of the cluster and this has recently been discussed by Tully & Shaya (1984) amongst others. The observed infall is apparently not consistent with $\Omega_0 = 1$ if the mass in the Universe has the same distribution as the visible galaxies (see, for example, Rees 1986) and this is another argument for biased galaxy formation.

Before I discuss the possible forms of the hidden mass one point should be made explicit. The virial theorem gives the total mass of a cluster in the form

$$M = k\langle v^2\rangle R/G, \tag{2.1}$$

where $\langle v^2\rangle$ is the average random velocity squared for the galaxies, R is an effective radius of the cluster and k is a constant of order unity. The redshift gives the velocity directly but R is deduced from an angular diameter and it is proportional to $1/h_0$, where

$$H_0 = 100h_0 \,\text{km s}^{-1}\,\text{Mpc}^{-1}\dagger. \tag{2.2}$$

The mean density of cluster matter is then proportional to h_0^2 just as is the critical density, ρ_{crit}. There are therefore no uncertainties in the contribution of clusters to Ω_0 in addition to those which we have already discussed. However, as we shall see in the next section, the required density in the form of baryons to give the correct element abundances does not involve h_0 so a view as to whether cluster hidden mass can (or must) be baryonic depends on the value of h_0.

No estimates of the hidden mass in clusters of galaxies lead to a value of Ω_0 which is equal to unity; something between 0.1 and 0.2 is a more characteristic value. This implies that, if Ω_0 does in fact equal unity, the space between clusters must be far from empty. As we shall see in the next section, if h_0 has its lowest generally suggested value (*ca.* 0.5), the inferred mass in clusters may not exceed the limit on the amount of baryonic matter allowed by the light element abundances. There is, therefore, no absolute necessity for the known hidden mass to

† 1 pc ≈ 30857×10^{12} m.

be non-baryonic, although this is not the case if $h_0 \approx 1$. This means that it is important to consider carefully what forms of baryonic matter are allowed by observations.

Consider first hidden mass in the Galaxy and in other galaxies. Bahcall & Casertano (1985) have pointed out that rotation curves are very similar from galaxy to galaxy and Rubin *et al.* (1985) have shown that they are independent of the morphological type of spiral galaxies. Bahcall & Casertano believe that this is itself strong evidence for the hidden mass being baryonic. The most obvious suggestion is that it is in the form of brown dwarfs or *Jupiters*, very low mass subluminous stars. This is regarded as a plausible solution for the hidden mass contribution to the Oort limit by Bahcall *et al.* (1985) but not by Reid & Gilmore (1984), who claim that the luminosity function turns over well before the lower mass limit of luminous stars is reached. Even if Reid & Gilmore are correct there could be a distinct distribution of subluminous stars, which cannot at present be ruled out observationally. The view that a luminosity function might sometimes be weighted to very low mass stars is supported by the observations by Fabian *et al.* (1984) of cooling flows in clusters of galaxies which are not producing very luminous stars. The present theoretical understanding of star formation is not adequate to rule out substantial variations in initial mass function.

A recent discussion of the possible forms of baryonic hidden matter has been given by Carr *et al.* (1984). The only other way in which stars could provide all of the matter, if h_0 is small, is through dead (black hole) remnants of massive objects between 10^2 and 10^6 $M_\odot$. Such black holes would be produced by the evolution of pregalactic Population III stars or possibly by the very first stars produced in galaxies. If such black holes contribute to the hidden mass in our Galaxy, they might be detected by the radiation produced by their accretion of interstellar matter. It appears that a significant contribution by black holes as massive as 10^6 $M_\odot$ can already be ruled out while the possibility that lower-mass black holes might also be detected has been the subject of recent studies by McDowell (1985) and Lacey & Ostriker (1985). A possible way of discovering compact objects at great distances from our Galaxy is that they might produce gravitational minilenses and this will be discussed by Rees. Another possible form of baryonic hidden matter is small solid particles. There seems to be no way in which these can be rejected as a result of observations. However, they must be regarded as an extremely implausible solution because they would require a very large production of heavy elements which are then tied up efficiently in solid particles and do not show in the general element abundances. This essentially rules them out.

Although all of the hidden matter might be baryonic if h_0 is small, cosmological considerations, to be discussed further in §3, rule this out if $h_0 \approx 1$. This has led Bekenstein & Milgrom (1984) to ask whether the observations necessarily imply the existence of large amounts of hidden mass. Milgrom (1983) had previously suggested a modification of Newton's law of gravitation on scales larger than that on which it has been directly tested in which the inertia of a body was a more general function of its acceleration than linear. Felten (1984) has argued that this modification of the law of gravitation raises more problems than it solves and that in particular it does not at present provide a framework for discussing cosmology. For our purpose it serves as a reminder that the laws of physics cannot be experimentally established for all conditions in the Universe. Another effective modification of the law of gravitation on scales much larger than those considered by Milgrom is the introduction of a cosmological constant, which is discussed in §3.

3. Cosmological constraints

I first discuss constraints within the context of the standard *hot* Big Bang cosmological theory and then comment on possible variants of the theory. The theory, of course, owes its origin to the observed redshifts in spectra of galaxies. The two key observations that led to its current popularity are the existence and near isotropy and black body form of the cosmic microwave radiation and the abundances of the light elements and isotopes ^{1}H, ^{2}H, ^{3}He, ^{4}He and ^{7}Li. As the Universe expands and cools, nuclear reactions produce the isotopes other than ^{1}H from neutrons and protons when the temperature is *ca.* 10^9 K. The precise abundances of the isotopes depend on the speed with which the Universe cools to the reaction temperature and on the number density of neutrons and protons (baryons) at the relevant time. The abundance of ^{4}He depends primarily on the rate of expansion of the Universe and those of ^{2}H, ^{3}He and ^{7}Li mainly on the baryon density. The baryon mass density today, ρ_{B0}, and the microwave temperature today, $T_{\gamma 0}$, are related to those from the time of nucleosynthesis onwards by $\rho_B/T_\gamma^3 \approx$ const., where the slight departure from precise equality is well understood; as we know $T_{\gamma 0}$, a knowledge of ρ_{B0} would determine the influence of baryon number on nucleosynthesis.

Particles other than baryons (and the accompanying electrons) and photons affect the comparison between theory and observation in two different ways. They may influence the speed of expansion of the early Universe and they may contribute to the total mass density of the Universe today. It is well known that the light element abundances cannot be explained within the standard model if the baryon density is sufficient to close the Universe. Thus, if $\Omega_0 = 1$, it is necessary for some matter to be in non-baryonic form. In the early Universe particles and antiparticles are in equilibrium with radiation and their number density is similar to that of photons, the precise value depending on whether they are fermions or bosons and on how many spin states they possess. As the temperature drops below the value $T \approx mc^2/k$ for any particle, the number density of the pairs drops as $\exp(-mc^2/kT)$ so that it is not long before both the kinetic energy and rest mass energy of the particles are negligible compared with the kinetic energy of the remaining relativistic particles. If the particles are stable, their number density *freezes* when they cease to interact effectively with other matter and radiation, we refer to this as decoupling. They then remain in the Universe until the present epoch and eventually they may make an important contribution to the mass density of the Universe because, once their annihilation ceases, the contribution of their rest mass to the energy density of Universe scales as R^{-3}, while the radiation density scales as R^{-4}; clearly they will only be important if their rest mass density exceeds that of the baryons, which behave in the same manner. Essentially the same considerations apply to very long-lived unstable particles. There are slightly different considerations for extremely massive particles such as magnetic monopoles whose density relative to baryons might be reduced by many orders of magnitude during an inflationary phase. The axion is another special case; it is light and very weakly interacting but it is never in thermodynamic equilibrium and never relativistic.

The known believed stable weakly interacting particles are neutrinos. They are thought to be of very low mass and this is certainly true of the electron neutrino. There are known to be three species of neutrino, electron, muon and tauon. They decouple from matter when $T \approx 1$ MeV$/k$. If all of their masses are significantly less than 1 MeV$/c^2$, we have the simplest description of the theory. At the time of nucleosynthesis the rate of expansion of the Universe is determined by the energy density in photons, electrons and positrons and neutrino–antineutrino pairs of three types. The relation between temperature and time is

$$t = (3c^2/32\pi G\alpha a)^{\frac{1}{2}}\, T^{-2}, \tag{3.1}$$

where the energy density is $\alpha a T^4$ and where α is $\frac{43}{8}$, if there are just three species of neutrino. If there is a further species of low mass neutrino, α is increased by $\frac{7}{8}$ and the Universe expands more quickly.

As explained earlier, the abundance of ^{4}He is primarily dependent on the expansion rate. The general accepted primeval value of the abundance by mass $Y_P \approx 0.23 \pm 0.02$ (see, for example, Pagel 1982; Kunth & Sargent 1983) can be in agreement with predictions of the theory if there are three species of low-mass neutrino and probably if only two of the known species prove to be of low mass; there are certainly no good constraints on the mass of ν_τ. The agreement between theory and observation might just remain if there were four species of light neutrino but it seems very unlikely if there are more; each additional neutrino species increases Y_P by *ca.* 0.015. The abundances of ^{2}H, ^{3}He and ^{7}Li can only be explained if $\Omega_{B0} \lesssim 0.1$; there is significant uncertainty in the critical value of Ω_{B0} both because the element production depends on ρ_{B0}, whereas Ω_{B0} contains h_0^2 whose value is not well known and because there are difficulties in deducing primeval abundances from present abundances. The primeval abundance of ^{4}He is itself uncertain for three different reasons; there are difficulties in reducing observations to obtain reliable abundances (Davidson & Kinman 1985); the whole picture would change if even one object was discovered with a reliable abundance below the currently accepted Y_P; it is only possible to observe helium in the oldest objects in galaxies but some helium synthesis might have occurred in pregalactic Population III stars. As explained in §2, if h_0 is small an allowable value of Ω_{B0} may include all the directly inferred hidden mass. The problem of light-element abundances will be discussed by Pagel.

Until recently, the only experimental information about the number of neutrino types arose from direct detection of the neutrinos. Now the discovery of the Z^0 boson which mediates the electroweak interaction gives further evidence. The width of the Z^0 is related to the number of possible decay channels which in turn depends on the number of neutrinos light enough to arise from Z^0 decay. Recent observations of the width of the Z^0 (Miller 1985) suggest that there may be 2 or 3 additional families of particle light enough to arise from its decay. If this is true (the present error bars on the result are rather large) and if all are neutrinos light enough to be relativistic at decoupling, this poses very serious problems for the hot Big Bang theory with regard to the light-element abundances, apart from any contribution which the neutrinos might make to the mass density of the Universe, which I shall discuss below. Of course, if there are more neutrino types this has very important implications for particle physics; presumably there are the same number of families of quarks and charged leptons, which must affect views about the form of any grand unified theory.

As mentioned above, all weakly interacting particles, whose mass density when they decouple is comparable to or greater than that of the baryons, make an important contribution to the mass density of the Universe. If the element abundances are to be correct, there must be mass in non-baryonic form if $\Omega_0 = 1$ and possibly if the mass in clusters of galaxies has its highest implied value. Non-baryonic hidden mass may be provided by the neutrinos if they have non-zero mass or by other at present hypothetical weakly interacting massive particles, such as axions, monopoles, photinos or other supersymmetric particles. If such particles have very low mass, they behave like additional neutrinos and affect the expansion of the Universe and the nucleosynthesis as well as the current mass density. If they have high mass, they are only important in ρ_0. It should be stressed, as has already been mentioned, that given their decoupling temperature the masses of the particles are constrained to within slightly more than one order of magnitude if they are to make an important contribution without leading to an incompatibility in the theory; clearly the constraint is almost precise if it is demanded that $\Omega_0 \equiv 1$. This implies that particle physics must provide particles of just the appropriate masses and interactions or that there must be families of particles which add up to the required mass value, proper account being taken of the number of spin states.

Early ideas of non-baryonic matter concentrated on the known neutrinos. Originally cosmology was used to provide constraints on neutrino masses, assuming that the cosmological theory was correct. Then the report of a measured mass for $\bar{\nu}_e \approx 30\ \mathrm{eV}/c^2$ (Lyubimov *et al.* 1980), which was uncomfortably large on the assumption that ν_μ and ν_τ are more massive, stimulated further interest. The contributions of neutrinos to the mass density can be written

$$\Omega_{\nu 0} = 0.4\left(\frac{n_\nu}{6}\right)\left(\frac{m_\nu c^2}{10\ \mathrm{eV}}\right)\left(\frac{T_{\gamma 0}}{2.7}\right)^3 \frac{1}{h_0^2}, \tag{3.2}$$

where n_ν is the number of neutrino types with neutrinos and antineutrinos counted separately and m_ν is the average neutrino mass. It was recognized that, if the hidden mass was in neutrinos, the distribution of baryons and neutrinos would differ in both galaxies and clusters and that this was what was needed to explain variation of mass:light ratio with position. It was also of interest to note that restrictions associated with Liouville's theorem and the Pauli exclusion principle permitted the right amount of mass to be accommodated in galaxies if neutrinos closed the Universe. It was also recognized that non-baryonic matter could influence the formation of structure in the Universe from an essentially structureless background and it is on this point that most recent interest has centred. Finally, if neutrino masses are found to be large enough that $\Omega_{\nu 0}$ is significantly greater than unity, there are problems for the standard theory.

The possible candidates for non-baryonic matter can be differentiated according to whether, like neutrinos, they are relativistic and thus effectively massless at the time when they decouple from matter or whether they are non-relativistic. The relativistic particles retain essentially the same temperature as the photons and the baryons, while the non-relativistic particles are colder and this means that the Jeans mass is different when matter and radiation decouple and self-gravitating fluctuations begin to enhance and form galaxies and clusters. If the hidden matter is neutrinos the first structures are of supercluster size but they are much smaller if the matter is cold. Much recent work has been concerned with which form of hidden matter is best able to produce the observed clustering properties of galaxies and whether or not this is compatible with $\Omega_0 = 1$. There have been claims (White *et al.* 1984) which have in turn been disputed (Melott 1985) that the observed galaxy clustering cannot arise if the hidden mass is in the form of neutrinos and that *cold dark matter*, which is non-relativistic at decoupling, is better able to give agreement. Even then $\Omega_0 = 1$ would appear to require the observed peaks of light not to be faithful tracers of where the mass is concentrated; this has led to the development of various ideas of biased galaxy formation in which galaxies only form at the highest mass peaks. This subject will be discussed by Frenk. It must be stressed that if the elementary-particle

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

physicist discovers more than one species of weakly interacting elementary particle, the cosmologist is not able to choose which one he will include; they must all be included. It is, of course, necessary to remember that the particles might have effects in branches of astronomy which are only remotely related to cosmology. One example is the suggestion that the centre of the Sun might contain such particles, which would influence and possibly resolve the solar neutrino problem (Steigman *et al.* 1978; Spergel & Press 1985).

It is possible that the standard model of the hot Big Bang will prove to be correct, by which I mean that there will be no gross discrepancies between the predictions of the completely homogeneous and isotropic model and observations and that an explanation can be obtained for those small departures from uniformity that are necessary to give an understanding of the observed structure. This might not be true and it is therefore necessary to ask whether there are any simple modifications of the theory which would provide the agreement. There are two, neither of which tends to be greeted with great enthusiasm by particle physicists. They address two separate problems; the possibility that the permissible theoretical age of the Universe according to the theory is less than the observed age of galactic objects and the failure of light element abundances to be in agreement. Essentially all estimates of ages of galactic globular clusters exceed $1/H_0$ if $h_0 = 1$ (see, for example, Tayler 1984).

The possible solution to the first problem is the introduction of a cosmological constant, Λ, into Einstein's equations. Einstein originally introduced Λ with the value $4\pi G\rho_0$ to produce a static world model. If a somewhat larger value of Λ is used, the Universe has a period of very slow expansion before entering the present observed stage of more rapid expansion. This means that the age of the Universe can exceed the reciprocal of Hubble's constant. In particular, a closed Universe can be reconciled with the ages of galactic objects even if H_0 has a small value. Although physicists are not completely happy with Λ being precisely zero, they also have no reason to expect that Λ should have a value of a few times $4\pi G\rho_0$, a value that enables it to be just important today when it has been unimportant for most of the past history of the Universe. It provides another problem of the type which the inflationary model was introduced to solve. It should perhaps be stressed that, if the suggestion that $\Omega_0 \equiv 1$ is ignored, the most attractive solution is to have $\Lambda = 0$, $h_0 = 0.5$ and all the mass baryonic, although this may raise problems with galaxy clustering.

The next possibility raises something which has not been mentioned explicitly earlier. The standard model assumes that the net baryon number and lepton numbers of the Universe are essentially equal. This would be the case for example, if at the present time the number of neutrinos and antineutrinos of each type in unit volume are equal and if the number of electrons equals the number of protons whether free or in compound nuclei. It is however possible to suppose that the Universe contains a larger number of one or more of the types of neutrino. If there is a large excess of either electron neutrinos or antineutrinos, the neutron: proton ratio at the time of nucleosynthesis is affected and the production of ^{4}He changed substantially. An excess of any type of neutrino speeds up the expansion of the Universe and also affects nucleosynthesis of all the light isotopes and it seems likely that there are enough free parameters to provide agreement with observations. The problem here is that the observed baryon: photon ratio is about 10^{-10}–10^{-9} and attempts have been made to explain this figure as a result of baryosynthesis in the very early Universe when grand unified particle theories break the conservation of baryon number. However, in this case one would expect lepton synthesis to be about as effective as baryon synthesis whereas a lepton to photon ratio of close to unity or greater would be needed to change nucleosynthesis significantly.

Finally, it must be mentioned that there are from time to time attempts to understand the structure of the Universe outside the context of the hot Big Bang, in particular in terms of a cold Big Bang. In such a theory the microwave radiation and the light elements must be produced by Population III stars which form before galaxies and the obvious constraints on the amount of baryonic matter in the Universe no longer hold. There are great difficulties in producing the light elements without overproducing heavy elements and in thermalizing stellar radiation to produce the microwave background, but these may not prove insuperable. It must be stressed that *cold* is not equivalent to low energy in the early Universe. In its very high density state matter is very degenerate and some high-energy phenomena involving elementary particles can still occur. However, essentially all of the detailed discussion of hidden matter in the context of cosmology would have to be modified significantly.

References

Bahcall, J. N. & Casertano, S. 1985 *Astrophys. J.* **293**, L7–10.
Bahcall, J. N., Hut, P. & Tremaine, S. D. 1985 *Astrophys. J.* **290**, 15–20.
Bailey, M. E. 1982 *Mon. Not. R. astr. Soc.* **201**, 271–280.
Bekenstein, J. & Milgrom, M. 1984 *Astrophys. J.* **286**, 7–14.
Carr, B. J., Bond, J. R. & Arnett, W. D. 1984 *Astrophys. J.* **277**, 445–469.
Davidson, K. & Kinman, T. D. 1985 *Astrophys. J. Suppl.* **58**, 321–340.
Faber, S. M. & Gallagher, J. S. 1979 *A. Rev. Astr. Astrophys.* **17**, 135–187.
Fabian, A. C., Nulsen, P. E. J. & Canizares, C. R. 1984 *Nature, Lond.* **310**, 733–740.
Felten, J. E. 1984 *Astrophys. J.* **286**, 3–6.
Fernley, J. A. & Bhavsar, S. P. 1984 *Mon. Not. R. astr. Soc.* **210**, 883–890.
Gott, J. R. III & Turner, E. L. 1977 *Astrophys. J.* **213**, 309–322.
Guth, A. H. 1982 *Phil. Trans. R. Soc. Lond.* A **307**, 141–148.
Kunth, D. & Sargent, W. L. W. 1983 *Astrophys. J.* **273**, 81–98.
Lacey, C. G. & Ostriker, J. P. 1985 *Astrophys. J.* **299**, 633–652.
Lucey, J. R. 1983 *Mon. Not. R. astr. Soc.* **204**, 33–43.
Lyubimov, V. A., Novikov, E. G., Nozik, V. Z., Tretyakov, E. G. & Kosik, V. S. 1980 *Phys. Lett.* B **94**, 266–268.
McDowell, J. C. 1985 *Mon. Not. R. astr. Soc.* **217**, 77–85.
Melott, A. L. 1985 *Astrophys. J.* **289**, 2–4.
Milgrom, M. 1983 *Astrophys. J.* **270**, 365–370.
Miller, D. J. 1985 *Nature, Lond.* **317**, 110.
Oort, J. H. 1932 *Bull. Astr. Insts. Neth.* **6**, 249–287.
Pagel, B. E. J. 1982 *Phil. Trans. R. Soc. Lond.* A **307**, 19–35.
Rees, M. J. 1985 *Mon. Not. R. astr. Soc.* **213**, 75p–81p.
Rees, M. J. 1986 In *Dark matter in the Universe* (*IAU symposium*, no. 117) (ed. G. Knapp & J. Komendy). Dordrecht: D. Reidel. (In the press.)
Reid, N. & Gilmore, G. 1984 *Mon. Not. R. astr. Soc.* **206**, 19–35.
Rubin, V. C., Burstein, D., Ford, W. K. Jr & Thonnard, N. 1985 *Astrophys. J.* **289**, 81–104.
Schaeffer, R. & Silk, J. 1985 *Astrophys. J.* **292**, 319–329.
Schramm, D. N. 1982 *Phil. Trans. R. Soc. Lond.* A **307**, 43–54.
Spergel, D. N. & Press, W. H. 1985 *Astrophys. J.* **294**, 663–673.
Steigman, G., Sarazin, C. L., Quintana, H. & Faulkner, J. 1978 *Astron. J.* **83**, 1050–1061.
Tayler, R. J. 1984 In *Observational tests of the stellar evolution theory* (ed. A. Meader & A. Renzini), pp. 549–562. Dordrecht: D. Reidel.
Tully, R. B. & Shaya, E. J. 1984 *Astrophys. J.* **281**, 31–55.
Turner, E. L., Aarseth, S. J., Gott, J. R. III & Mathieu, E. D. 1979 *Astrophys. J.* **228**, 684–695.
White, S. D. M., Davis, M. & Frenk, C. S. 1984 *Mon. Not. R. astr. Soc.* **209**, 27p–31p.