SELF-SIMILAR COSMOLOGICAL MODELS

by

DAVID ALEXANDER B.Sc.

Thesis

submitted to the

University of Glasgow

for the degree

of Ph.D.

Department of Physics and Astronomy,

The University,

Glasgow G12 8QQ.

November 1988

For my Father

A curious thing about tensors

is tensors have traces and norms.

Their tops are made out of vectors.

Their bottoms are made out of forms.

There's stress and pressure

and one that measures

the distance from P to Q.

But the most wonderful thing about tensors,

is the one called $G_{\mu\nu}$.

D. Alexander and A.G. Emslie,

composed in ''The Aragon'',

Byres Rd., Glasgow, May 1987.

CONTENTS

Page

PREFACE

In this thesis, cosmological models, which admit self-similar symmetries, are examined. Symmetries in cosmology have become increasingly important since the formulation of the Einstein field equations demonstrating the close correspondence between geometry and physics. Physical investigations of gravitational systems have benefited greatly from the applications of geometric techniques and, in particular, from the consideration of various symmetries which lead to a simplification of the relevant equations. One such symmetry is that of self-similarity, which was initially developed as a physical symmetry in the study of hydrodynamics and is now associated with the more global symmetries of conformal and homothetic motions. Self-similar symmetry is particularly useful in the study of cosmology, since the Universe can be treated as a hydrodynamic fluid (or a geometrical manifold) in which there are no characteristic scales and so may be expressible by the techniques of self-similarity.

In Chapter 1, a general review of the current developments in the study of cosmology is given. In particular, the 'intrusion' of particle physics into the realm of cosmology is discussed. The application of particle physics theories, which has resulted in a better understanding of the Universe at early epochs and has helped to explain the origin of large-scale structure in the Universe, is also outlined briefly. Finally, exotic cosmological theories which attempt to go beyond Einstein's theory of general relativity are addressed.

The application of differential geometry to cosmology is discussed in Chapter 2. The description of the Universe as a four-dimensional manifold is considered and the various symmetries which may be imposed on such a manifold are described. The symmetry of self-similarity is then introduced together with a discussion of its development in hydrodynamics and its applications to the study of the dynamics of the Universe.

Spatially-inhomogeneous spacetimes with a non-zero cosmological constant are investigated in Chapter 3. The role of self-similarity is discussed and solutions of this spacetime, which admit a similarity symmetry, are considered. Integrals of the motion are determined and these are related to the degree of anisotropy and inhomogeneity of any given solution. The solutions found are then discussed in the context of the cosmological "no-hair" theorems which consider the effect of a large vacuum term on the expansion of the Universe.

In Chapter 4, the effects of viscosity and shear on the evolution of a cosmological model are considered. A self-similar analysis is carried out in which the viscosity coefficients vary in a prespecified manner, and two different classes of solution are investigated. These solutions differ in the choice of the equation of state, which is chosen to represent the extreme cases of a 'viscous dust' and a 'stiff' Universe. The self-similar stiff solutions are then developed to consider the growth of primordial black holes in the early Universe in Chapter 5.

Chapter 6 includes a brief review of the work of the thesis and suggests a few interesting lines for future research.

The original work of this thesis is contained in Chapters 3-5 and also in the Appendix. The contents of Chapter 3 have been accepted for publication in *Monthly Notices* and different aspects of this work have also appeared in various conference proceedings. The work of Chapters 4 and 5 is currently being developed for publication.

ACKNOWLEDGEMENTS

The work in this thesis was carried out while the author was a research student in the Department of Physics and Astronomy, University of Glasgow and a research assistant in the Department of Physics, University of Alabama in Huntsville. I extend my thanks to the staff and students of both these departments, who have made my last three years both interesting and enjoyable. I would also like to thank them for tolerating my numerous inane prattlings and stupid questions.

I would like to thank my joint supervisors, Robin Green in Glasgow and Gordon Emslie in Huntsville, for providing a calming influence or a kick, whenever one or the other was required. I am indebted to both Robin and Gordon for their friendship, encouragement and perserverance. I also thank Gordon for introducing me to the delights of Huntsville-Alabama, parachuting, Wendys and Margueritas (Long Live Casa Gallardo!!!), and for providing me with the occasional foray into the world of solar physics.

Other members of the department deserve special mention. A standard, and well deserved acknowledgement, in Astronomy theses from Glasgow, is to Professor P. A. Sweet and, like my predecessors, I am grateful to 'The Prof.' for his encouragement and for the many discussions of things astronomical and physical. Thanks: to Professor John Brown for helping me to procure funds for my trips to Alabama, Oxford and Cambridge and for providing me with the opportunity to continue my career in research; to Drs. John Simmons and Ken McClements (now at Oxford) for allowing me to play squash under the pretence of 'collecting data'; to Drs. Alex MacKinnon and Bert van den

Oord for straightening out my naive attempts at understanding solar physics; to Dr. Alan Thompson for sharing his knowledge of the multifarious computer systems which abound in the Astronomy Group; to Mr. Andrew Liddle for being another cosmologist, although on the other side of the fence; to Daphne, whose organisational and coffee-making abilities have helped smooth the way for many a student's research; to Christine for typing the page numbers.

I would like to thank Professor I. S. Hughes for allowing me to carry out my research in the Department of Physics and Astronomy and to the Science and Engineering Research Council for providing the necessary funding.

Thank-you to all my friends and family; to the Physics 'mob' for their sarcasm, wit and general abhorrence of discussing physics on a Friday night; to my good friends Charmaine, George and 'Aunty' Phamie for feeding me on Sundays and for keeping me in touch with the real world; to Ladye Wilkinson, for fun, friendship, Pină Coladas and hospitality in Alabama (a bold woman if ever there was one). In particular, I would like to thank my mother, Cathy, and my brothers and sisters, Thomas, Angela, Paula, Robert and Tracy, for introducing some normality and, if not sanity, at least a different form of insanity into my life. To Jack and Margaret Sellar, I extend my warm appreciation for everything they have done for me in the last six years.

Finally, this thesis would not have been possible without the love, support, proof-reading, cooking and tolerance of my fiancé Carolyn Sellar. I cannot thank Carolyn enough and I hope I can give as much in return.

David Alexander

SUMMARY

The Universe today is observed to be extremely homogeneous and isotropic on large scales. The dipole anisotropy of the microwave background, due to the relative motion of the Earth, is measured to be less than one part in 10^4. The quadropole component, due to intrinsic anisotropies, is even smaller. Thus, any viable mathematical or physical description of the large scale properties of the Universe must encompass the observational evidence and reflect this large degree of uniformity.

The most popular, and certainly the most successful, description of the Universe at the present epoch is provided by the Friedmann-Robertson-Walker (FRW) cosmological models. These spherically symmetric models consider the Universe as an isotropic, spatially homogeneous, perfect fluid matter distribution, which is in a state of dynamic evolution. All of the FRW cosmologies exhibit an expansion, i.e. the volume of the spatial sections varies with time, during some stage of their evolution, in agreement with the observed expansion of the Universe. An important consequence of this behaviour is that it leads to a singularity at a finite time in the past when the volume of the spatial sections becomes zero and matter becomes infinitely dense and infinitely hot (the hot Big Bang scenario). The isotropy and homogeneity of the Universe at the present epoch, cannot necessarily be extrapolated back to these earlier times. Certainly, there must exist inhomogeneities on small scales at all epochs in order to produce the observed structure, such as galaxies, clusters and superclusters. This raises the question of the effect of anisotropy on the initial stages of the evolution of the Universe.

In this thesis we consider cosmological models which differ significantly from the FRW descriptions. We consider the effect of a large cosmological constant (vacuum energy term) on the behaviour of a spherically symmetric anisotropic universe, characterised by different expansion rates in the radial and transverse directions. The analysis is simplified considerably by imposing the condition that the model admits a self-similar symmetry. The techniques of similarity and dimensional analysis are employed to obtain a class of spatially inhomogeneous solutions to the Einstein field equations with a non-zero cosmological term. These solutions are found to contain some which tend asymptotically to a de-Sitter FRW solution and thereby extend the cosmological "no-hair" theorems, which state that under certain restrictions any model containing a large positive cosmological term will evolve to a de-Sitter cosmology at late times. Such models are attractive since they tend to isotropic spacetimes.

Similarity methods are also applied to the study of an anisotropic spacetime with an imperfect fluid as source. The fluid description of the cosmology is chosen to include the dissipative processes of shear and bulk viscosity but to neglect the effects due to the existence of magnetic fields, heat conduction or acceleration along the flow lines. In order to obtain a self-similar description of such a fluid we must impose certain conditions on the form of the viscous coefficients of bulk and shear. This allows a degree of tractability but restricts the physical significance of the models. Solutions are found for which the matter distribution acts as (i) a 'presureless fluid' with an equation of state given by $T^1{}_1=0$ and (ii) a 'stiff' fluid with equation of state, $T^1{}_1=-T^0{}_0$. The conditions under which the Universe may attain either of these extreme properties are discussed in relation to the physical processes occurring in the matter distribution at different epochs. It is found that the presence of viscosity has a marked effect on the dynamics of the Universe, particularly at early times.

The self-similar viscous models with a stiff equation of state are then considered with respect to the formation of black holes in the early Universe. The difficulties of obtaining a smooth continuation of the viscous solutions from the Universe particle horizon to a black hole event horizon are discussed in view of the limitations encountered in the non-viscous black hole solutions.

Finally, the possibility of future investigations inspired by the considerations of this thesis are discussed. In particular, the determination of a geometric symmetry corresponding to self-symmetry of the second kind and the formation of a self-consistent similarity treatment of imperfect fluid cosmologies are deemed important. Possible lines of research to these ends are considered.

1. REVIEW OF CURRENT COSMOLOGICAL IDEAS

1.1 Preamble

The study of cosmology has relatively recent origins. Its beginning must be placed around 1929, when Hubble discovered the expansion of the universe. The most amazing thing about this discovery is the universality of this expansion. *All* galaxies are moving away from us, and moving with velocities which increase with their distance. Therefore, the observed expansion is neither a local phenomenon nor a random statistical event. The whole universe expands, all the galaxies move away from each other with enormous velocities which, at large distances, approach the speed of light. This discovery verified some of the most daring predictions of Einstein's theory of relativity.

The first theorists to construct models of an expanding universe, using general relativity, were de-Sitter (1917), Friedmann (1922) and particularly Lemaitre (1927). However, the majority of astronomers did not take these models seriously since the concept of a dynamic universe was contradictory to most beliefs at the time. This situation lasted until the discovery of Hubble (1929), after which the whole previous way of thinking was altered. For the first time the study of the universe as a whole became the object of serious physical research, subject to observational constraints. The great advancement of cosmology that followed was due to systematic research in observations *and* theory. Hubble initiated a large scale study of the universe, starting from the nearby galaxies. Galaxies can be regarded as the basic ingredients of the universe, its "atoms". Modern astronomical techniques have taken the subject far beyond the nearby galaxies to distant objects from which light may take billions of years to reach us.

The subject of cosmology is concerned mainly with this extragalactic world. It is a study of the large-scale nature of the universe extending to distances of giga-parsecs, a study of the overall dynamic and physical behaviour of a myriad of galaxies spread across vast distances and of the evolution of this enormous system over several billion years.

1.2 The Observable Universe

It became clear from the catalogue of the positions of bright galaxies, compiled by Shapley and Ames (1932), that the galaxies segregated into compact clusters, many of which appeared spherically symmetric. Abell (1958) chose a homogeneous sample of such clusters and noticed that apart from clusters of galaxies there are clusters of clusters of galaxies due to second-order clustering. That is, he noticed that rich clusters had a tendency to segregate into larger structures, called *superclusters*, whose size were of the order of 50Mpc, (1Mpc $\equiv 3\times10^{22}$m). A typical galactic scale is 30-50kpc. These superclusters have as many as 10 rich clusters and masses of 10^{15}-$10^{17}M_\odot$. The largest superclusters could have as many as 10^5 member galaxies.

There is evidence that there is a supercluster around our own galaxy, called the Local Supercluster. It is a flat ellipsoidal system of 15Mpc cross-section and 1Mpc thickness, which includes the local group of galaxies. Studies of the radial velocities and dynamics of the 50,000 galaxies in the Local Supercluster show that it is rotating and expanding. From the rotation of this system it is estimated that its total mass is $10^{15}M_{\odot}$. The whole local supercluster also moves with respect to other distant superclusters with a velocity of the order of $500kms^{-1}$. Most galaxies belong to such large dynamic structures. In fact, it is estimated that ≈90% of all galaxies belong to clusters and superclusters. The clusters of galaxies are generally spherical in shape whereas almost all superclusters are flat. For this reason superclusters are sometimes called "pancakes", after Zel'dovich (1970).

At first glance the observations of clusters and superclusters indicate that their distribution is random. However, after detailed observation and data analysis it becomes clear that their distribution is not uniform. There exist huge areas in the universe which contain almost no galaxies. These are appropriately called "voids". It is estimated that only 10% of space is occupied by superclusters while the rest does not contain any luminous matter. The voids may reach dimensions of up to 100Mpc. Recent theories on the structure of superclusters can, in fact, explain the creation of condensations of matter in spherical form and also the formation of large scale filamentary and flat structures, (Peebles 1965, Zel'dovich 1970 and Saarinen et al. 1987). Numerical simulations of the large-scale structure, give a honeycomb structure, in which the cells of the honeycomb are the voids and the walls are the superclusters, (White et al. 1987).

The universe as a whole appears to be isotropic and homogeneous on very large scales. Isotropy means that the universe looks the same in every direction, homogeneity means that the universe will appear the same to any observer, independently of their position in the universe. In other words, all observers will measure the same density and generally the same properties of the universe.

This is termed the 'cosmological principle'. A proof of the homogeneity of the distribution of galaxies is based on the observation that the number of galaxies up to a magnitude m+1 is about four times the number of galaxies up to magnitude m, as is expected if the density of galaxies in space is constant. On small scales, the distribution of galaxies is inhomogeneous but becomes increasingly more homogeneous as the scale increases. The greatest degree of homogeneity is exhibited by the microwave background radiation (MBR).

The MBR was first discovered by Penzias and Wilson (1965) and became one of the major cosmological discoveries of all time. This radiation is the remnant of the radiation of the early universe and uniformly fills the whole of space. It is an extremely diffuse radiation which comes uniformly from all directions and corresponds to a black body spectrum of approximately 3°K. The implication of this is that the MBR is not due to stars or galaxies but to the early concentration of matter in the universe when its temperature was about 3000°K, at which temperature hydrogen recombined and the mean free path of photons in the universe became as large as the horizon. Evidence for the origin of this radiation comes from the fact that its spectrum is very nearly that of a black body. The temperature of the MBR has been cooled by the general expansion of the universe.

Thus, on very large scales, the universe is observed to be extremely isotropic and homogeneous and any viable cosmological model must contain this exactly or at least in the limit. On the scale of superclusters ($\simeq$50Mpc), however, observations show that the universe has some filamentary and bubble-like structure, cf. Figure 1.1, making it generally inhomogeneous. There are several theories which attempt to explain this large-scale structure.

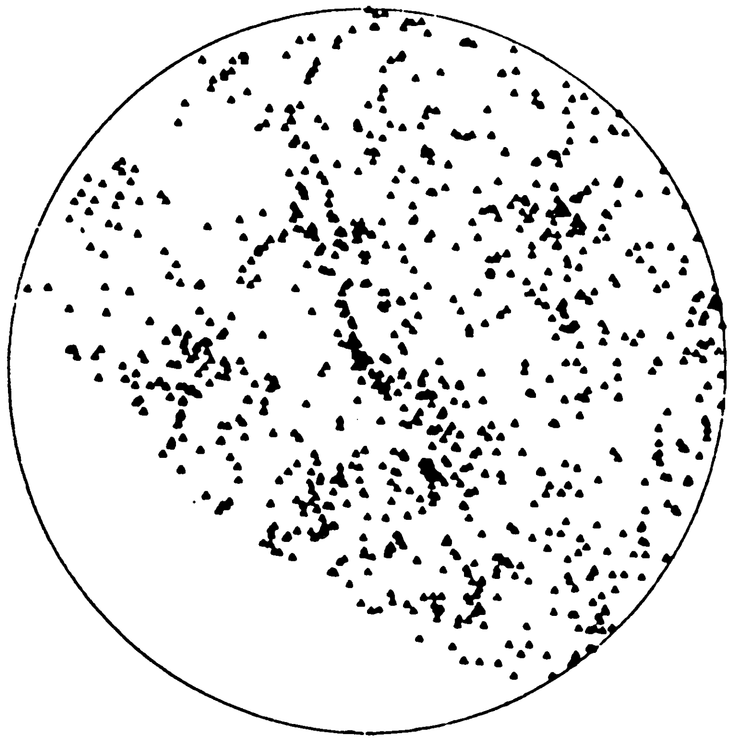


Figure 1.1 Equal area projections of the galaxy distribution in the northern sky from the CfA survey volume limited to 4000 kms^{-1} (cf. White 1987).

1.3 Theories for Large-Scale Structure in the Universe

The so-called "Big Bang" model of the universe has been extremely successful. It describes how all matter and energy came into existence at a single point in space and time and then expanded, which quite naturally explains why other galaxies seem to be rushing away from our own. The model also predicts that the universe is filled with a low level of radiation, left over from the big-bang, and the observation of this radiation by Penzias and Wilson became the first major success of the big-bang model. This and subsequent successes, such as predicting the abundance of the elements, has given us the confidence to trace the history of the universe back into the first second of existence.

One of the questions that remains unanswered, however, is why the universe is "lumpy" and how it got that way. The big bang model treats the universe as completely smooth and uniform. We described in the previous section that, on a very large scale, matter does indeed appear to be spread out evenly everywhere. However, on smaller scales, a great deal of structure exists. Recent observations reveal structures such as huge empty regions ('voids'), the largest ≈60Mpc in diameter (Kirschner et al. 1983), giant 'filaments', i.e. roughly linear overdense regions in the distribution of galaxies about 100Mpc long and 5Mpc across (Giovanelli and Haynes 1982) and in more complete surveys most galaxies appear to lie on the surfaces of 'bubbles', ≈50Mpc across (de Lapparent et al. 1986).

Before we consider the formation of these complicated large-scale structures we must address the problem of how the basic matter condensations are formed. We shall concentrate on the mechanism of gravitational instability based on the work of Jeans (1902). [Another mechanism for the initial formation of matter condensations was that of turbulence in the early universe, proposed by Weizsäcker (1951): However, serious arguments have been advanced against this theory].

If we consider the universe to be uniformly filled with gas, then a small local perturbation in the density may be enhanced or damped. Indeed, a local density excess will grow as it causes a stronger local gravitational field which will tend to attract even more matter. On the other hand, the gas pressure will tend to disperse any density enhancement and restore the initial homogeneity. Jeans (1902) found that small scale perturbations are quickly dispersed, while large scale perturbations become enhanced. In the latter case, the density in the perturbation increases continuously with time. This is called *gravitational instability* or *Jeans instability.* Such an instability finally creates a concentration of matter which may evolve to form a star, a galaxy or even a cluster of galaxies. The amount of matter condensed in this way depends on the initial density of the gas and the local sound speed, at which speed the local density perturbations propagate. The minimum mass required for the onset of gravitational instability is called the *Jeans mass,* M_j, and its radius is known as the *Jeans length,* λ_j. In a sphere of radius greater than λ_j, gravity overcomes the gas pressure and causes a concentration of matter. The reverse happens for a sphere with radius less than λ_j, i.e. the pressure of the gas overcomes gravity and the perturbation is damped.

Before the time of the recombination of hydrogen, the Jeans length was very large because the sound speed at that time approached the speed of light, since matter and radiation were strongly coupled. The Jeans mass increased until shortly before recombination, when it was $\approx 10^{17} M_{\odot}$, much greater than the mass of galactic clusters. After recombination, matter and radiation ceased contributing to the pressure and the sound speed suddenly fell to a few kilometres per second. The corresponding Jeans mass also dropped to $10^5 M_{\odot}$, that is, comparable to the mass of a globular cluster.

We distinguish two extreme types of matter condensations which form via gravitational instability: (a) isothermal, and (b) adiabatic. In the former case, the temperature inside the perturbation is the same as the cosmic temperature. This is achieved by the free movement of photons which remain uniformly distributed while the matter is clumped. In the latter case, the ratio of photons to baryons is the same inside and outside the perturbation, so that the temperature increases along with the density. Each type of perturbation has its own implications for the future evolution of structure in the universe. Consequently, two diifferent theories have been proposed for the formation of galaxies and clusters of galaxies.

The theory of isothermal perturbations has been mainly proposed by Peebles (1965). It states that any isothermal perturbation in the initial distribution of matter in the universe does not evolve before the time of recombination, t_{rec}. The perturbations simply follow the expansion of the universe. After t_{rec}, however, every perturbation which is greater than the Jeans mass starts to grow, i.e. every perturbation $\geq 10^5 M_{\odot}$, the mass of a globular cluster. After the formation of these condensations we have two opposite effects proceeding together. On the one hand, these clusters break up into small condensations which ultimately form stars (fragmentation). On the other hand, the same clusters concentrate in larger and larger groups, which make up the galaxies, groups of galaxies, clusters and superclusters. The larger the scale of concentration, the longer it takes to be formed.

There are two basic arguments supporting this theory. The first is that a study of observational data on the galaxy distributions shows that there are no distinctive scales for groups of galaxies up to superclusters, cf. Peebles (1980). The second argument is based upon numerical calculations done with models of the expanding universe (e.g. Aarseth et al. 1979). They found that points initially uniformly distributed (each point representing a galaxy) tend to segregate into groups which tend to increase in size as the universe expands. An interesting aspect of this theory is that large areas devoid of galaxies are formed in between the concentrations of galaxies. As time passes, the voids increase in size while the concentrations become more compact. This picture, therefore, seems to explain several characteristics of the observed universe.

The theory of adiabatic perturbations was proposed by Zel'dovich (1970) and his collaborators. A characteristic property of adiabatic perturbations is that small condensations are destroyed by viscosity during the epoch before recombination. Only concentrations more massive than $10^{13}M_{\odot}$ can survive until t_{rec}, when they can collapse since they are then greater than the Jeans mass. According to this theory, the large scale structure (superclusters) formed first, while the structure on smaller scales (clusters etc.) were formed later by the fragmentation of these initial concentrations. This is exactly the inverse process to the one suggested by the theory of isothermal perturbations. The superclusters, according to this theory, are not even approximately spherical but are flat like "pancakes". For this reason, this theory is also known as the "pancake theory". These "pancakes" eventually fragment into galaxies. (The Peebles picture is often referred to as the *bottom-up* scenario whereas the Zel'dovich picture is called the *top-down* scenario).

This theory has several attractive characteristics. For example, in a picture of the distribution of galaxies (Figure 1.1) one can distinguish several elongated structures, consisting of galaxies, groups of galaxies and clusters, reminiscent of Zel'dovich's pancakes. The numerical experiments are consistent with this picture if we assume that the points represent particles, rather than galaxies.

It is premature to say which of these two theories best describes the formation of galaxies and groups, clusters and superclusters of galaxies. What is common to both is that the initial perturbations in the distribution of matter in the universe, which led to the formation of galaxies, were very small and appeared during the first stages of the expansion of the universe, (they probably existed already at the Planck time, i.e. $t_p \simeq 10^{-43}$s). Thus, the isotropy and homogeneity of the early universe, according to these theories, was almost exact.

One current field of study, which incorporates the above theories in order to produce the large-scale structure, is that of the theories of dark matter in the universe. Dark matter is the 'unseen' matter which most astronomers believe surrounds the luminous stars and galaxies and makes up the vast bulk of the mass of the universe. Dark matter betrays itself by the gravitational effect it has on the matter we can see. Observational evidence shows that dark matter is present on all distance scales, from within the close neighbourhood of the sun to the rotation of galaxies themselves, in the dynamics of clusters and superclusters and also in the expansion of the universe, (cf. Kormendy and Knapp 1987).

At present, dark matter is in a non-gaseous, effectively collisionless form, and therefore the evolutionary phases of the structure in the universe can be studied, quite easily, by N-body numerical methods (e.g. White et al, 1984). There are essentially two forms that this dark matter can take, i.e. it can be composed of baryonic or non-baryonic matter.

Baryonic Dark Matter

Primordial nucleosynthesis constrains the fraction of density of the universe contributed by baryons (luminous and dark) to be

$$0{\cdot}056h^{-2} \leq \Omega_{baryons} \leq 0{\cdot}14h^{-2} \quad , \qquad (1.1)$$

where the Hubble constant is chosen to be $H_0=50h(kms^{-1}Mpc^{-1})$ (observationally $1 \leq h \leq 2$) and $\Omega=\rho/\rho_C$, the ratio of the density of the universe to the critical (or closure) density. Since the measured value of Ω, dynamically, is $\approx 0{\cdot}1-0{\cdot}2$, then there may not be a problem and all of the dark matter could be baryonic in the form of 'Jupiters' or black holes. Such constituents of dark matter may eventually be detected.

Non-Baryonic Dark Matter

Belief in the inflationary universe scenario (see §1.4) strongly biasses most cosmologists and there is almost universal agreement that our universe should be flat with $\Omega=1$. Due to the constraints imposed on baryonic matter, discussed above, this seemingly suggests that most of the matter in the universe is non-baryonic. Also the existence of galaxies and clusters today requires that perturbations in the density must become non-linear before the present epoch. In a baryonic universe, for adiabatic perturbations at recombination, this implies that present-day fluctuations in the microwave background radiation should be much larger than present observational upper limits.

One of the currently fashionable possibilities is that the dark matter consists of relic WIMPs (Weakly Interacting Massive Particles) left over from the very hot, early epoch of the universe. The early universe and modern particle theories working together have provided a very generous list of candidates for the dark matter (cf. Turner 1987). For the purpose of illustration we shall consider only two.

The standard model of particle physics, [a gauge theory which undergoes spontaneous symmetry-breaking at a temperature $T \approx 300$ Gev: $SU(3) \times SU(2) \times U(1) \longrightarrow SU(3) \times U(1)$, $(1\mathrm{GeV} \simeq 10^{13}\mathrm{K})$], supplies no candidates, other than the rather exotic quark nuggets, for dark matter beyond ordinary baryons in some non-luminous guise. Virtually all extensions of the standard model provide us with a generous supply of dark matter candidates. The two we shall consider are massive neutrinos and axions. [Table 1.1 provides a summary of the conversion scales between temperature, energy, size of the universe and time after the big bang for a hot big bang model].

Massive neutrinos are a product of the standard model extension known as the Majoron model. (The symmetry broken in this theory is the lepton number). We know neutrinos exist and if they have a mass then they would seemingly be ideal dark matter candidates. If neutrinos are massive then it can be shown that the density of neutrinos relative to the critical density is given by

$$\Omega_\nu = 0 \cdot 12(n_\nu/n_\gamma)h^{-2}\sum(m_\nu)_{eV} \qquad , \qquad (1.2)$$

Temperature	Energy	Size of Universe	Time after big bang (s)	Remarks
3K	3.10^{-4}eV	1	$\approx 10^{18}$	Present epoch
3000K	0.3eV	10^{-3}	10^{13}	Recombination of Hydrogen
10^{9}K	0.1MeV	10^{-9}	100	Big-Bang Nucleosynthesis
10^{11}K	10MeV	10^{-11}	0.01	
10^{13}K	1GeV	10^{-13}	10^{-6}	Quark/Hadron transition
10^{15}K	100GeV	10^{-15}	10^{-10}	End of electro-weak unification
10^{27}K	10^{14}GeV	10^{-27}	10^{-34}	End of grand unification
$>10^{31}$K	$>10^{19}$GeV	$<10^{-31}$	$<10^{-43}$	Planck era - Quantum gravity

Table 1.1 Conversion factors between temperature and energy for significant times in the history of the very early universe.

where n_γ is the photon number density in the microwave background and we are summing over the number of neutrino species. In the standard model $n_\nu/n_\gamma \simeq 3/11$. So one species of mass $\approx 25h^2$eV suffices to give Ω_ν=1.

Another popular, though more exotic, dark matter candidate is the axion. Peccei and Quinn (1977) proposed extending the standard model further by adding one additional Higgs doublet to the Majoron model. (We can add as many scalar (or Higgs) fields to the theory as we like by relating them to the free energy of the system). This extension introduces another symmetry (the PQ symmetry) which is also spontaneously broken. The existence of such a broken symmetry leads to a new light pseudoscalar boson called the axion.

The mass of the axion, its lifetime and its coupling to ordinary matter are all determined by the symmetry-breaking scale of the PQ symmetry, f_{PQ}, viz.,

$$\begin{aligned} m_a &\simeq 10^{-5}\text{eV}(10^{12}\text{GeV}/f_{PQ}) \quad , \\ \tau(a\rightarrow 2\gamma) &= 10^{41}\text{yrs.}(f_{PQ}/10^{12}\text{GeV})^5 \quad , \\ g_{aee} &\approx m_e/f_{PQ} \quad , \end{aligned} \tag{1.3}$$

where g_{aee} is the coupling of the axion to the electron. Thus, f_{PQ} is required to be $>10^8$GeV from helium burning constraints in various stars. It can be shown that if the energy density is to be of order unity, $\Omega_a \simeq 1$, then we require a PQ breaking scale of $\simeq 10^{12}$GeV, corresponding to an axion mass of 10^{-5}eV. Thus, for the allowed values of f_{PQ} we have

$$10^{-5}\text{eV} \leq m_a \leq 10^{-1}\text{eV} \quad . \tag{1.4}$$

What then are the implications of these WIMPs for the formation of structure in the universe?

It is well known that density perturbations in a self-gravitating fluid, in which the mean free path of the fluid constituents is finite, will undergo Landau damping (cf. Bond and Szalay 1983). For instance, the damping scale, λ_{FS}, for a massive neutrino is of the order of 40Mpc(m/30eV) whereas for an axion it is $<10^{-5}$Mpc. Physically, the damping scale λ_{FS} is the comoving distance that a WIMP could have travelled since the big bang. The scale $\leq$1Mpc corresponds to galactic scale. The relationship to the galactic scale neatly separates the WIMPs into three categories:

(i)	Cold	$\lambda_{FS} \ll 1$Mpc e.g. axions	galactic size perturbations survive free-streaming
(ii)	Warm	$\lambda_{FS} \simeq 1$Mpc	
(iii)	Hot	$\lambda_{FS} \geq 1$Mpc e.g. massive ν's	only perturbations on scales much larger than galactic scales survive free-streaming.

We can see this more clearly in Figure 1.2, where we see the power spectra at late times in a universe now dominated by WIMPs. The quantity $k^3|\delta_k|^2$ is the local power in plane wave perturbations of scale $\lambda=2\pi/k$, and δ_k is the amplitude of the relative density fluctuations in some particle or radiation field. Objects of this size will condense out of the general expansion when $k^3|\delta_k|^2 \approx 1$. Until non-linear effects are important the spectra shown evolve by increasing their amplitude while maintaining their shape.

We see that the characteristic scale for hot dark matter, such as massive neutrinos, is

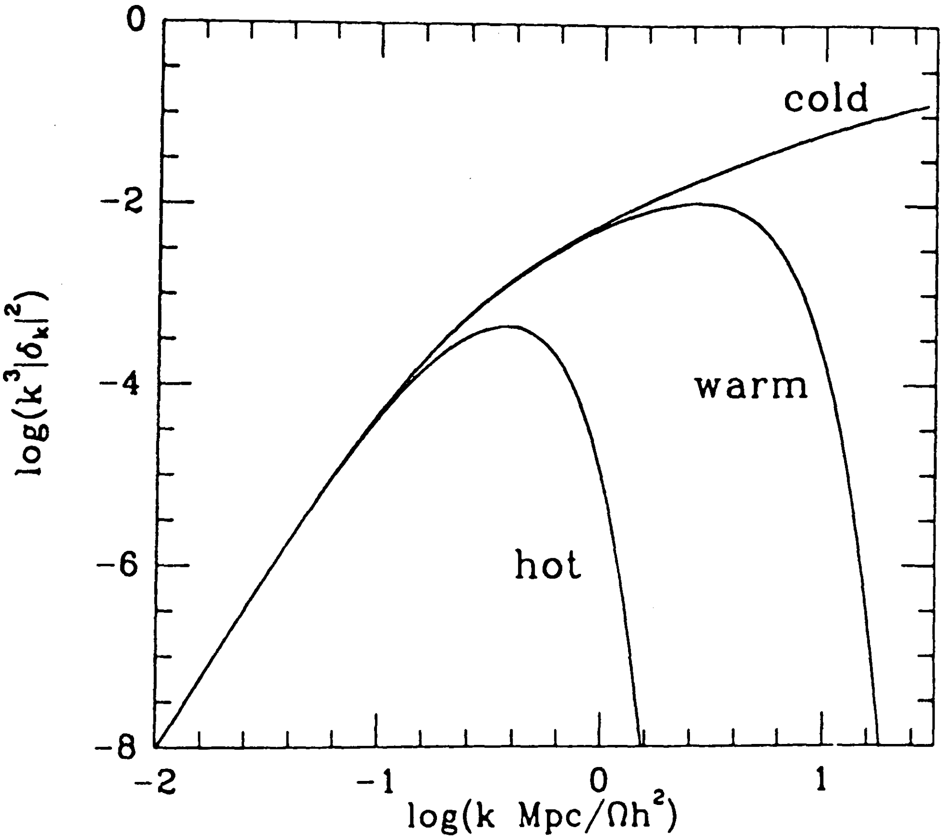


Figure 1.2 Linear power spectra as a function of spatial frequency at late times in a universe dominated by collisionless particles. The three cases are differentiated by the random velocities of the of the particles involved, (cf. White 1987).

$$\lambda \simeq \frac{17}{\Omega h^2}\,\mathrm{Mpc} \simeq 17(100\mathrm{eV}/m_X)\,\mathrm{Mpc} \qquad , \tag{1.5}$$

which is of the order of supercluster size (m_X is the mass of the particle). Thus, structure in a neutrino-dominated universe will grow according to a variant of Zeldovich's "pancake" scenario. On the other hand, an axion-dominated model will cluster hierarchically in the manner discussed by Peebles (1965).

Cosmologies dominated by cold dark matter produce mass distributions which fit the observed galaxy distribution, (i) if Ω=0·1–0·2 and galaxies trace the mass distribution or (ii) Ω=1, $H_0 \leq 50\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ and galaxies form preferentially in high density regions (cf. biassed galaxy formation models, e.g. Dekel and Silk 1986). These cold dark matter model catalogues differ from the real data in that their clusters are somewhat tighter and the associated velocities somewhat higher. (Cold dark matter models can, therefore, reproduce the observed galaxy-galaxy correlation function of Peebles (1980) but not the cluster-cluster correlation). If Ω is indeed unity galaxies cannot trace the mass. Rather they must be over-represented by a factor of about five in the dense regions from which dynamical mass estimates are obtained, (Kaiser 1985).

The major opponent to the dark matter models for the large-scale structure of the universe is the theory of galaxy-formation based on cosmic strings. The cosmic string model does not preclude the existence of dark matter but the mechanism which generates the structure is somewhat different.

In the spontaneously broken gauge theories of elementary particle physics there are, in addition to the fundamental particles of the theory, topological entities, which form as defects in the process of breaking the symmetry. (These objects correspond to classical configurations of the gauge and Higgs fields). A class of these Grand Unified Theories (GUTs) leads to the prediction of topological entities which are line singularities and are referred to as cosmic strings (Vilenkin 1985).

GUTs all begin with the assumption that at the very high energies of the first moments after the big bang, there was no distinction between three of the four fundamental forces of nature (electromagnetism, weak interactions and strong interactions). Soon after the big bang, the symmetry broke and energy settled into fundamental particles, such as quarks and leptons. However, it was postulated that when this occurred, at about 10^{-35}s, frozen bits of unified field got trapped in long 'cosmic strings' (Kibble 1976). These defects contained remnants of the high energy that existed just after the big bang. The existence of cosmic strings is highly speculative. Nevertheless, many unified theories do predict that the universe would fill up with a network of such strings, as defects at the time of the symmetry breaking, which is inherent to these theories. These cosmic strings would be very heavy (typically 10^4Kgcm^{-1}) and very thin ($\approx 10^{-7}$cm) and they would have a very strong gravitational field. (It can be shown that the production of cosmic strings in the very early universe leads to isothermal perturbations in the matter distribution of a definite spectrum and amplitude (Vilenkin 1985), which allows for a theory of structure formation in the late universe).

Cosmic strings are found to occur either in the form of closed loops or as infinitely long strings (Turok 1987). Most ($\approx$80%) of the strings are actually in long 'infinite' strings as large as the universe horizon size. The remainder are in the form of a scale invariant distribution of closed loops. The infinite strings that form are not straight but meander about in the form of Brownian random walks, and the whole collection of strings forms a network that permeates all of space. The mean velocity of a piece of string is of the order of $10^{-1}c$ (Albrecht and Turok 1985). Thus the bits of string frequently intersect.

The evolution of a network of cosmic strings depends crucially on what happens when two strings intersect. For instance, if cosmic strings were to pass right through each other, then the physical length in string would expand as fast as the scale factor of the universe, a(t), and hence the energy density in strings would only decay as $a^{-2}(t)$, compared to the energy density in radiation which falls off as $a^{-4}(t)$. Thus the energy in strings would rapidly become the predominant form of matter-energy in the universe. A universe dominated by cosmic strings would look very different from the one that we observe today. Cosmic strings would be plainly visible all around us and the additional energy of the cosmic strings would cause the universe to expand much faster than the observed rate, e.g. in a radiation-dominated FRW period, the energy density in the strings would cause the universe to expand as $\simeq t$. If on the other hand, strings, as they cross, could also break and reconnect the other way, long strings would form loops (Figure 1.3) and this would avoid the scenario of a universe dominated by strings.

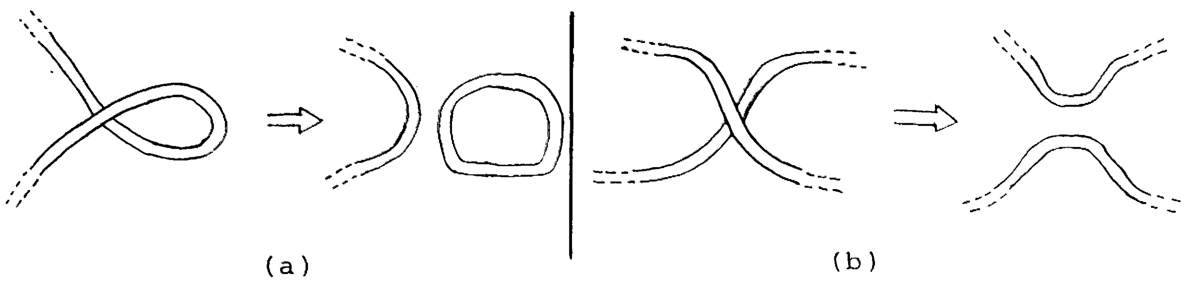


Figure 1.3 Intercommutation of cosmic strings: (a) a single self-intersecting string, (b) two infinite strings intersect.

As a string moves around, it loses energy by radiating gravitational waves. This effect will eventually cause a loop (unlike an infinite string) steadily to decrease in size until nothing remains but radiation. It is this conversion of the energy of the string to radiation which prevents strings from dominating. Shellard (1987) discovered that provided their relative speed is less than 0.9c, two intersecting cosmic strings always break and reconnect to produce smaller loops.

The expansion of the universe, characterised by the Hubble length, strongly influences the evolution of strings. At any time, the infinite strings have 'wiggles' on them that are about the Hubble length in size. These 'wiggles' cause the infinite strings to cross themselves and produce loops, also about a Hubble length in size. As the Hubble length increases, the loops which form are correspondingly larger. Once fromed, the tension in the loops cause them to oscillate. Oscillating mass gives rise to gravitational radiation and so the loops decay by radiating gravitational waves. It can be shown (Brandenberger 1987) that the loops decay completely into radiation after about 10^6 oscillations. Thus, at any time there is a 'debris' of loops left behind by the network, ranging in size from the Hubble length down to zero. (In fact, the network seems to evolve in a self-similar manner with the Hubble length characterising the scale of the network. The existence of this self-similar evolution greatly reduces the complexity of the calculations involved).

Zel'dovich (1980) and Vilenkin (1981) suggested that strings could produce density fluctuations sufficient to explain the galaxy formation in the universe. The gravitational fields of the loops accrete matter leading to the build up of the structures which exist in the universe. A loop accretes a mass proportional to its own mass, so smaller loops accrete a smaller amount of matter. Large loops not only accrete more matter but also the smaller loops around them. Thus, in this scenario, smaller loops formed galaxies and larger loops formed clusters of galaxies. The evolution of the network determines the number of loops of different sizes. The size of the loop also determines its mass and therefore how much matter it will accrete. The mass of a loop is its length times the mass per unit length, μ, of the cosmic string, a quantity that is not uniquely predicted by the underlying field theory. The mass per unit length, μ, depends on the value of the (string-generating) scalar field for which the potential energy is at a minimum and so should be the same for all loops.

We can determine μ by counting the number of galaxies and then try to predict from the theory which size of loop appears in the same quantity. We can then choose the value of μ which gives loops of the right size to accrete a galaxy. This procedure can be repeated for clusters of galaxies. Remarkably these two independent determinations of μ give the same value, Turok and Brandenberger (1985). The value obtained also lies within the range most preferred in the underlying field theory.

A model based on cosmic strings should predict more than just the total number of galaxies or clusters. The distribution of these objects should also be a reflection of the distribution of loops of cosmic string in space. Clusters offer a clearer test because they are too far apart for gravity to have moved them around very much since their formation. A simple way to measure the degree of clustering of the distribution of objects is to use the two-point correlation function (Peebles 1980). For the observed clusters of galaxies, this is found to be identical with that of the corresponding calculations for loops of cosmic strings. Cosmic strings are, therefore, more likely to explain the existence of voids, filaments and sheets in the universe. However, they do have a problem when it comes to predicting the observed matter distribution on scales of order 30-50h^{-1}kpc, i.e. galactic scales. We conclude then by stating that cosmic strings do offer an intriguing alternative to the scenario of a WIMP-dominated universe as a viable model for structure formation. The dark matter models are able to produce a good agreement with the observed galaxy-galaxy correlation function but fail on the cluster-cluster correlations. Thus, we should not rule out the possibility that some combination of these models may be more profitable than the individual models themselves.

Cosmic strings were found to originate in the very early universe. Therefore, it would seem that in order to fully comprehend the evolution of the universe we must consider the contribution made by phenomena occurring in the early stages of this evolution.

1.4 The Early Universe

The models of Friedmann (1922) and Lemaitre (1927) and the discovery of extragalactic recession by Hubble (1929) established securely the concept of an expanding universe. A simple extrapolation back in time leads directly to an initial big-bang state of high density. The idea of a hot and dense early universe was put on a firm foundation by the discovery of the 3°K background radiation by Penzias and Wilson (1965) and its identification by Dicke et al. (1965). However, prior to this discovery the idea had played an active and a prominent role in the work of Gamow and collaborators, with excellent reviews in Gamow (1953) and Alpher et al. (1953). Advocates of cold big bang theories, in which the microwave background does not have a primordial origin, face the problems of providing a mechanism which generates the observed thermal background and producing the observed cosmic helium abundance. With the assumption of a hot big bang, the early universe becomes an extremely fascinating and physically intricate subject for study.

The thermal history of the standard big bang from a temperature of 10^{12}K is illustrated in Figure 1.4. The standard model of the early universe studied by Gamow (1948) is free of any pronounced anisotropy and large inhomogeneities. The extreme early universe ($T>10^{12}$K) contains all manner of particles and antiparticles. By the time the temperature has dropped to 10^{12}K the hordes of hadrons existing in an earlier era have almost completely disappeared, leaving behind a few surviving nucleons and a rapidly diminishing population of pions. The universe then consists mainly of leptons, antileptons and photons and enters the *lepton era*. As far as density is concerned, the universe is lepton-dominated.

As the declining temperature approaches $\simeq 8\text{x}10^{11}$K (cooled by the expansion of the universe) muon pairs begin to annihilate. Shortly thereafter both the muon and electron neutrinos decouple when their rapidly increasing interaction time exceeds the expansion time. The surviving electrons and photons remain in a state of thermal equilibrium until the temperature approaches $\simeq 4\text{x}10^{9}$K. Electron pairs then annihilate significantly faster than they are created and we reach the *radiation era*.

The radiation era lasts until the recombination of hydrogen occurs at 3000°K. During most of this long period the universe is radiation-dominated and contains only a trace of matter. Toward the

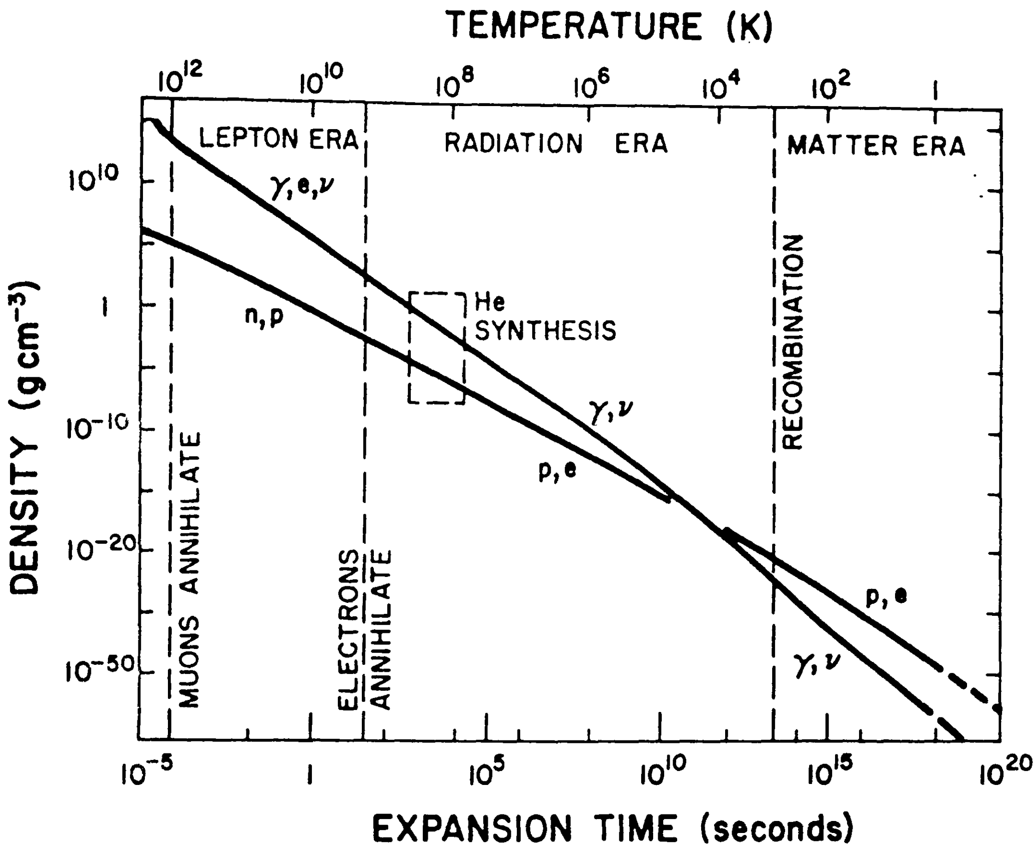


Figure 1.4 Thermal history of a standard big bang universe to a temperature of 10^{12}K. Note that recombination does not coincide with the universe becoming matter-dominated.

tail-end of the radiation era matter is increasingly important, and either slightly before or slightly after the recombination epoch the universe becomes matter-dominated. The radiation era, however, does not necessarily terminate at the instant the universe is matter-dominated. The important fact is that the contents of the universe remain radiation-pressure dominated until the recombination epoch. Helium is synthesised during the early stages of the radiation era and later when the Hubble mass has increased, the various precursory inhomogeneities of galaxy formation take effect. A more detailed discussion of the physical processes occurring in the different regimes of the early universe is given by Harrison (1973).

Thus, the standard model, whereby the scale factor of the universe follows the Friedmann equations, allows us to extrapolate backwards from the present extremely isotropic universe, at least on large scales, to $t \approx 10^{-2}$ seconds. However, to completely describe the evolution of the universe, we need to know what happens during that first 10^{-2}s. The remarkable developments in elementary particle physics, in the search for a unified theory of the forces of nature, have allowed cosmologists to 'probe' the very early stages of the universe.

The gauge theories of the particle physicists have been very successful, so far, in describing and predicting the behaviour of fundamental particle interactions, where the forces controlling the particles are described by the spontaneous breaking of the symmetries imposed by the theories. The point of spontaneous symmetry breaking (SSB) is that although the laws of physics may be intrinsically symmetrical, that symmetry is not manifest below a certain temperature or energy level, because the lowest energy state of the system (the vacuum) is a particular solution of the equations that does not possess their symmetry. In spontaneously broken gauge theories the properties of the vacuum state are of vital importance and are described by the vacuum expectation value of a scalar field - the Higgs field (Abers and Lee 1973). The appearance of a non-zero vacuum expectation value signals SSB.

The mathematical elegance and experimental vindication of the SU(3), strong force symmetric gauge, and SU(2)xU(1), electro-weak gauge theory, models has led to their incorporation within proposals for the grand unification of the strong and the electro-weak interactions. In such models a spontaneous breakdown of complete symmetry between the strength and properties of these three interactions occurs when the temperature falls to 10^{28}K, an energy of about 10^{15}GeV. Such energies are never likely to be attained by terrestrial particle accelerators. However, we do have a 'theoretical laboratory' in which to test our ideas, i.e. the early universe. According to the hot big bang model, temperatures corresponding to average particle energies as large as 10^{19}GeV should have been reached in the early universe.

Guth (1981) proposed a new picture for the early stages of the hot big bang model which provided a 'natural' explanation for a collection of cosmological problems. The picture was dubbed the 'inflationary universe '. In particle physics theories which undergo SSB, the Lorentz invariant energy density associated with the vacuum changes during a phase transition and creates an effective cosmological constant, $\rho_0 \simeq O(T_C{}^4)$, where T_C is the critical temperature of the phase transition. As the universe cools below T_C, bubbles of the low-temperature phase (asymmetric vacuum) nucleate and grow and eventually the entire universe is in the asymmetric phase. However, Guth pointed out that if the nucleation rate was sufficiently small, then the universe would remain in the high-temperature phase (symmetric vacuum) for a non-negligible period: the symmetric vacuum is metastable. During this interval the initial cosmological term, ρ_0, would soon begin to dominate the expansion dynamics and the universe would expand exponentially and 'supercool', erasing its previous history. (Exponential expansion was first discussed by de-Sitter in 1917). When the transition to the asymmetric vacuum state does occur, an enormous latent heat (the energy difference between the two vacuum states) is released, reheating the universe so its subsequent evolution is that of the standard big bang model. As a result of the exponential expansion phase in its early evolution, the portion of the universe that is observable today should be extremely uniform and expanding at a critical rate (that is, the ratio of the potential to the kinetic energy in the universe, Ω, should be equal to unity).

Without this de-Sitter phase, in which the size of the universe is greatly inflated over what one would expect early on, the present uniformity of the universe and the proximity of its expansion rate to the critical value would remain a mystery. Our only explanation, other than some inflationary model, would be to appeal to very special initial conditions, i.e. the value of the Hubble constant would have to be fine-tuned to an accuracy of one part in 10^{55}. Guth's inflationary universe dispensed with the need for special initial conditions. Furthermore, it solved the somewhat embarrassing problem of the over-production of magnetic monopoles during the SSB phase transition of the Grand Unified Theories (GUTs). In the earlier, more conventional, models of the GUT phase transition so many monopoles were predicted to arise that they would contribute a density more than 10^{12} times larger than the maximum allowed by the observed deceleration of the universe, (Preskill 1979, Zel'dovich and Khlopov 1979). In the inflationary model, the period of accelerating expansion dilutes the monopole density to a small and observationally permissible level. Unfortunately for this model, once the universe reaches this symmetric phase, it remains trapped there.

Linde (1982), Albrecht and Steinhardt (1982) and Hawking and Moss (1982) analysed a different class of GUTs (those which undergo SSB of a characteristic type first studied by Coleman and Weinberg 1973) and discovered that in these models the advantageous features of Guth's inflationary model could be retained whilst the difficulty of escaping from the symmetric vacuum of de-Sitter space could be overcome. In their second generation model - the so-called 'new inflationary model' - the universe never gets trapped in the symmetric vacuum state but instead simply takes a very long time to evolve from the symmetric to the asymmetric vacuum state, and while the universe is evolving from the symmetric to the asymmetric state it expands exponentially due to the large energy density of the symmetric vacuum. In this new model the evolution to the true vacuum takes long enough for sufficient exponential expansion to occur to explain the uniformity, expansion rate and monopole-free composition of the present-day universe.

One problem with the new-inflationary model is that the matter inhomogeneities, which are spontaneously generated by quantum fluctuations during the de-Sitter phase, are found to be far too large to lead to galaxies, rather everything would evolve to form superdense objects or black holes. Also, the inflationary models predict that today the density parameter Ω should be equal to unity to a very high degree of precision. However, the best astronomical determinations of Ω all consistently suggest a much smaller value, less than 0·2. A possible way out of this conflict would be if the universe was dominated ($\Omega \approx 1$) by non-baryonic dark matter, which coalesces on very large scales so that the observations to date would not have been sensitive to the presence of this matter.

The fact that inflationary models predict the observed isotropy of the universe does not preclude the existence of significant anisotropies and inhomogeneities before the onset of inflation. Indeed several authors (Diósi et al. 1984, Waga et al. 1986) have raised the possibility that bulk viscosity in the early universe could be the driving force of an accelerated expansion akin to inflation. These authors have suggested that bulk viscosity arising around the time of a GUT phase transition could lead to a negative pressure thereby driving an inflationary expansion. In order for structure such as galaxies and clusters of galaxies to form, the universe must develop density inhomogeneities at some time during its evolution. It is unrealistic to assume that the universe could contain such inhomogeneities without being at least somewhat anisotropic, since density fluctuations tend to generate shear motions via tidal stresses (see Liang 1974 and Barrow 1977). Quite apart from the anisotropy which is generated in this way it is feasible that the universe started off endowed with a lot of primordial anisotropy. Thus, if we are going to feed inhomogeneities into the initial conditions of the universe (to explain the galaxies), we are equally entitled to feed shear into the initial conditions and, in this situation, the anisotropy must dominate the dynamics of the universe at early enough times.

However, the universe cannot be shear-dominated indefinitely. The density associated with induced shear can never be more than comparable to the matter-radiation density and even primordial shear can only dominate the density of the universe until some time t_s. This is because shear energy decreases with redshift like z^6, whereas the matter-radiation density decreases more slowly (like z^4 before matter-radiation equilibrium and z^3 thereafter). Thus, even if the universe starts off shear-dominated, it will not remain so for ever. Indeed, calculations of the effect of shear on cosmological nucleosynthesis (Barrow 1976) indicate that t_s cannot exceed $\simeq$1s. We might anyway expect most primordial anisotropy to have been dissipated before t=1s by collisional and collisionless dissipative processes (Misner 1967).

Relaxing the requirement of isotropy, i.e. permitting the cosmological flow to rotate and shear as it expands, allows more freedom in the choice of solutions. Various cosmological models of this kind have been formulated in which the rates of expansion in different directions are unequal (Taub 1951, Ellis and MacCallum 1969). A more physical interpretation of these models is that very-long-wavelength gravitational standing waves are present throughout the evolving matter distribution. Such a gravitational wave field determines preferred directions and orientations in space. If this picture is simultaneously true everywhere, the models are said to be spatially homogeneous (cf. Bianchi models; Bianchi 1897, Ryan and Shepley 1975).

The simplest anisotropic model is the Kasner model (Kasner 1921) in which the vorticity and acceleration of the flow lines are absent and in which shear and expansion are completely specified once the expansion rate in one direction is known. Such Kasner solutions can be easily extended to higher dimensional cosmologies and these will be discussed in the next section. Suffice it to say that within the very generous confines of 'conventional' cosmology the early universe has a great deal to offer to the understanding the universe in its entirety. (A collection of papers on the subject of the physics of the very early universe, is presented in Gibbons et al. 1983).

1.5 Exotic Cosmologies

In the preceding sections we have discussed the various 'conventional' methods of devising cosmological models in an attempt to explain the universe as it exists, within the framework of the general theory of relativity. There have been models which attempt to 'usurp' general relativity by providing an alternative theory (e.g. Brans and Dicke 1961, Smalley 1974). None of these alternatives, however, have been able to match the phenomenal success of relativity theory in satisfying most of the available observational tests. Less radical, but more enlightening, have been the extensions of general relativity to higher dimensions (e.g. Kaluza-Klein models, supersymmetry theories, superstring theories and extended Kasner models). All of these theories exist in an attempt to unify the fundamental forces of nature.

In general relativity, the force of gravity appears as a result of distortions in four-dimensional space. Kaluza (1921) was interested in what would happen if the equivalent equations were expressed in five dimensions. (Note that in 1921 there was no physical justification for this). When Kaluza wrote down the five-dimensional equivalent of the equations of general relativity, by the addition of an extra spatial dimension, he found that they automatically divided into two sets of equations in four dimensions. One set corresponded to the familiar gravitational influence while the other set exactly corresponded to the equations of electromagnetism (Maxwell's equations). Five-dimensional "reativity" seemed to unify the two forces known in 1921.

Kaluza proved his results only for the case where the fields were weak, (i.e. $g_{\mu\nu}=\eta_{\mu\nu}+h_{\mu\nu}$, $|h_{\mu\nu}|\ll 1$, $\eta_{55}=1$), and the velocity was small ($v/c \ll 1$). However, Klein (1926a) showed that these two constraints were irrelevant, unification should not depend on the fields being weak and the velocities small. Klein employed the results obtained by the rapid developments of quantum physics in the early 1920's, particularly the development of Schrődinger's equation (Schrődinger 1926). Klein took Kaluza's five-dimensional theory and translated it into quantum terms by writing down a version of Schrődinger's equation with five variables (each one effectively corresponding to a dimension) instead of four. He showed that this five-dimensional Schrődinger's equation had solutions which corresponded, respectively, to gravitational and electromagnetic waves in four-dimensional space.

In these early studies no real attempt was made to justify the use of five dimensions nor to explain where the extra dimension was hidden. However, Klein (1926b) suggested that the extra fifth dimension could be 'rolled up' or 'compactified' so that it was undetectable in the everyday world. (The usual analogy is a hosepipe, which when viewed from a long way away, looks like a one-dimensional line. However, when you look more closely, it turns out to be a two-dimensional object. Each point on the 'line' is a circle around the circumference of the tube). Klein suggested that every point in three-dimensional space might really be a tiny circle looping around a fourth spatial dimension. Calculations suggest that each loop of "string" would be about 10^{-30} to 10^{-33} centimetres across, i.e. of the order of the Planck length. This compactification argument has become standard in the study of higher dimensional cosmologies today, known collectively as Kaluza-Klein models.

The original Kaluza-Klein theory helped unify the two known fundamental forces at that time, gravity and electromagnetism. However, in the decades that followed, experiments in particle physics led to the discovery of two more fundamental forces of nature; the strong interaction and the weak interaction. Witten (1981) demonstrated that the Kaluza-Klein approach could be extended to unify the strong, weak and electromagnetic interactions, if we employed a minimum of eleven dimensions in all, ten spatial and one temporal. There are two natural ways for the resulting eleven-dimensional space to compactify, either four dimensions curl up into insignificance, leaving a seven-dimensional world, or seven dimensions compactify, leaving four dimensions behind. The 'odd' force out in these considerations is gravity. As yet there has been little success in the search for a consistent quantum theory of gravity. The idea of a supergravity has been developed, however, which like Einstein's theory is a geometrical theory of gravity. Supergravity goes beyond general relativity and attempts to unify gravity with the other forces of nature. Theories of supergravity can be made to work in different numbers of dimensions, but only up to a maximum of eleven.

Today, the most favoured variation of the Kaluza-Klein approach is ten-dimensional superstring theory, which involves fundamental particles that are one-dimensional strings, not mathematical points, and also incorporates a version of supergravity. The justification for working in ten or eleven dimensions comes purely from theoretical particle physics and is not founded in any observations. Present accelerators have probed matter at distances as small as 10^{-16}cm without finding any evidence of extra dimensions, which is not too surprising as the extra dimensions are expected to have a size characteristic of the Planck length ($\approx 10^{-33}$cm).

Supersymmetry (Duff et al. 1986) is the symmetry which interchanges fermions and bosons. In a supersymmetric theory there is a bosonic counterpart for every fermion and vice versa. There is no evidence for such a symmetry in the world around us, e.g. there is no massless fermionic partner for the photon, or scalar partner for the electron. The motivation for supersymmetry is that mathematically it is very elegant and it is the ultimate symmetry we have available to impose. When supersymmetry is made a gauge symmetry (this is called supergravity) it leads to a generally covariant theory, i.e. it automatically incorporates general relativity into the theory. Thus, it offers the hope of unifying gravity with the other forces. Supersymmetry also offers the hope of clearing up the discrepancy of the weak and GUT symmetry breaking scale encountered in all GUTs. This discrepancy is some twelve or so orders of magnitude in a typical GUT. Although we are free to set these scales to very different energies, quantum corrections restrict this and tend to raise the weak scale up to the GUT scale (or the highest scale in the theory). Supersymmetry can be used to stabilise this discrepancy once it is initially set.

Since there is no evidence that we exist in a supersymmetric universe, supersymmetry must also be a broken symmetry. In order to stabilise the weak scale the supersymmetry breaking scale must effectively occur at the weak scale. This means that the supersymmetric partners, or spartners, of all known particles must have masses of the order of the weak scale, where "of the order" means between a few GeV and a TeV. The scalar partners of the quarks are called squarks; the scalar partners of the leptons are called sleptons; the fermionic partners of the photon, gluon, W, Z and graviton are the photino, gluino, Wino, Zino and gravitino respectively. The fermionic partners of the Higgs particles are referred to as Higgsinos.

Because of an additional symmetry that most supersymmetry/supergravity models have (called R-parity) the lightest spartner is stable, and because the effective supersymmetry breaking scale is of the order of the weak scale, the interactions of spartners with ordinary particles are about as strong as the usual weak interactions. This makes the lightest spartner an ideal candidate WIMP (see dark matter discussion in §1.3).

Almost all supersymmetry/supergravity models are supersymmetric GUTs. The unification scale in these theories is higher than in normal GUTs, more like 10^{16}GeV (compared to 10^{14}GeV) and these theories are supposed to describe physics up to 10^{19}GeV. Therefore, supersymmetry/supergravity models also predict all of the additional particles that GUTs do - magnetic monopoles, massive neutrinos, axions and even cosmic strings in some cases.

Superstring theories (Schwarz 1985) combine the ideas of supersymmetry, gauge symmetry, extra dimensions and one new one, strings (not to be confused with cosmic strings). The basic idea is that the fundamental particles are not point-like, but rather are string-like, one-dimensional entities and such theories can only be consistently formulated in ten dimensions. Superstring theories unify all the forces of nature (including gravity) in a finite quantum theory and are almost unique (only five string theories are known to exist). In principle, starting from the superstring (which describes physics at or above the Planck scale) we can calculate everything - the masses of all the fermions, the GUT, etc. When viewed at large distances the loops look like point particles. The so-called point-like limit of a superstring theory is supposed to be a supersymmetry/supergravity GUT. All the WIMP candidates predicted by supersymmetric GUTs are also predicted by superstring theories.

Thus, GUTs attempt to describe physics up to around 10^{14}GeV, supersymmetric GUTs up to 10^{19}GeV and superstring theories at energies above 10^{19}GeV.

All of the higher-dimensional cosmologies, discussed above, rely on the compactification of the additional dimensions. One simple way of understanding how these extra dimensions can be incorporated into general relativistic cosmology is to consider the Kasner solutions discussed briefly in the last section. The Kasner solutions (Kasner 1921) were the prototypes for cosmological models with great asymmetry in a few degrees of freedom. The four-dimensional Kasner metric is given by

$$ds^2 = dt^2 - t^{2m}dx_1{}^2 - t^{2n}dx_2{}^2 - t^{2p}dx_3{}^2 \quad , \qquad (1.6)$$

where m, n and p are constants satisfying the constraints,

$$m + n + p = m^2 + n^2 + p^2 = 1 \quad . \qquad (1.7)$$

Thus, each t=constant hypersurface of this model is a flat three-dimensional space. This model represents an expanding universe, since the volume element is constantly increasing. However, it is an *anisotropically* expanding universe. The separation between standard (constant x_1, x_2, x_3) observers is $t^m \Delta x_1$, if only their x_1-coordinates differ. Thus distances parallel to the x_1-axis expand at one rate, $R_1 \propto t^m$, while those along the x_2-axis can expand at a different rate, $R_2 \propto t^n$. Most remarkable perhaps is the fact that along one of the axes, distances contract rather than expand. This contraction shows up mathematically in the fact that equations (1.7) require one of m, n or p, say p, to be non-positive:

$$-1/3 \leq p \leq 0 \qquad . \qquad (1.8)$$

Thus, we can immediately see the extension to higher dimensions, e.g. five (four spatial and one temporal). The metric would then take the form

$$ds^2 \;=\; dt^2 - t^{2m}dx_1{}^2 - t^{2n}dx_2{}^2 - t^{2p}dx_3{}^2 - t^{2q}dx_4{}^2 \qquad , \qquad (1.9)$$

where we have introduced the new 'scale factor', t^q, for the additional dimension. The constraints on the expansion rates are then

$$m + n + p + q \;=\; m^2 + n^2 + p^2 + q^2 \;=\; 1 \qquad , \qquad (1.10)$$

and again we see that at least one of the expansion rates must be non-positive, q say. Thus, compactification of the extra dimension follows quite naturally in such a model. As $t \to \infty$, $t^q \to 0$ and the dimension x_4 is 'lost' while the remaining dimensions grow large. This analysis can be continued for as many extra dimensions as we like, being only restricted by the fact that observationally all but three of the spatial dimensions must become vanishingly small.

Finally, we would like to consider another five-dimensional (5D) cosmology, which is interesting in that it draws on actual observations of the universe. One observational feature of the universe is the mass. If we allow the somewhat unusual property that the rest mass of a particle changes with time, then it is found that at least some of the observations not accounted for by Einstein's four-dimensional (4D) theory of gravitation, may be explained (Wesson 1984). The approach of this version of cosmology starts from the equivalent of adding an extra dimension to everyday space by multiplying time by the speed of light to obtain a measure of distance. The constant of gravity, G, can also be used to convert masses into distances. The length ct is a coordinate in 4D special and general relativity. But the parameter Gm/c^2 (where m is the rest mass of a particle) also has units of length and so can be used as a coordinate, to give a 5D version of general relativity.

Relativity in four dimensions implies that the strength of gravity is constant, a fact which has been verified by many experiments. However, when we examine carefully the calculations on which this conclusion is based, a curious fact emerges. Because of the nature of the underlying physical laws, if the strength of gravity is proportional to the time that has elapsed since the birth of the universe in the big bang, the properties of astronomical systems are almost exactly the same as they are if Gm/c^2 is a constant. This parameter is therefore a natural measure of the 'strength' of the gravitational force associated with a particle of mass m. Thus, in nature it seems to make no difference whether the rate of change of this parameter is steady and finite or zero. Thus, we can allow for a steady rate of change of the rest mass, m, keeping G as strictly constant.

A similar argument, for a varying Gm/c^2, was given by Dirac

(1938) in what he called the Large Numbers Hypothesis, where dimensionless ratios of the physical constants of nature are found to be typically of the order of 10^{40}. For example, if we compare the relative strength of the electrical and the gravitational forces between the electron and the proton we find that a large dimensionless number is obtained, given by

$$\frac{e^2}{Gm_p m_e} = 2.3\text{x}10^{39} \quad , \qquad (1.11)$$

where e is the charge of the electron, G is the gravitational constant and m_p, m_e are the masses of the proton and electron, respectively. Similarly, if we compare the length scale associated with the universe, c/H_o, and the length associated with the electron, $e^2/m_e c^2$, we obtain the ratio

$$\frac{m_e c^3}{e^2 H_o} = 3.7\text{x}10^{40} h_o^{-1} \quad . \qquad (1.12)$$

Dirac pointed out that (1.12) contained the Hubble constant, H_o, and therefore the magnitude computed in this formula varies with the epoch in the standard Friedmann model. If so, the near equality of (1.11) and (1.12) has to be a coincidence of the present epoch in the universe, unless the constant in (1.11) also varies in such a way as to maintain the state of near equality with (1.12) at all epochs. This would imply that at least one of the so-called constants involved in (1.11), e, m_p, m_e, and G, must vary with epoch. Because G has macroscopic significance, whereas the other constants are atomic quantities, Dirac postulated that the gravitational constant must vary with time, in such a manner as to keep the ratios (1.11) and (1.12) of roughly the same magnitude.

There is no observation that we can devise which could be capable of detecting this particular kind of variation. In effect, this is a 5D equivalent to the feature of special relativity that there is no preferred 'frame of reference' so that the laws of physics are the same in all frames moving at constant velocity relative to one another. In this 5D theory, an equivalent rate of change refers to "velocity" along the fifth-axis - the one described in terms of mass.

This description of the universe agrees with all observations to date which is not too surprising since, in the limiting case where the rest masses of particles vary infinitely slowly, the equations become the familiar equations of 4D relativity. Provided the rate at which mass is changing today is small, there will be no reason to expect any observations to conflict with the predictions of the theory of relativity. The age of the universe is 10^{10} years and according to the 5D theory, the amount of mass in the universe today has built up at a steady rate over all that time. So the rate at which the mass of a proton, say, is increasing today is no more than one part in 10^{10} each year, far below detectable limits.

There is a set of solutions to the 5D equations that allow the rest masses of particles to increase from zero at time zero, Wesson (1986a). According to this theory, in the beginning there was no mass. The same rate of growth of mass which is one part in 10^{10} per year today, means that the mass of a proton doubled during the second "year" of the life of the universe. It also means that, unlike the usual theory, the universe did not start in a big bang. If this were true it would completely change our ideas on the origin of the universe, the nature of the cosmic microwave background and the origin and evolution of galaxies.

2. GLOBAL SYMMETRIES IN COSMOLOGY

2.1 Introduction

Differential geometry has a major role to play in the study of modern theoretical physics. In the nineteenth century, physicists were content to 'live' in the three dimensional world of Euclidean geometry, happy in the realisation that the physical laws of nature could be expressed as differential equations. Euclidean geometry allowed them to develop powerful analytic techniques with which to solve these differential equations and any further applications of geometry were neglected.

However, two developments in this century significantly altered the balance between geometry and physical analysis in the outlook of the modern physicist. The first was the development of the theory of relativity, according to which the Euclidean three-space is only an approximation to the correct description of the physical world. The second was the realisation, principally by the mathematician Cartan, that the study of geometry leads naturally to the development of certain analytic tools (e.g. the Lie derivative and exterior calculus) and certain concepts (e.g. the manifold and the identification of vectors with derivatives) that have an important function in the applications of physical analysis. Because it has developed this intimate connection between geometrical and analytical ideas, modern differential geometry has become increasingly more important in theoretical physics, where it has led to a greater simplicity in the mathematics and a more fundamental understanding of the physics.

The key to differential geometry's importance is that it studies the geometrical properties of continuous spaces, in which most physical problems are embodied, whether it be a physical three-dimensional space, a four-dimensional spacetime or a phase space. The most basic of these geometrical properties go into the definition of the differentiable manifold, which is the mathematically precise substitute for the word 'space'.

A manifold is essentially a space which is locally similar to Euclidean space in that it can be covered by coordinate patches. This structure permits differentiation to be defined, but does not distinguish intrinsically between different coordinate systems. Thus, the only concepts defined by the manifold structure are those which are independent of the choice of a coordinate system, (Hawking and Ellis 1973, Schutz 1980).

The theory of relativity, on which the study of theoretical cosmology is founded, is based on the consideration of the universe as a four-dimensional differentiable manifold, where ordinary three-space is combined with time into one unified coordinate system (Einstein 1905, 1915). [There are cosmological models which deal with higher dimensional manifolds, such as the five-dimensional Kaluza-Klein models dicussed in the last chapter, but these are rather exotic and for the moment we prefer to remain conventional]. We have noted above that an n-dimensional manifold, M, is locally similar to Euclidean space, denoted $\mathbb{R}^n$, so that each point on M can be identified with a set of 'coordinates' in $\mathbb{R}^n$. This allows us to define a coordinate system on the manifold.

We now introduce the important concept of a vector field on the manifold. A vector field is closely tied to the concept of differentiablity. Consider a function f on the manifold. The change in f between the points P and Q depends on the vector **PQ** and the function itself. If P and Q are in the same coordinate patch (Δx^σ the difference in their coordinates),

$$f(Q)-f(P) \equiv \Delta f \approx \Delta x^\sigma(\partial f/\partial x^\sigma) = \text{vectorial derivative.} \tag{2.1}$$

The dependence of Δf on displacement is contained in the linear differential operator

$$\Delta x^\sigma(\partial/\partial x^\sigma) \equiv \Delta x^\sigma \partial_\sigma \quad , \tag{2.2}$$

to be thought of as the vector **PQ.**

Modern differential geometry refines this idea of a vector as follows: (1) Take the limit as $\Delta x^\sigma \to 0$ to define a local concept (*tangent vector*) which preserves the directional properties of **PQ.** (2) Ensure that this concept is independent of coordinates. (3) Define the concept of vector field, consisting of a tangent vector at each point of the manifold.

Once vectors are defined in a coordinate-free manner it is convenient to define the concept of a differential form. Differential forms are especially useful in describing antisymmetric, covariant tensor fields. We define a differential form (of first degree), also called a one-form, as a linear operator on vector fields. That is, if $\tilde{\omega}$ is a one-form and $\mathbf{U}$ a vector, $\tilde{\omega}(\mathbf{U})$ is a function, so that $\tilde{\omega}(\mathbf{U})(P)$ is a real number, where P is a point on the manifold.

If $\{\mathbf{X}_\mu\}$ is a basis, we define a set of one-forms $\{\tilde{\omega}^\mu\}$ by

$$\tilde{\omega}^\mu(\mathbf{X}_\nu) = \delta^\mu{}_\nu = \begin{cases} 1 \ , & \mu=\nu \\ 0 \ , & \mu\neq\nu \end{cases} \quad , \tag{2.3}$$

($\delta^\mu{}_\nu$ is the Kronecker delta). The functions $\tilde{\omega}^\mu(\mathbf{X}_\nu)$ are the constant functions $\delta^\mu{}_\nu$. These $\tilde{\omega}^\mu$ are called the *duals* of $\mathbf{X}_\mu$. As with vectors, $\tilde{\omega}=b_\sigma\tilde{\omega}^\sigma$ is required to be independent of the choice of cooordinates.

If we assume the manifold to be that of a four-dimensional spacetime then any point on the manifold is termed an 'event'. Any spacetime event can be labelled by four coordinates, (x^0,x^1,x^2,x^3), where x^0 is the 'time' assigned to the particular event and x^1,x^2,x^3 are its space values. Having labelled each event in spacetime we can proceed to determine the spacetime interval between any two events.

In general relativity the interval between two events in spacetime can be expressed by the metric,

$$ds^2 = g_{\mu\nu}\Delta x^{\mu}\Delta x^{\nu} \quad , \qquad (2.4)$$

where Greek indices run from 0-3 and we apply the Einstein summation convention. Δx^{μ} is the difference in the μ^{th} coordinate value between the two events and $g_{\mu\nu}$ is the component form of the *metric tensor* and is dependent on the geometry of the spacetime.

The metric tensor, $\underline{\mathbf{g}}$, is extremely important in cosmology as it is used to define the geometric structure of the spacetime. It is defined as a linear function which associates a number (the 'dot product') with two vectors, and can be written as

$$\underline{\mathbf{g}}(\mathbf{V},\mathbf{U}) = \underline{\mathbf{g}}(\mathbf{U},\mathbf{V}) \equiv \mathbf{U}.\mathbf{V} \quad . \qquad (2.5)$$

These components form an n x n symmetric matrix. If it happens that the matrix is the unit matrix, we say the metric tensor is the Euclidean metric and the vector space is called Euclidean space. In this space the transformations between coordinate systems (Cartesian bases) are given by the orthogonal matrices. These matrices form a group, which is called the Euclidean symmetry group. If the metric tensor takes the form used in special relativity, namely,

$$g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1) \quad , \qquad (2.6)$$

the metric tensor is said to be pseudo-Euclidean and the vector space is called the Minkowski spacetime. This manifold is one of the most important manifolds in physics. The transformation matrices between bases in Minkowski spacetime also form a group, called the Lorentz group L(n). (Note that throughout this thesis we shall adopt the timelike convention, i.e. the metric signature is (+−−−)).

The fundamental dependence of the metric tensor on the geometry of the spacetime associated with general relativity, allows the effects of curvature to be included into the general description of a cosmology. Thus derivatives of functions, vectors or tensor fields take on a more complicated form as terms must be added to compensate for such curvature effects (Misner, Thorne and Wheeler 1973). We therefore need a generalisation of the concept of a partial derivative in order to set up field equations for physical quantities on a manifold. We obtain such a generalised derivative, *the covariant derivative*, by introducing some extra structure in the form of an affine connection on the manifold, (cf. Hawking and Ellis 1973). Affine connections allow us to define the concept of parallelism on a manifold, i.e. we can compare vectors at different points on the manifold. An affine connection is a rule for *parallel transport*, for moving a vector along a curve without changing its direction.

It can be shown (cf. Ryan and Shepley 1975) that the affine connection takes the form

$$\begin{aligned}\Gamma^{\mu}{}_{\nu\lambda} \;=\;& \tfrac{1}{2}g^{\mu\sigma}(g_{\sigma\nu,\lambda} + g_{\sigma\lambda,\nu} - g_{\nu\lambda,\sigma}) \\ &+ \tfrac{1}{2}(-C^{\mu}{}_{\nu\lambda} + g_{\nu\sigma}g^{\mu\tau}C^{\mu}{}_{\tau\nu} + g_{\sigma\lambda}g^{\mu\tau}C^{\sigma}{}_{\tau\nu})\end{aligned}, \qquad (2.7)$$

where $(,\tau)$ denotes partial derivative with respect to coordinate x^{τ}, i.e.

$\partial/\partial x^{\tau}$, and where the C's are the "structure coefficients", which, if non-zero, express the non-commutativity of the basis tetrads used. If the coordinate system is chosen to be holonomic (the basis tetrad $\mathbf{e}_{\mu}=\partial/\partial x^{\mu}$ is a coordinated basis), all structure coefficients are zero and the affine connection reduces to the Christoffel form;

$$\Gamma^{\mu}{}_{\nu\lambda} = \frac{1}{2}g^{\mu\sigma}(g_{\sigma\nu,\lambda} + g_{\sigma\lambda,\nu} - g_{\nu\lambda,\sigma}) \quad . \tag{2.8}$$

This will become important later.

Having introduced the affine connection on the manifold, we may define the covariant derivative of a vector field **Y** along vector **x** as,

$$\nabla_{\mathbf{x}}\mathbf{Y} = Y^{\mu}{}_{;\nu}x^{\nu}\mathbf{e}_{\mu} = \left[\frac{\partial Y^{\mu}}{\partial x^{\nu}} + \Gamma^{\mu}{}_{\nu\lambda}Y^{\lambda}\right]x^{\nu}\mathbf{e}_{\mu} \quad , \tag{2.9}$$

where the Γ's are the components of the affine connection and $\mathbf{e}_{\mu}$ is the basis vector $\partial/\partial x^{\mu}$.

The "curvature" of the manifold, defined by the Riemann curvature tensor, measures the non-commutativity of covariant derivatives in spacetime:

$$R^{\lambda}{}_{\mu\nu\tau}x^{\mu} = x^{\lambda}{}_{;\nu\tau} - x^{\lambda}{}_{;\tau\nu} \quad . \tag{2.10}$$

For a spacetime to be flat, i.e. to have zero curvature, the left hand side of equation (2.10) must be zero for all events on the spacetime topology. In other words, in a flat spacetime covariant derivatives defined by different vector fields on the manifold commute. From the definition of the covariant derivative, equation (2.9), we see that given an affine connection, the Riemann tensor takes the form,

$$R^{\lambda}{}_{\mu\nu\tau} = \Gamma^{\lambda}{}_{\mu\tau,\nu} - \Gamma^{\lambda}{}_{\mu\nu,\tau} + \Gamma^{\sigma}{}_{\mu\tau}\Gamma^{\lambda}{}_{\sigma\nu} - \Gamma^{\sigma}{}_{\mu\nu}\Gamma^{\lambda}{}_{\sigma\tau} \quad , \tag{2.11}$$

where again (,) denotes partial derivative.

In general relativity, the coordinates are chosen such that the affine connection introduced on the manifold is usually the Christoffel connection, the components of which are symmetric in their lower two indices (see above), thereby directly introducing the metric into the definition of curvature.

The path of any particle in the manifold of general relativity is affected by the curvature of the manifold. The matter in turn defines the geometry through Einstein's field equations:

$$R_{\mu\nu} - \tfrac{1}{2}\mathbb{R}g_{\mu\nu} = -8\pi T_{\mu\nu} \quad , \qquad (2.12)$$

with $R_{\mu\nu}=R^{\lambda}{}_{\mu\nu\lambda}$ being the components of the Ricci tensor (the contraction of the Riemann tensor), $\mathbb{R}$ the Ricci scalar, $\mathbb{R}=g^{\mu\nu}R_{\mu\nu}$, and $T_{\mu\nu}$ the components of the stress-energy tensor. We have chosen units such that c=G=1.

The field equations (2.12) are a complicated set of coupled, non-linear partial differential equations. In cosmology, we simplify these equations by imposing symmetries on the solution.

2.2 Symmetries of Spacetime

In this section we are interested in the symmetries which may be imposed on a four-dimensional manifold representing the spacetime of general relativity. Such symmetries help to reduce the complexity of the Einstein field equations, which define the geometry of spacetime given the matter distribution. We will express the symmetry of the models obtained in a coordinate-free manner by the use of differential forms and vector fields defined in the preceding section.

A symmetrical cosmological model is a manifold, M, on which the metric is invariant under a certain (specified) group of transformations. That is, each operation of the symmetry group corresponds to a map of M onto itself. This map (isometry) carries a point P into another point Q at which the metric is the same, when expressed in a coordinate independent way.

The description of the invariance of a metric under a group (*Lie group*) of isometries is achieved by directing attention to the infinitesimal transformations (*Lie algebra*) in the group. Other members of the group can be obtained from the infinitesimal members by *exponentiation* (repeated application of the infinitesimal members, Helgason 1962). Thus, a symmetric cosmological model is found by imposing the structure of a Lie algebra, although Lie group terminology is used. Before considering particular symmetries of the spacetime it is necessary to introduce some useful concepts.

To describe an infinitesimal transformation it is convenient to revert to coordinates (Misner 1964). Consider a point P_o in a neighbourhood, N, in which coordinates x^μ (μ=1,...,n) are used. A point P in N will have coordinates $x_P{}^\mu$. An *infinitesimal transformation* is of small effect and therefore carries points in N', a small neighbourhood of P_o which lies within N, into other points of N. Our transformation may be described in N' by n functions f^μ of the coordinates x^μ. The point P is carried to the point Q in N with the coordinates $\bar{x}_Q{}^\mu$:

$$f^\mu(x_P{}^\mu) = f^\mu(P) = \bar{x}_Q{}^\mu \qquad . \tag{2.13}$$

An infinitesimal transformation has the form

$$f^\mu(P) = x_P{}^\mu + \epsilon a^\mu(P) \qquad . \tag{2.14}$$

The number ϵ is meant to be so small that points in N' are carried only to points in N. The vector field $\mathbf{X}=a^\mu\partial_\mu$ describes the magnitude and direction of the transformation, ($\partial_\mu=\partial/\partial x^\mu$ is a basis vector).

A transformation acting on a space induces a transformation which carries a vector at point P into a vector at the image point Q. [This statement holds for any tensor, not just a vector which is a tensor of type (1,0), and so is completely general]. This transformation of vectors defines a new vector, whose value at Q is the "same" as the value of the vector at P. It can be shown that a vector $\mathbf{Y}=b^{\mu}\partial_{\mu}$ will change by the formula

$$Y_{new} = b^{\mu}{}_{new}(Q)\partial_{\mu} = [b^{\mu}(P)+\epsilon a^{\mu}{}_{,\nu}b^{\nu}(P)]\partial_{\mu} \quad . \qquad (2.15)$$

The value $\mathbf{Y}_{new}$ is what one would expect to see at Q if $\mathbf{Y}$ did not change in the direction given by the transformation vector field $\mathbf{X}$. $\mathbf{Y}-\mathbf{Y}_{new}$ is the observable change in the vector $\mathbf{Y}$. The measure of this change is,

$$\begin{aligned}(Y-Y_{new})^{\mu} &= b^{\mu}(Q)-b^{\mu}{}_{new}(Q) = b^{\mu}(Q)-b^{\mu}(P)-\epsilon a^{\mu}{}_{,\nu}b^{\nu}(P)\\ &= b^{\mu}(x^{\alpha}+\epsilon a^{\alpha}) - b^{\mu}(x^{\alpha}) - \epsilon a^{\mu}{}_{,\nu}b^{\nu}(P) \quad .\end{aligned}$$

Dividing by ϵ and letting $\epsilon\to 0$, we have the *Lie derivative* of Y with respect to $\mathbf{X}$,

$$\mathcal{L}_{\mathbf{X}}\mathbf{Y} = (b^{\mu}{}_{,\sigma}a^{\sigma} - a^{\mu}{}_{,\sigma}b^{\sigma})\partial_{\mu} \quad . \qquad (2.16)$$

This expression is simply the commutator of $\mathbf{X}$ and $\mathbf{Y}$,

$$\mathcal{L}_{\mathbf{X}}\mathbf{Y} = [\mathbf{X},\mathbf{Y}] \quad . \qquad (2.17)$$

The Lie derivative can easily be extended to arbitrary tensors, by requiring that it acts as a differentiation with respect to the tensor product $\tilde{\omega}^1\otimes\tilde{\omega}^2$, where $(\tilde{\omega}^1\otimes\tilde{\omega}^2)(\mathbf{U},\mathbf{V})=\tilde{\omega}^1(\mathbf{U})\tilde{\omega}^2(\mathbf{V})$, and that $\mathcal{L}_{\mathbf{X}}f=\mathbf{X}f$, where f is a function. In the coordinated system used above, if the tensor, $\underline{\mathbf{T}}$, has components $T^{\alpha\beta}{}_{\gamma\delta}$, $\mathcal{L}_{\mathbf{X}}\underline{\mathbf{T}}$ is given by,

$$(\mathcal{L}_{\mathbf{X}}\mathbf{T})^{\alpha\beta}{}_{\gamma\delta} = T^{\alpha\beta}{}_{\gamma\delta,\sigma}a^{\sigma} - T^{\sigma\beta}{}_{\gamma\delta}a^{\alpha}{}_{,\sigma} - T^{\alpha\sigma}{}_{\gamma\delta}a^{\beta}{}_{,\sigma}$$

$$+ T^{\alpha\beta}{}_{\sigma\delta}a^{\sigma}{}_{,\gamma} + T^{\alpha\beta}{}_{\gamma\sigma}a^{\sigma}{}_{,\delta} \quad . \tag{2.18}$$

All of the commas (partial derivatives) may be replaced by semi-colons (covariant derivative) without affecting the correctness of this relation. Thus, the Lie derivative is independent of both metric and connections. A comprehensive discussion of Lie derivatives and their application is given by Yano (1955).

A transformation which leaves the metric $\mathbf{g}$ invariant is called an *isometry*. An infinitesimal isometry is described by a vector $\mathbf{v}$, called a *Killing vector* (Killing 1892), which is said to *generate* isometries. A Killing vector thus satisfies

$$\mathcal{L}_{\mathbf{v}}\mathbf{g} = 0 \quad . \tag{2.19}$$

In other words, the derivatives of the functions $g_{\mu\nu}$ in the direction of $\mathbf{v}$ are zero. That is, the geometry of the manifold is left completely unchanged by a translation of points through the infinitesimal displacement $\epsilon\mathbf{v}$, where ϵ is small.

Equation (2.19) leads to the Killing equation for the components a_{μ} of the contravariant form of $\mathbf{v}$ in an arbitrary basis, viz.,

$$a_{\mu;\nu} + a_{\nu;\mu} = 0 \quad , \tag{2.20}$$

(cf. Yano and Bochner 1953). Thus, a vector field $\mathbf{v}(P)$ generates an isometry if and only if it satisfies Killing's equation, (2.20). It is important to notice that if $\mathbf{v}_1$ and $\mathbf{v}_2$ are two Killing vectors, then the linear combination $a_1\mathbf{v}_1+a_2\mathbf{v}_2$ is a Killing vector if a_1,a_2 are two constants. However, if a_1,a_2 are functions of position, $a_1\mathbf{v}_1+a_2\mathbf{v}_2$ is a vector field, but not necessarily a Killing vector. The commutator of two Killing vectors, $[\mathbf{v}_1,\mathbf{v}_2]$, is also a Killing vector.

The set of isometries of a manifold M has the structure of a group: An associative product is defined (the product of isometries A and B is A followed by B), an inverse exists for each element, and a unit transformation (the identity) exists. The group of isometries is the *symmetry group* of M. Isometries are obtained from the Killing vectors by exponentiation in the same way that group elements are obtained from the infinitesimal generators which form the Lie algebra of the group.

Many manifolds of interest in physics have metrics and it is therefore of considerable interest whenever the metric is invariant with respect to some vector field, cf. equation (2.19). We stated, above, that such a vector field is termed a Killing vector field. A convenient way of identifying a Killing vector is to find a coordinate system in which the components of the metric are independent of a certain coordinate, then the basis vector for that coordinate is a Killing vector.

As an example we shall determine the Killing vector fields of the three-dimensional Euclidean space. The metric in Cartesian coordinates has components,

$$g_{ij} = \delta_{ij} \quad , \tag{2.21}$$

which is independent of x, y and z. Therefore, $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$ are Killing vectors. The same metric in spherical polar coordinates has components,

$$\begin{aligned} g_{rr} &= \frac{\partial}{\partial r}\cdot\frac{\partial}{\partial r} = 1 \quad , \\ g_{\theta\theta} &= \frac{\partial}{\partial \theta}\cdot\frac{\partial}{\partial \theta} = r^2 \quad , \\ g_{\phi\phi} &= \frac{\partial}{\partial \phi}\cdot\frac{\partial}{\partial \phi} = r^2\sin^2\theta \quad , \end{aligned} \tag{2.22}$$

and we see that $\partial/\partial\Phi$ is a Killing vector.

In general, therefore, the presence of a symmetry means that a coordinate x^0 may be chosen such that

$$\mathbf{v} = \partial/\partial x^0 \quad , \quad g_{\mu\nu,x^0} = 0 \quad . \tag{2.23}$$

Then,

$$ds^2 = g_{\mu\nu}(x^1,x^2,x^3)dx^\mu dx^\nu \quad .$$

Common examples occur when the metric is stationary, in which case the Killing vector, $\mathbf{v}$, is timelike, when the metric has axial symmetry (and is invariant under the coordinate transformation $\Phi \rightarrow \Phi$+constant where Φ is the usual polar angular coordinate and $\mathbf{v}=\partial/\partial\Phi$), or when the geometry has the same cross-section for all points on one coordinate axis, say the z-axis (and the metric is thus invariant under $z \rightarrow z$+constant and $\mathbf{v}=\partial/\partial z$).

Other assumptions on the spacetime may be that the Weyl curvature tensor belongs to a particular Petrov type, that the metric has some special form and so on. In some of these special cases the field equations may be simplified to such an extent that they can be fully integrated. For example, Kinnersley (1969) found *all* vacuum metrics for which the Weyl tensor is Petrov type D. In general, however, some particular assumption is insufficient to allow for such full integration.

The groups classified by Petrov (1969) serve as isometry groups of a four-dimensional manifold with the appropriate metric signature. A list of all of the three-dimensional Lie algebras, each of which uniquely determines the local properties of a three-dimensional group, has been given by Ryan and Shepley (1975). Each of these may be used as the isometry group of a spatially-homogeneous cosmological model, the Bianchi Type universes. [Bianchi (1897) studied the set of all three-dimensional spaces which are homogeneous in the sense that there exist vectors **V** which leave the metric invariant]. Such spatially-homogeneous models are consistent with the observed distribution of matter in the universe today, e.g. Seldner et al. (1977).

Bianchi's studies showed that there are only nine independent Lie groups which satisfy the homogeneity condition. These nine groups are labelled Bianchi Types I to IX and can be identified by the values of the structure constants, $C^{\mu}{}_{\nu\lambda}$, of the particular group, given by the commutator of the Killing vectors associated with the homogeneity, i.e. $[\xi_{\nu},\xi_{\lambda}]=C^{\mu}{}_{\nu\lambda}\xi_{\mu}$. These structure constants are equal to the structure constants of the Lie algebra introduced in the affine connection (2.7).

One of the most important of the isometry groups is that of the Bianchi Type IX spaces, to which the Friedmann-Robertson-Walker universe belongs. Bianchi Type IX spaces are invariant under SO(3), the special orthogonal group in three dimensions, which is isomorphic to the three-dimensional rotation group. SO(3), a subgroup of the Euclidean symmetry group O(3), consists of matrices with determinant +1 and can, therefore, be shown to be the group of rotations. (The remaining matrices of O(3) can be interpreted as inversions). Thus, any matrix of SO(3) is equivalent to successive rotations in independent two-dimensional planes. This means that when a model universe M has the invariance group SO(3), the invariant hypersurfaces are taken to be three-spheres and the manifold is spherically symmetric about any point. If the three-spheres are spacelike then any fourth invariant vector, $\mathbf{Y}_0$, will be timelike. This vector may be chosen freely and one convenient choice is to take the vector perpendicular to the spacelike S^3's and of unit length, i.e. we choose a synchronous coordinate system. Thus, in terms of the dual one-forms, the metric of M will then be of the form,

$$\mathbf{g} = \underline{\underline{\mathbf{ds}}}^2 = \tilde{d}t^2 - g_{ij}(t)\tilde{\omega}^i \otimes \tilde{\omega}^j \quad . \tag{2.24}$$

Each g_{ij} is a function of the proper time, t, alone. The $\tilde{d}t$, $\tilde{\omega}^i$ are the duals of $\mathbf{Y}_0$, $\mathbf{Y}_i$, respectively.

In the case of the FRW metric, the g_{ij}, expressed in this $\tilde{\omega}^i$ basis, would have the form

$$g_{ij} = G^2 \delta_{ij} \qquad \text{with} \quad G=G(t) \quad . \tag{2.25}$$

The fact that g_{ij} is diagonal and has three equal entries shows that the metric of the FRW universe is *isotropic*. In other words, the FRW universe has symmetries in addition to the homogeneity of spacelike sections which is granted by invariance under SO(3). This additional symmetry of the FRW universe - its isotropy - may be expressed by the statement that its metric is invariant under rotations about any axis in a homogeneous three-space H(t).

Another example of a universe which is invariant under SO(3) is the Taub universe, Taub (1951). This (vacuum) model is rotationally invariant about only one axis in each three-space and has the metric,

$$\underline{\underline{\mathbf{ds}}}^2 = dt^2 - b_1{}^2(\tilde{\omega}^1)^2 - b_2{}^2[(\tilde{\omega}^2)^2+(\tilde{\omega}^3)^2] \quad . \tag{2.26}$$

The manifold is again $S^3 \times \mathbb{R}$. As we see, by the form of this metric, where the b's are functions of t only, the Taub universe is spatially homogeneous with invariance group SO(3).

A more general matter-filled SO(3)-homogeneous model may be imagined in which g_{ij} is not diagonal as a function of t (nor may be made diagonal by changing the choice of ω^i). This model exhibits rotation and anisotropy, cf. the models described in Chapter 4 of this thesis.

Symmetries based on isometry groups and Killing vector fields, therefore, provide extremely useful methods for solving the rather complicated Einstein equations describing a cosmology. However, there exist other kinds of assumptions which can be made on the metric tensor and thereby help to reduce the complexity of the field equations. For instance, the assumption of the existence of a second or higher order Killing tensor rather than that of a Killing vector. Killing tensors yield constants of motion and enable, for example, the separation of the Hamilton-Jacobi equation in most Petrov type D vacuum solutions, in particular the Kerr-Newman solution, (cf. Hughston and Sommers 1973). However, Killing tensors are not easily handled and we will not consider them further.

One other type of assumption which may be made is that there exist symmetries such as homothetic motions, conformal motions and curvature collineations as discussed for example, by Katzin et al. (1969). We shall be mostly concerned with the homothetic and conformal motions as these relate directly to the symmetries of similarity solutions of the first kind. The possibility of higher order self-similar symmetries will also be discussed.

2.3 Self-Similar Symmetries

In this section we shall approach the subject of self-similarity from two directions. The first will continue the discussion of the previous section and consider the equivalence of self-similar motions (of the first kind) with homothetic motions in a manifold. The second will be to use the dimensional methods developed by Sedov (1959) and Zel'dovich and Raizer (1967), in the study of hydrodynamical fluids. Although the second treatment is somewhat less mathematically aesthetic than the first, it does provide some useful physical insight into the symmetries imposed on the motion of the physical variables of any given problem. This division of "tactics" also highlights the two, often distinct, approaches to doing cosmology; treating the universe as (i) a manifold on which to apply the tools of differential geometry and (ii) a fluid which satisfies continuity relations, equations of motion etc. The fact that the two methods are complimentary is neatly encapsulated in the Einstein field equations of general relativity, with 'geometry' on the left hand side of the equations and 'physics' on the right.

A. Geometrical Approach

Perhaps the simplest generalisation of a Killing motion is a homothetic one, M^{c}Intosh (1980). In this case the vector $\mathbf{v}$ satisfies the equation,

$$(\mathcal{L}_{\mathbf{v}}\mathbf{g})_{\mu\nu} = \Phi g_{\mu\nu} \quad , \tag{2.27}$$

where Φ is a constant. If, on the other hand, Φ is an arbitrary scalar function, then $\mathbf{v}$ is termed a conformal vector field. Collinson and French (1967) showed that in a non-flat empty spacetime a conformal motion must be a homothetic one (unless it is of Petrov Type N).

Thus, if we are to use these symmetries to integrate the field equations to obtain new solutions, there is no need to consider the possibility of conformal motions with Φ non-constant and we shall therefore take $\Phi=2k$ in what follows.

To emphasise the symmetry inherent in homothetic motions, equation (2.27) can be written in the form,

$$\mathcal{L}_{\mathbf{v}}(g^{-1/n}g_{\mu\nu}) = 0 \quad , \quad g \equiv |\mathrm{Det}(g_{\mu\nu})| \quad , \tag{2.28}$$

where n is the dimension of the manifold. Thus, we see immediately the geometric object which is left invariant under a homothetic motion. In general, if a space admits an infinitesimal point transformation with respect to which the Lie derivative of some geometric object vanishes, then it also admits a one-parameter invariance group (physical symmetry) of this geometric object, (Yano 1955). Examples of such symmetries include; the conformal or homothetic motions discussed above, Killing motions where the invariant geometric object is the metric tensor, $g_{\mu\nu}$, itself (see §2.2), affine collineations which demand invariance of the Christoffel symbols, $\Gamma^{\lambda}{}_{\mu\nu}$, and curvature collineations for which the Riemann curvature tensor, $R^{\lambda}{}_{\mu\nu\tau}$, is the relevant geometric object defining the symmetry. A discussion of these various types of symmetry was given by Katzin et al. (1969). We shall restrict our attention to the homothetic (conformal) motions.

To obtain a physical notion of equation (2.27) let us form the *conformal* Killing equation, using equation (2.18), viz.,

$$g_{\mu\nu,\sigma}a^{\sigma} + g_{\mu\sigma}a^{\sigma}{}_{,\nu} + g_{\sigma\nu}a^{\sigma}{}_{,\mu} = 2kg_{\mu\nu} \quad , \tag{2.29}$$

where $\mathbf{v}=a^{\mu}\partial_{\mu}$. Restricting ourselves to two dimensions we have for a flat two-space,

$$ds^2 = dx^2 + dy^2 \qquad ,$$

that equation (2.29) reduces to the expression

$$a^{\mu}{}_{,\nu} + a^{\nu}{}_{,\mu} = 2k\delta^{\mu}_{\nu} \qquad , \tag{2.30}$$

(μ,ν=1,2, $g_{\sigma\sigma}$=1), which is just the flat-space form of equation (2.20) for a conformal symmetry. If we take the components of a^{μ} to be a^1=u, a^2=w, then equations (2.30) reduce to the following:

$$\begin{aligned} \frac{\partial u}{\partial x} &= k = \frac{\partial w}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial w}{\partial x} \end{aligned} \qquad , \tag{2.31}$$

which are the Cauchy-Riemann equations for transformations in complex analysis. Thus, Killing's equation for a conformal symmetry can be regarded as the general representation of the Cauchy-Riemann equations in curved space.

For a homothetic motion, a coordinate x^0 can be chosen such that,

$$\mathbf{v} = \partial/\partial x^0 \qquad , \qquad g_{\mu\nu,x^0} = 2kg_{\mu\nu} \qquad , \tag{2.32}$$

and the line element can then be written as

$$ds^2 = \exp[2kx^0]h_{\mu\nu}(x^1,x^2,x^3)dx^{\mu}dx^{\nu} \qquad . \tag{2.33}$$

For example, one common metric which admits a homothetic motion is that of the Einstein-de Sitter cosmology, for which

$$ds^2 = dt^2 - t^{4/3}[dx^2+dy^2+dz^2] \qquad , \tag{2.34}$$

in which case the homothetic vector is

$$\mathbf{v} = t\frac{\partial}{\partial t} + \frac{1}{3}\left[x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} \right] \qquad . \tag{2.35}$$

In this form all of the coordinates are scaled, under the action of (2.35), although not all by the same amount. The geometry is mapped along the congruence of curves, whose tangent vectors are **v**, to one of the same "shape", but where lengths are changed by a fixed amount. Another obvious example of a homothetic motion is the mapping along the axis of symmetry of a cone such that the cross-sections remain of similar shape but increase or decrease in size. The word self-similar is often used to describe a homothetic mapping. In this thesis our main consideration is to investigate the properties of cosmological models which admit self-similar symmetries. Let us, therefore, study the homothetic motions in a little more detail.

In the preceding section we defined the Lie derivative in terms of infinitesimal transformations. For a spherically symmetric spacetime, which admits a Killing vector $\partial/\partial\phi$, the infinitesimal transformation corresponding to the conformal motion (2.27), is defined by

$$t' = t+\epsilon\beta(r,t) \quad ; \quad r' = r+\epsilon\alpha(r,t) \quad ; \quad \theta'=\theta \quad ; \quad \phi'=\phi \quad , \tag{2.36}$$

where α and β are two arbitrary functions of r and t, and ϵ is an infinitesimal parameter. The associated *conformal* Killing vector is,

$$\mathbf{v} = (\beta,\alpha,0,0) \quad , \tag{2.37}$$

with the conformal Killing equation given by

$$a_{\mu;\nu} + a_{\nu;\mu} = 2kg_{\mu\nu} \quad , \tag{2.38}$$

where $\mathbf{v}=a^{\mu}\partial_{\mu}$. Given a metric of the form

$$ds^2 = e^{\sigma}dt^2 - e^{\omega}dr^2 - r^2(d\theta^2+\sin^2\theta d\phi^2) \quad , \tag{2.39}$$

we see that only four equations in (2.38) are not identically zero:

$$2\frac{\partial\beta}{\partial r} + \beta\frac{\partial\sigma}{\partial t} + \alpha\frac{\partial\sigma}{\partial r} = 2k \quad , \tag{2.40}$$

$$e^{\sigma}\frac{\partial \beta}{\partial r} + e^{\omega}\frac{\partial \alpha}{\partial t} = 0 \quad , \qquad (2.41)$$

$$2\frac{\partial \alpha}{\partial r} + \alpha\frac{\partial \omega}{\partial r} + \beta\frac{\partial \omega}{\partial t} = 2k \quad , \qquad (2.42)$$

$$\alpha = kr \quad . \qquad (2.43)$$

From equations (2.43) and (2.41), we see that the infinitesimal transformation is necessarily of the form, (Munier et al. 1980),

$$\begin{aligned} t' &= t + \epsilon\beta(t) \\ r' &= r + \epsilon kr \end{aligned} \quad . \qquad (2.44)$$

Following the analysis of Munier et al. (1980) we introduce new 'scaled' coordinates:

$$r^* = r \quad ; \quad \frac{\partial t^*}{\partial t} = \frac{t^*}{\beta(t)} \quad . \qquad (2.45)$$

Substituting the transformed system, $(r^*,t^*,g_{\mu\nu}{}^*)$, back into the conformal Killing equation, we find that a new dimensionless variable emerges,

$$\xi = rt^{-k} \quad , \qquad (2.46)$$

where we have dropped the asterisks. The variable ξ is an invariant of the infinitesimal group (2.44) and is termed the self-similar variable.

It is remarkable that in general relativity every conformal homothetic group is equivalent because of the possible rescaling of the variables, (2.45). Self-similarity in relativity theory is, therefore, an extremely general process.

In the particular case, where the homothetic constant is unity (k=1), then the invariant of the group reduces to $\xi=r/t$, [or equivalently $\xi = t/r$, see Chapter 4], and the metric coefficients, g_{00} and g_{11}, are expressible in terms of ξ alone.

Note that when the source of the gravitational field is a perfect fluid it is a consequence of the self-similar (homothetic) motion and the transformation properties of the Einstein tensor, $G_{\mu\nu}$, that the four-velocity, u^{μ}, is conformally invariant. That is,

$$u^{\mu}{}_{;\nu}v^{\nu} - v^{\mu}{}_{;\nu}u^{\nu} = -u^{\mu} \quad . \tag{2.47}$$

Another consequence of the self-similar symmetry is that only equations of state of the form

$$p = \alpha\rho \quad , \quad (\alpha \text{ constant}) \tag{2.48}$$

are possible, (Cahill and Taub 1971). A comprehensive discussion on the uses of self-similarity in general relativity is given by Eardley (1974).

B. Classical Hydrodynamical Approach

Similarity solutions in classical hydrodynamics have been a fruitful source of models for physical systems having no intrinsic scale of length or mass, or time. In this sub-section we will discuss the development of similarity techniques in mechanics from their origin in dimensional analysis. This subject is dealt with comprehensively in the textbooks by Sedov (1959) and Zel'dovich and Raizer (1967).

Every phenomenon in mechanics is determined by a series of variables, such as energy, velocity and stress. Problems in dynamics reduce to the determination of certain functions and characteristic parameters. The relevant physical laws and geometrical relations are represented as functional equations, usually differential equations. In purely theoretical investigations, we use these equations to establish the general qualitative properties of the motion and to calculate the unknown physical variables by means of mathematical analysis. However, very often the problem cannot be formulated mathematically because the mechanical system to be investigated is too complex to be described by a satisfactory model. In general, we begin every investigation of a natural phenomenon by finding out which physical properties are important and looking for mathematical relations between them which govern the behaviour of the phenomenon.

Many phenomena cannot be investigated directly and to determine the laws governing them we must perform experiments on similar phenomena which are easier to handle. Theoretical analysis is needed when formulating such experiments to determine the values of particular parameters of interest. In general, it is very important to select the non-dimensional parameters correctly; there should be as few parameters as possible and they must reflect the fundamental effects in the most convenient way. This preliminary analysis of a phenomenon and the choice of a system of definite non-dimensional parameters is made possible by dimensional analysis and similarity methods.

We call quantities dimensional if their numerical values depend on the scale used, i.e. on the system of measurement units. Quantities are non-dimensional when their values are independent of the system of measurement units. The subdivision of quantities into dimensional and non-dimensional is to a certain extent a matter of convenience. However, the quantities of length, time and mass (or energy, or force) are usually regarded as dimensional, whereas angles and the ratio of lengths are non-dimensional.

In practice, it is sufficient to establish the units of measurement for three quantities; precisely which three depends on the particular conditions of the problem. In physical investigations it is convenient to take the units of length, time and mass. (Such a system of measurement in cosmology is obtained by choosing $G=c=1$, where G is the gravitational constant and c is the speed of light, allowing all dimensions to be measured in one unit, usually length).

In particular, dimensional and similarity theory is of a great value when making models of various phenomena. The basic idea of modelling is that the information required about the character of the effects and the various quantities related to the phenomenon under natural conditions can be derived from the results of experiments with models. Modelling is based on an analysis of physically similar phenomena. We replace the study of the natural phenomenon, which interests us, by the study of a physically similar phenomenon, which is more convenient and easier to reproduce. Physical similarity can be considered as a generalisation of geometric similarity. Two geometric figures are similar if the ratio of all the corresponding lengths are identical.

There are various ways of defining dynamical or physical similarity. We shall adopt the definition of similar phenomena used by Sedov (1959), viz.,

'Two phenomena are similar, if the characteristics of one can be obtained from the assigned characteristics of the other by a simple conversion, which is analogous to the transformation from one system of units of measurement to another.'

The 'scaling factor' must be known in order to accomplish this conversion. Further, the necessary and sufficient conditions for two phenomena to be similar are that the numerical values of the non-dimensional coefficients forming the basic system are constant. These conditions are called *similarity criteria*.

For instance, all the non-dimensional quantities in the problem of steady, uniform motion of a body in an incompressible, viscous fluid are defined by two parameters: the angle of attack α and the Reynolds number R. The conditions of physical similarity are represented by the relations

$$\alpha = \text{constant} \quad , \qquad R = \frac{vd\rho}{\mu} = \text{constant} \quad , \tag{2.49}$$

where v is the velocity of the fluid, d is the scale size of the body, ρ is the fluid density and μ is the dynamic viscosity. Thus, in fluid mechanics, flows of the same type with the same Reynolds number are similar. This is dubbed the law of similarity, (Reynolds 1883).

We can extend this law of similarity to the situation where the two fluid motions being compared belong to the same fluid, but at different times. The type of motion in which the distributions of the flow variables remain similar (in the above sense) to themselves with time and vary only as a result of changes in scale is called *self-similar*. For example, the motion of a compressible medium, in which the dimensionless parameters depend only on the combination

$$\frac{x}{bt^k} \ , \ \frac{y}{bt^k} \ , \ \frac{z}{bt^k} \quad , \tag{2.50}$$

where x, y, z denote Cartesian coordinates, t is the time and b is a constant with dimensions LT^{-k}, will be called self-similar with a centre of similarity at the origin of the coordinate system. (This corresponds to the conformal motion where the dimensionless independent variable is given by equation (2.46)). It is easy to discover the general character of all problems for which self-similarity exists. It is sufficient for the system of dimensional characteristic parameters, prescribed in part by supplementary conditions and in part by boundary or initial conditions, to contain not more than two constants with independent dimensions other than length or time.

Generally speaking, for self-similarity to exist in the motion of a compressible fluid it is necessary that the formulation of the problem should not contain a characteristic length or time.

To fix our ideas, let us consider the motion of a fluid in one-dimension, i.e. all the properties of the fluid depend only on one geometric coordinate and on the time. It can be shown that the only possible one-dimensional motions are produced by spherical, cylindrical and plane waves (Liubimov 1956). We can distinguish the problems which can be solved by the methods of dimensional analysis, i.e. by analysing the dependent variables and the fundamental parameters of one-dimensional motion. The basic physical variables in the Eulerian approach are the velocity v, the density ρ and the pressure p and the characteristic parameters are the linear coordinate r, the time t and the constants which enter into the equations, the boundary and the initial conditions of the problem.

Since the dimensions of the quantities ρ and p contain the mass, at least one constant a, the dimensions of which also contain the mass, must be a characteristic parameter. Without loss of generality, it can be assumed that its dimensions are

$$[a] = ML^{k}T^{s} \quad , \qquad (2.51)$$

where we will use the notation [ ] for the dimensions of any quantity. We can then write for the velocity, density and pressure

$$v = \frac{r}{t}V \quad , \quad \rho = \frac{a}{r^{k+3}t^{s}}R \quad , \quad p = \frac{a}{r^{k+1}t^{s+2}}P \quad , \qquad (2.52)$$

where the 'reduced functions', V, R and P, are arbitrary quantities and, therfore, can depend only on non-dimensional combinations of r, t and other parameters of the problem. In the general case, they are functions of two non-dimensional variables. However, if an additional characteristic parameter b can be introduced with dimensions independent of those of a, the number of independent variables which can be formed by combining a and b is reduced to one. (Incidentally, it is interesting to note that similar functions, called homology-invariant variables, were introduced by Bondi and Bondi (1949) in the theoretical study of stellar structure to aid calculations of the relevant equations and to avoid the instabilities which occurred in other methods. These h-variables, as they were called, played a similar role to the reduced functions, in that they greatly simplified the equations involved).

Since the dimensions of the constant a contain the mass, we can chose the constant b, without loss of generality, so that its dimensions do not contain the mass, i.e.

$$[b] = L^{m}T^{n} \quad . \qquad (2.53)$$

The single non-dimensional independent variable in this case will be $r^{m}t^{n}/b$, which can be replaced, for $m \neq 0$, by the variable

$$\xi = \frac{r}{b^{1/m}t^{\delta}} \quad , \quad \text{where } \delta = -\frac{n}{m} \quad . \qquad (2.54)$$

If m=0, V, R and P depend only on the time t and the velocity v is proportional to r. The solution depending on the independent variable may contain a number of arbitrary constants. [It was noted by

Stanyukovich (1955) that, in addition to the power-law self-similarity of equation (2.54), it is also possible to have exponential self-similarity, in which $\xi = re^{-kt}/A$, where k and A are constants. The majority of problems of practical interest have a power-law character].

This argument shows that, when the characteristic parameters include two constants with independent dimensions in addition to r and t, the partial differential equations satisfied by the velocity, density and pressure in the one-dimensional unsteady motion of a compressible fluid can be replaced by ordinary differential equations for V, R and P. Such motions are called self-similar.

In the case of a perfect, inviscid, non-heat-conducting fluid the equations of motion, continuity and energy take the form

$$\frac{\partial v}{\partial t} + \frac{v\partial v}{\partial r} + \frac{1}{\rho}\frac{\partial p}{\partial r} = 0 \quad ,$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial r}(\rho v) + (\nu-1)\frac{\rho v}{r} = 0 \quad , \qquad (2.55)$$

$$\frac{\partial}{\partial t}\left[\frac{p}{\rho^{\gamma}}\right] + \frac{v\partial}{\partial r}\left[\frac{p}{\rho^{\gamma}}\right] = 0 \quad ,$$

where γ is the adiabatic index; $\nu=1$ for planar flow, $\nu=2$ for flow with cylindrical symmetry and $\nu=3$ for flow with spherical symmetry. These equations do not contain any dimensional constants. Consequently, the question of the self-similarity of the motion is determined by the number of parameters with independent dimensions introduced by the remaining conditions of the problem. If there are only two of these the motion is self-similar.

One illuminating way of describing the physical effect of self-similar motion on the fluid variables is demonstrated in Figure 2.1, which displays the density and velocity profiles in a centred

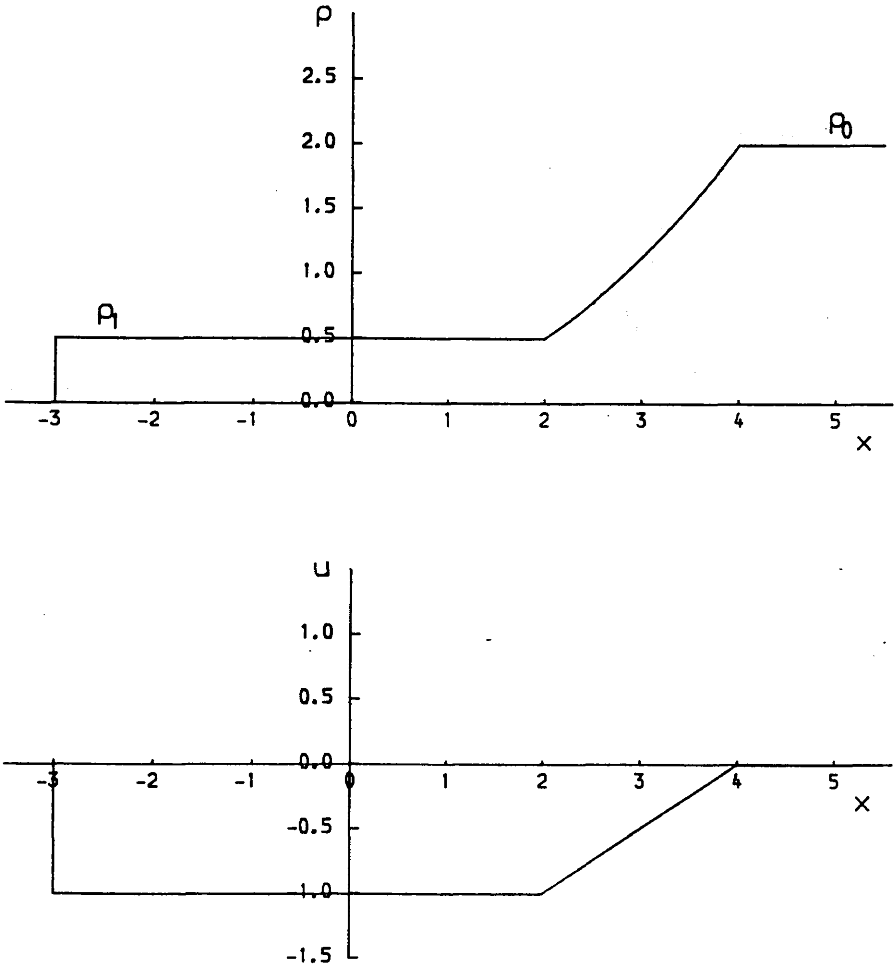


Figure 2.1 Density and velocity profiles in a centred rarefaction wave produced by the motion of a receding piston moving at a constant velocity.

rarefaction wave resulting from the influence of a piston receding with a constant velocity, (cf. Zel'dovich and Raizer 1967). If the motion is self-similar the distributions of all quantities with respect to the x-coordinate will change with time without changing their form; they remain similar to themselves. If we were to draw the profiles shown in Figure 2.1, using as the abscissa not x but the ratio x/t, we would obtain a "frozen" picture, one which does not vary with time. [A centred rarefaction similarity solution has $\delta=1$, $m=-n$, cf. equation (2.54)].

Many phenomena in nature, although not exactly self-similar throughout their evolution, do exhibit self-similar behaviour in the limit as $t\to\infty$, i.e. in a region far from the initial conditions and the influence of the boundary conditions. The existence of such a limiting solution corresponds to the concept of intermediate asymptotics as reviewed by Barenblatt and Zel'dovich (1972).

An intermediate asymptotic regime is a region in which the behaviour of the solutions is no longer dependent on the details of the initial and/or boundary conditions, but where the system is still far from being in a state of equilibrium. For example, in the case of a strong explosion, discussed by Barenblatt and Zel'dovich (1972), where there is a phase change across the resulting shock front, the solution appropriate to the intermediate asymptotic region corresponds to the instantaneous release of an amount of energy E in an infinitesimal domain of radius R_0. In this example, it can be said that the intermediate asymptotic solutions do not 'remember' either the energy E or the size R_0 of the domain, in which the energy is released at the initial time, independently but a combination of these quantities.

The study of self-similar motions is, therefore, of great physical interest. The fact that it is possible to reduce a system of partial differential equations to a system of ordinary differential equations for new reduced functions simplifies the problem from the mathematical standpoint and in a number of cases makes it possible to find exact analytic solutions.

To complete this section we note that there exist two rather different kinds of self-similar solutions. The first type possess the property that the similarity exponent δ and the exponents of the dimensional constant b in all scales are determined either by dimensional considerations or from the conservation laws. Problems of this type always contain two parameters with independent dimensions. These parameters are used to construct one parameter whose dimensions contain the unit of mass, a, and another parameter, $A \equiv b^{1/m}$, that contains only the units of length and time. With the second parameter A it is possible to construct a dimensionless combination, the similarity variable $\xi = r/At^{\delta}$. The dimensions of the parameter A, $LT^{-\delta}$, are determined by the similarity exponent δ. Motions of this kind are called self-similar of the first kind and were considered extensively above.

In self-similar problems of the second kind, the exponent δ cannot be found from dimensional considerations or from the conservation laws without solving the equations. In this case, the determination of the similarity exponent requires that the ordinary differential equations for the reduced functions be integrated. (It turns out that the exponent is found from the condition that the integral curve must pass through some singular point, as otherwise the boundary conditions cannot be satisfied. The existence of such a singular point is linked directly to the fact that the self-similar solution of the gasdynamic equations is only physically meaningful if it is single-valued. That is, each value of the independent variable ξ should correspond to unique values of the reduced functions of the problem, e.g. $\xi(V)$, $\xi(R)$, $\xi(P)$ should not have extrema. A more comprehensive discussion of this issue is given for the imploding shock wave problem in Zel'dovich and Raizer (1967), Vol. 2, Ch. XII.2).

Examination of solutions to specific problems of the second kind shows that in all these cases the initial conditions of the problem contain only one dimensional parameter with the unit of mass but lack another which could be used to form the parameter A. This circumstance eliminates the possibility of determining the number δ from the dimensions of A. Actually, of course, the problems do have a dimensional parameter A, with dimensions $LT^{-\delta}$, relevant to it, otherwise it would be impossible to construct the dimensionless combination $\xi=r/At^{\delta}$. However, the dimensions of this parameter (i.e. δ) are not dictated by the initial conditions of the problem, but rather are found from the solution of the equations. Thus, for instance, if the self-similar motion originated as a result of some non-self-similar flow that approaches a self-similar regime asymptotically, then the value of A can only be found by a numerical solution of the complete non-self-similar problem in which it is possible to follow the transition of the non-self-similar motion into the self-similar one. As an example of another problem which admits a self-similarity of the second kind, consider a cosmological fluid with a non-zero cosmological constant Λ (cf. Chapter 3). There are three dimensional constants in this problem; G (units $M^{-1}T^{-2}L^{3}$), c (LT^{-1}) and Λ (T^{-2}). The gravitational constant G immediately gives us a parameter which contains the units of mass. However, two independent dimensionless variables can be formed from the remaining two constants and the independent variables r and t, namely,

$$\xi = \frac{r}{ct} \quad , \quad \tau = \frac{r\Lambda^{1/2}}{c} \quad , \tag{2.56}$$

and we therefore cannot define a unique similarity exponent α. (For the variable ξ, α=1, for τ, α=0). Thus, we cannot define a self-similar solution of the first kind. However, it is found that by treating Λ as a strict constant and relating it to the energy density of the vacuum the system admits a self-similar symmetry of the second kind. The the analysis leading to this result will be discussed more fully in Chapter 3).

We noted above that self-similarity of the first kind can be represented as a homothetic (or conformal) motion of the underlying spacetime. However, in the more complex situation of a self-similar symmetry of the second kind, no such conformal analogue exists. It is to be expected that similarity motions of the second kind correspond to the global invariance of some geometric object but as yet no such object has been identified, (see the Appendix).

2.4 Applications of Self-Similarity

The investigations of self-similar motions, whether in geometric form, equation (2.27), or hydrodynamic form, equation (2.54), have applications in many branches of physics and astrophysics, particularly in the study of theoretical cosmology. Self-similar solutions are often the leading terms in an asymptotic expansion of a non-self-similar evolution, in a regime where the motion has "forgotten" to a considerable extent about the initial conditions and as such are likely to be frequently encountered in nature. Thus, self-similar solutions have greater physical interest than merely being a special class of mathematical solutions. This property of similarity solutions concerns the concept of intermediate asymptotics mentioned earlier. It is the purpose of this section to discuss a selection of these investigations in an attempt to demonstrate the wide-ranging applications of self-similar solutions in physics.

The self-similar description in fluid mechanics of a rarefaction wave in a compressible medium has been extended to the regime of laser fusion plasma physics. It was found (Varey and Sander 1969) that the expansion of a plasma into a vacuum produces an electro-acoustic (ion) rarefaction wave propagating into the plasma. Allen and Andrews (1970) considered a self-similar treatment to successfully describe this effect for a plasma bounded by a positively-charged sheath. For a similarity solution to be valid, in this situation, the processes of ionisation are assumed to be negligible over the time scale involved and the theory must be restricted to the case in which the plasma boundary moves with constant velocity. Denavit (1979) also considered the expansion of a collisionless plasma into a vacuum using a particle simulation code. The results of these simulations confirmed the existence of an ion front and verified the general features of self-similar solutions behind this front. The assumption of self-similarity also yields a linearly decreasing ion acoustic speed which is in good agreement with the computations.

A self-similar solution of the asymptotic type (i.e. a solution which is approached in the limit as $t \to \infty$) was also considered in the investigation of the ablative heat wave formed when a dense body is suddenly brought into contact with a thermal bath (Pakula and Sigel 1985). The self-similarity is a consequence of the fact that the solid may be considered as infinitely dense in the limit. It was found that there is a range of validity of the self-similar solutions which is dependent on the temperature of the thermal bath (T) and the time (t). The boundaries of the valid region are formed by three straight lines in the (T,t)-space; the first marks the boundary where local thermodynamic equilibrium becomes invalid, the second separates the heat wave regime from the ablative heat wave regime and the third marks the boundary of non-negligible radiation pressure. Such a study is important as it provides a good example of a system which is asymptotically self-similar. (One would expect this type of solution to occur more frequently in nature than the exact self-similar solutions). Brown and Emslie (1988) also found a solution where the validity of the self-similar symmetry is confined to a particular region of space and time when considering the heating of solar flares by an electron beam.

Further applications of self-similar flows can also be found in the hydrodynamic investigations of the propagation of shocks in accretion disks (Gaffet and Fukue 1983) or cosmological media (Bertschinger 1983), in the study of gravitationally bound stellar clouds (Henriksen and Turner 1984) and in the investigation of gravitational clustering (Efstathiou 1983).

Gaffet and Fukue (1983) used a self-similar analysis to construct a family of solutions that describes freely propagating shock waves in accretion disks, i.e. the case where there is no energy source other than the initial instantaneous release of the outburst energy. This family of solutions is interesting in that it displays a self-similarity of the second kind (cf. the strong explosion example in the preceding section) and when the shock wave is sufficiently weak, in comparison with the gravitational force, a critical point appears through which the solution should pass in order to have physical meaning. This critical point corresponds to the sonic limit of the motion. Beyond this point the regime becomes supersonic. On the contrary, for sufficiently weak shocks, no such critical point exists and the regime remains everywhere subsonic. The solutions are then self-similar of the first kind. Bertschinger (1983) also applied the analysis of Sedov (1959), regarding the use of similarity methods in the description of blast waves, in an attempt to find self-similar solutions for adiabatic shock waves in cosmologically varying media (homogeneous in space but time-varying and self-gravitating). In this investigation the shock wave formed satisfies the Sedov solution for adiabatic evolution in a constant medium initially, but it is modified by the cosmological expansion to form a self-similar thin dense shell which cools by adiabatic expansion and is unstable to gravitational fragmentation and collapse. The shock solution obtained is applied to the explosive amplification model of galaxy formation proposed by Ostriker and Cowie (1981). In this model rapid star formation and supernova explosions occurring in a collapsing protogalaxy power a galactic wind, producing a shock wave propagating into the intergalactic medium. Bertschinger showed that, for exact similarity solutions of shock waves in homogeneous cosmological media to exist, the background universe has to be Einstein-de Sitter. (Two other cases are possible but these have no gravity and are therefore of no immediate interest). He, further, states that the formation of galaxies in the self-similar shell may inject more energy into the intergalactic medium, causing a shock wave to continue propagating radially outward. This shock may fragment to form a second generation of galaxies, with parameters typical of small galactic clusters.

Henriksen and Turner (1984) considered the relation between internal velocity dispersions and cloud sizes in an ensemble of galactic molecular clouds. They argued that the ensemble of clouds could be regarded as elements in a self-similar regime of compressible turbulence. In this scenario the distinction between correlations "inside clouds" and "among clouds" becomes rather obscure (because no cloud is isolated) at least below some maximum scale. Rather each is an element of a larger unit and contains subunits. This hierarchical distribution is very suggestive of a self-similar process. Self-similar symmetry is well known to be present in turbulent flow (e.g. Cantwell 1981). Moreover, Henriksen and Turner extended such analyses to derive scaling laws for the turbulence of these gravitationally bound clouds.

The work of Efstathiou (1983) considered self-similar gravitational clustering in an attempt to explain the observed nature of the distribution of galaxies. If the clustering pattern does obey some simple similarity scaling then the clustering at some early time, apart from a change in length scale, would be statistically indistinguishable from the pattern observed today. The relevance of such a simplifying assumption to the actual clustering pattern is provided by the power-law shape of the two-point correlation function and the simple forms of higher order correlation functions (Peebles 1980).

Cosmologies, based on a similar hierarchical structure to that described by Efstathiou (1983), have been proposed by many authors on the strength of the observations of de Vaucouleurs (1970) that the density of matter in the universe is not uniform on a cosmological scale. Instead, de Vaucouleurs found that the density varies roughly as $d^{-1.7}$ for large distance, d. Bonnor (1972) attempted to model this density law by considering a specific Bondi-Tolman dust solution. His solution is chosen specifically so that the density on constant t slices varies as $r^{-1.5}$ for large r. However, the optical equations describing the propagation of light are not simple in this solution so that comparison with observation is difficult. Dyer (1979) followed the lead of Bonnor, but instead of assuming a particular density law, he assumed self-similarity. This assumption simplifies the analysis considerably although it is found to be only valid if the self-similar solution is used to represent a limited region of the universe. (A similar model was considered by Wesson (1979). However, his observational relations also have only limited regions of applicability).

The observed hierarchy of the structure of the universe also prompted the self-similar investigations of Henriksen and Wesson (1978a,b) which we will deal with in great detail in Chapter 4. [Later in this thesis we will also discuss, extensively, the use of similarity solutions in the study of the growth of primordial black holes (cf. Chapter 5). To obtain a significant distribution of black holes in the universe today, it was postulated that at some stage the black holes formed in the very early universe grew at the same rate as the universe particle horizon (Carr and Hawking 1974). The absence of any distinct length scale in such a scenario (all length scales grow at the same rate) suggests that a similarity solution may be applicable].

The study of self-similar solutions have also provided considerable insight into the more complicated results obtained from numerical solutions for the formation of structure in the universe. Fillmore and Goldreich (1984) investigated self-similar solutions in an attempt to describe the collapse of cold, collisionless matter within a background Einstein-de Sitter universe. They found that their results for planar symmetry display the same qualitative features found in other simulations (e.g. Melott 1983) and their spherically symmetric similarity solutions are compatible with the extended flat rotation curves observed in spiral galaxies (Rubin et al. 1980).

Finally, self-similarity (in its geometric form) has numerous applications in mathematical cosmology, i.e. the determination of exact solutions to the field equations. Many of these applications will be discussed in detail in the later chapters of this thesis. However, this section would not be complete without mentioning some of them.

Wainwright et al. (1979) found a number of inhomogeneous cosmological solutions of the Einstein field equations which have an irrotational perfect fluid, with equation of state $p=\rho$ (ρ is the energy density), as source. These solutions admit a two-parameter Abelian group of local isometries, but in general do not admit a third isometry and are thus classified as inhomogeneous. McIntosh (1978) demonstrated that all but one of the Wainwright et al. solutions admit homothetic motions which together with the two Killing motions span spacelike hypersurfaces. Each of these models thus admits a three-parameter similarity group of motions and is an example of a self-similar cosmology. The possible relevance of the equation of state $p=\rho$ as regards the matter content of the universe in its early stages, and in particular its relevance to self-similar cosmologies will be discussed at length in Chapters 4 and 5 of this thesis.

It was proved by McIntosh (1975, 1976) that, in order for a homothetic vector to be non-trivial in a perfect fluid solution of the field equations, either $p=\rho$ or the model is tilted. [A tilted solution is one in which the fluid velocity is not orthogonal to the group orbits, e.g. a radial tilting velocity will be of the form $u_\mu=(\alpha,\beta,0,0)$, where α,β may be functions of the coordinates r and t, cf. King and Ellis (1973)]. The solutions of Wainwright et al. (1979) agree with the statement of this theorem and moreover show that there are solutions with $p=\rho$ with non-trivial homothetic motions [$\Phi\neq 0$ in (2.27)] in the tilted and non-tilted cases.

The general conclusions of this chapter are, therefore, that differential geometry has an increasingly important role in the study of cosmology. By treating the universe as a four-dimensional manifold, we can employ the geometric techniques of topology to impose global symmetry conditions on spacetime and thereby simplify the field equations describing the behaviour of such 'symmetric' solutions. These geometric symmetries then manifest themselves as physical symmetries by requiring that the matter distribution obey certain conditions such as, spherical symmetry, homogeneity, isotropy, etc.

In particular, the success of self-similar symmetries in hydrodynamics, which are the 'physical' representations of conformal symmetries in geometry, has prompted their appliance to the realm of the cosmological fluid. One might reasonably expect a strongly self-gravitating system which evolves in size through many orders of magnitude, either expanding or contracting, to "forget" its initial conditions and eventually become scale-invariant. Thus, the necessary condition for self-similar flow, i.e. that the properties of matter in the system are scale-free, makes such solutions desirable as descriptions of cosmological fluids. It is on this assumption that the present treatise is undertaken.

3. SELF-SIMILAR INHOMOGENEOUS SPACETIMES WITH A COSMOLOGICAL CONSTANT

3.1 Introduction

The use of the cosmological constant, Λ, in the application of general relativity to cosmology has been the subject of much controversy. It is a new independent constant of nature, like the gravitational constant, G, and the speed of light, c, and should be avoided if possible. However, the Einstein equations are made more general by the inclusion of Λ, and it does seem unlikely that Λ is exactly zero. Einstein first introduced the Λ term in order to avoid the prediction of general relativity that the universe was dynamic, which was contradictory to most beliefs at the time. The constant was necessary to obtain stationary (non-expanding) solutions to the field equations and thus model a universe of constant radius. Einstein's argument was that a non-zero Λ requires empty spacetime to be curved, and this is contrary to the spirit of Mach's principle, that there is a connection between the local inertial behaviour of matter and the distant parts of the universe, of which Einstein was a firm believer. [Mach's principle is not based on any quantitative theory, but rather arose from the observation that the local inertial frame, earlier identified by Newton as absolute space, is one relative to which the distant parts of the universe are non-rotating, cf. Mach 1912]. To confuse the issue further, Hubble (1929) discovered that the spectrum of distant galaxies was redshifted by an amount directly proportional to the apparent distance of the galaxy from the Earth and, thus, concluded that the universe was expanding and, therefore, not static. This observation resulted in many authors returning to the simplicity of the original field equations with $\Lambda=0$.

If we are to allow solutions in which the cosmological constant is non-zero, we must address the origin of such a term. The cosmological constant represents an energy density which is usually associated with the vacuum, (cf. Linde 1979). The notion that the vacuum can act as a source of energy provides the basis for many of the current theories of elementary particle physics and early universe cosmology, (e.g. Guth 1981, Brandenberger 1987 and Turner 1987).

Elementary particle theories not only allow for a non-zero vacuum energy density but also strongly suggest that it should have a large value. A universe with a large Λ would be vastly different from the one we actually observe. The energy of the vacuum generates a gravitational field that reveals itself as a change in the geometry of spacetime. Therefore, a large vacuum energy density would have a profound effect on the evolution of the universe. As an example, consider the radius of curvature for an isotropic (Friedmann) universe. Whether the universe is closed (e.g. Einstein static model) or open (e.g. de-Sitter model), the radius of curvature (Hubble radius for an open solution) is given by $R \approx \Lambda^{-1/2}$ (in units of $c=1$). Thus, in either type of solution, if the cosmological constant was positive and large, the universe would be extremely small, obviously, contrary to all observational evidence.

Observational cosmology allows us to place a strict limit on the magnitude of the cosmological constant, (vacuum energy density), at the present epoch. Number counts of galaxies in many different regions of the universe have allowed us to determine the geometry of those regions, which then gives us a direct measure of the effect, if any, the cosmological constant may have on this geometry, (Loh 1987).

The results indicate that the magnitude of the cosmological constant must be smaller than $1/(10^{23}\mathrm{km})^2$, some 122 orders of magnitude smaller than the value predicted on the basis of the standard model of elementary particle physics, (e.g. Cheng and Li 1984). This extremely high prediction of the standard model for the cosmological constant is based on the assumed independence of the free parameters of that model. The discrepancy with observational limits suggests that this assumption is spectacularly wrong. As yet there has been little progress made in reconciling the vanishingly small cosmological constant 'observed' within the context of the standard model.

Cosmologies which contain a large Λ, may still be viable descriptions of the universe at very early epochs. Indeed, such models have proved to be very useful as a source of an accelerated expansion causing the universe to evolve from an anisotropic state to the extremely isotropic and homogeneous universe we observe today, (cf. Wald 1983 and Jensen and Stein-Schabes 1987). Such 'inflationary' scenarios, (Gibbons et al. 1983), are considered attractive because they hold out the hope that the present state of the universe can be explained without the necessity of imposing very special conditions on the initial state of the universe. The large value of Λ required in these inflationary models may be transitory, and could undergo one or more phase transitions, altering its value significantly, possibly bringing it to within observational limits, (Henriksen, Emslie and Wesson 1983, hereafter HEW). The importance of these accelerated expansion models demand that we discuss them in more detail .

HEW have shown that Einstein's equations with a constant vacuum energy density ("cosmological constant") term can be solved by exploiting a self-similarity of the second kind, (e.g. Zel'dovich and

Raizer 1967), to obtain non-trivial analytic solutions. These solutions include non-empty, spherically symmetric, inhomogeneous models which tend to the de-Sitter spacetime at small radial distances and/or large cosmic times. Barrow and Stein-Schabes (1984) considered dust-filled exact inhomogeneous solutions of the Szekeres type (Szekeres 1975) and showed that these solutions tended asymptotically to the de-Sitter spacetime, thus providing a specific example of the cosmic "no-hair" theorems of Hawking and Moss (1982) and Wald (1983). Recently Jensen and Stein-Schabes (1987) have extended these "no-hair" theorems and demonstrated that under quite general conditions *any* inhomogeneous solution with a positive cosmological constant and a non-positive three-curvature of space will approach the de-Sitter solution at late times. The HEW solutions, however, have a positive scalar three-curvature and are thus interesting examples of asymptotic de-Sitter solutions which lie outside the scope of these theorems. In fact, many attempts have been made to extend these cosmic "no-hair" theorems by considering different initial assumptions to those of Jensen and Stein-Schabes (1987). For instance, Ponce de Leon (1987) considered spherically symmetric, $\Lambda>0$, solutions with a positive scalar three-curvature on which he imposed the, physically reasonable, dominant, weak and, more importantly for this discussion, strong energy conditions of Hawking and Ellis (1973), together with the positive pressure criterion, $p>0$ for all time. The solutions obtained were found to be ever-expanding, overcoming the premature recollapse problem, but with an asymptotic behaviour different from de-Sitter. The relevance of the energy conditions within cosmological models will become apparent later, when we discuss the behaviour of our solutions.

The solutions discussed by all these authors are based on a synchronous coordinate system and thus belong to a different class of solutions to that of HEW. On the other hand, Gőtz (1988) investigated solutions in a non-synchronous gauge which had a matter field equivalent to an empty universe with a positive cosmological constant. These solutions do not display the asymptotic time behaviour predicted by the "no-hair" conjecture. Like the HEW solutions they have positive space curvature, but possess plane rather than spherical symmetry. The HEW solutions are, however, non-empty and have the distinctive feature of predicting an inflationary-type negative pressure at early times, making them attractive candidates for describing the evolution of an inhomogeneous, exponentiating, inflationary bubble into a present-day Friedmann universe. The presence of the cosmological constant, which we will take to be positive throughout, is of some considerable interest since this term, in its role as a vacuum energy density, may be strongly linked to the symmetry-breaking phase changes predicted by gauge theories of elementary particle physics (Guth 1981). In addition, the negative *material* pressures which exist at early times in the HEW solutions may be advantageous to the creation of particles in some symmetry-breaking particle theories (Brout, Englert and Gunzig 1978).

HEW discussed in detail an inhomogeneous self-similar spacetime in which, however, both the matter density and spatial sections of the manifold were homogeneous. It is the purpose of the present work to extend this analysis to a more general inhomogeneous similarity solution in which the matter density is *also* inhomogeneous. Solutions with an exponential expansion (i.e. inflationary models) tend to wash out any global inhomogeneities in the spacetime leaving behind a very smooth universe (Guth 1981). However, by introducing an intrinsic inhomogeneity into the spatial sections of the manifold, inhomogeneities may still exist on local scales at the end of the inflationary stage.

We proceed, then, to find a class of spatially, inhomogeneous, spherically symmetric solutions with the prescribed self-symmetry. These solutions are set in a physical context by demanding that they satisfy both the weak and dominant energy conditions for a Type I matter field (Hawking and Ellis 1973). We shall find that by imposing these energy conditions our inhomogeneous solutions cannot always be extended back to arbitrarily small times and/or large distances from the origin. This difficulty is overcome by patching our solution to an isotropic model, enabling us to obtain a global solution.

3.2 Formation of Self-Similar Solutions

The sign convention and notation are the same as in HEW, with c=G=1. [Choosing our units in this way is equivalent to introducing a constant length L_o, a constant energy density ρ_o and a constant mass m_o which satisfy

$$G\rho_o L_o^2 = c^4$$

$$m_o c^2 = \rho_o L_o^3 \quad .$$

In the subsequent discussion we shall measure lengths in units of L_o and suppress this constant]. In the ideal fluid case under discussion, the general form of the field equations,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = -8\pi T_{\mu\nu} \quad , \tag{3.1}$$

can be written as

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = -8\pi[(\rho_m+p_m)u_\mu u_\nu - p_m g_{\mu\nu}] \quad , \tag{3.2}$$

where $G_{\mu\nu}$ is the Einstein tensor, $T_{\mu\nu}$ is the energy-momentum tensor, $R_{\mu\nu}$ is the Ricci tensor, $g_{\mu\nu}$ is the metric tensor, $\mathbb{R}$ is the scalar curvature and Λ is the cosmological constant. The right hand member of equation (3.2) is the form of the energy-momentum for a perfect fluid matter distribution, where ρ_m is the density, p_m is the thermodynamic pressure and u_μ is the velocity four-vector of the fluid.

As in HEW, we identify

$$\Lambda = 8\pi\rho_V = -8\pi p_V \qquad , \tag{3.3}$$

where ρ_V is the energy density of the vacuum, and set, (cf. Henriksen 1982),

$$p = p_m + p_V \qquad ; \qquad \rho = \rho_m + \rho_V \qquad , \tag{3.4}$$

to obtain the field equations in standard form,

$$R_{\mu\nu} - \frac{1}{2}\mathbb{R}g_{\mu\nu} = -8\pi[(\rho+p)u_\mu u_\nu - pg_{\mu\nu}] \qquad . \tag{3.5}$$

We have thus included the cosmological constant on the right hand side of (3.5), in the role of a vacuum energy density. Thus, ρ and p are decoupled into matter and vacuum terms, each with its own separate equation of state: for the vacuum $\rho_V=-p_V$, (e.g. McCrea 1951 and Linde 1979), while the matter equation of state, $p_m=p_m(\rho_m)$, may be taken as the sixth equation that completes the definition of the problem. Rather than specify such an equation of state *a priori*, HEW searched for solutions which were self-similar in some dimensionless variable ξ, i.e. a typical physical variable, f say, can be written as $f(r,t)=g(\xi[r,t])$, with a suitable form of ξ.

Adopting a spherically symmetric metric, viz.,

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$$

$$= e^{\sigma(r,t)}dt^2 - e^{\omega(r,t)}dr^2 - R^2(r,t)d\Omega^2 \quad . \qquad (3.6)$$

with $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$, and a non-synchronous coordinate system which is comoving and such that

$$u_{\mu} = e^{\sigma/2}(1,0,0,0) \quad . \qquad (3.7)$$

Substituting equations (3.6) and (3.7) into the expression for the energy-momentum tensor, $T_{\mu\nu}$, we find that

$$T^{\mu}{}_{\nu} = g^{\mu\tau}T_{\tau\nu} = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & -p & 0 \\ 0 & 0 & 0 & -p \end{bmatrix} \quad . \qquad (3.8)$$

Einstein's equations in the form (3.5) reduce to a set of four partial differential equations, (cf. Zel'dovich and Novikov 1971) viz.,

$$e^{-\omega}\left[\frac{2R_{rr}}{R} - \frac{R_r\omega_r}{R} + \frac{R_r{}^2}{R^2}\right] - e^{-\sigma}\left[\frac{R_t\omega_t}{R} + \frac{R_t{}^2}{R^2}\right] - \frac{1}{R^2} = -8\pi\rho \quad , \qquad (3.9a)$$

$$-e^{-\omega}\left[\frac{R_r\sigma_r}{R} + \frac{R_r{}^2}{R^2}\right] + e^{-\sigma}\left[\frac{2R_{tt}}{R} - \frac{R_t\sigma_t}{R} + \frac{R_t{}^2}{R^2}\right] + \frac{1}{R^2} = -8\pi p \quad , \qquad (3.9b)$$

$$-\frac{e^{-\omega}R_{rr}}{R} + \frac{e^{-\sigma}R_{tt}}{R} - \frac{e^{-\sigma}R_t}{2\,R}\left[\sigma_t - \omega_t\right] - \frac{e^{-\omega}R_r}{2\,R}\left[\sigma_r - \omega_r\right]$$

$$+ \frac{e^{-\sigma}}{4}\left[2\omega_{tt} + \omega_t{}^2 - \omega_t\sigma_t\right] - \frac{e^{-\omega}}{4}\left[2\sigma_{rr} + \sigma_r{}^2 - \sigma_r\omega_r\right] = -8\pi p \quad , \qquad (3.9c)$$

$$2R_{tr} - \sigma_r R_t - \omega_t R_r = 0 \quad , \qquad (3.9d)$$

where subscript 't' denotes $\partial/\partial t$ and subscript 'r' denotes $\partial/\partial r$. Following Misner and Sharp, (1964) and Podurets, (1964), we introduce the function m(r,t), which is defined to be the mass within a comoving radius r and is given by

$$m = \int_0^r 4\pi(R(x,t))^2\rho(x,t)\frac{\partial R}{\partial x}dx = \int_0^r 4\pi R^2\rho R_r dr \quad . \tag{3.10}$$

Then clearly,

$$m_r = 4\pi\rho R^2 R_r \quad . \tag{3.11}$$

We also note that from equations (3.9a) and (3.9d) we find

$$-2m_r = e^{-\omega}\left[2RR_rR_{rr}-RR_r{}^2\omega_r+R_r{}^3\right] - e^{-\sigma}\left[2RR_tR_{tr}-RR_t{}^2\sigma_r+R_rR_t{}^2\right] - R_r \ ,$$

which can then be integrated to obtain

$$m = \frac{1}{2}\left[R + e^{-\sigma}RR_t{}^2 - e^{-\omega}RR_r{}^2\right] \quad . \tag{3.12}$$

In deriving equation (3.12) we have used the fact that $m\to 0$ as $R\to 0$ to remove an arbitrary function of t, assuming that there are no singularities in $e^{-\omega}$ and/or $e^{-\sigma}$. Thus we have, from equations (3.9b) and (3.9d), that

$$m_t = -4\pi R^2 R_t p \quad . \tag{3.13}$$

Now, since the vectorial divergence of the energy-momentum tensor vanishes identically, i.e.

$$T^{\mu\nu}{}_{;\nu} = 0 \quad ,$$

where (;) denotes the covariant derivative, we can obtain the equations expressing the law of conservation of energy and momentum. After some simple but tedious algebra, these equations are most concisely written as

$$\omega_t = \frac{-2\rho_t}{(p+\rho)} - \frac{4R_t}{R} \quad , \tag{3.14a}$$

$$\sigma_r = \frac{-2p_r}{(p+\rho)} \quad . \tag{3.14b}$$

We have thus obtained the Einstein equations in physical form which describe the evolution of a perfect fluid cosmology, i.e. equations (3.11)-(3.14).

The presence of the cosmological constant prevents there being a simple (first kind) self-symmetry, since Λ introduces a fundamental scale. However, by transforming to canonical coordinates t', r', (Bluman and Cole 1974), such that the appropriate self-similar variable is just ξ=t'/r', we can define a similarity symmetry of the second kind. This proves to be always possible, provided the equation of state $p_m(\rho_m)$ is allowed to be found as part of the solution.

This transformation to canonical coordinates is taken in the form

$$dt = e^{\Delta\sigma(t')/2}dt' \quad ; \quad dr = e^{\Delta\omega(r')/2}dr' \quad . \tag{3.15}$$

In these coordinates we may take the similarity variable to be ξ=t'/r' and make the usual self-similar ansatz, (Cahill and Taub 1971, Henriksen and Wesson 1978a), viz.,

$$\begin{aligned} 8\pi\rho_m &= \frac{\eta(\xi)}{r'^2} \quad , & 8\pi p_m &= \frac{P(\xi)}{r'^2} \quad , \\ m_m &= \frac{r'M(\xi)}{2} \quad , & R &= r'(S(\xi) \quad , \\ \sigma &= \sigma(\xi) \quad , & \omega &= \omega(\xi) \quad , \end{aligned} \tag{3.16}$$

where η, P, M, S, σ and ω are all dimensionless functions of the self-similar variable ξ and m_m is the material component of the gravitational mass (3.10).

Substituting (3.16) into the equations (3.11)-(3.14) gives the field equations in the form

$$M' = -PS^2S' \quad , \tag{3.17}$$

$$M - \xi M' = \eta S^2(S-\xi S') \quad , \tag{3.18}$$

$$\sigma' = -\frac{2}{\xi^2(P+\eta)}\frac{d}{d\xi}(\xi^2 P) \quad , \tag{3.19}$$

$$\omega' = \frac{-4S'}{S} - \frac{2\eta'}{(P+\eta)} \quad , \tag{3.20}$$

$$1 - \frac{M}{S} - \frac{8\pi\rho_V}{3}(r')^2S^2 = \exp[-(\omega(\xi)+\Delta\omega(r'))](S-\xi S')^2 - \exp[-(\sigma(\xi)+\Delta\sigma(t'))]S'^2 \quad , \tag{3.21}$$

where (') now denotes the differential $d/d\xi$.

By choosing $8\pi\rho_V = \Lambda > 0$, the only way to maintain the assumed symmetry is to set

$$e^{-\Delta\sigma} = \frac{\Lambda(t')^2}{3} \quad ; \quad \Delta\omega = 0 \quad , \tag{3.22}$$

whence (3.21) is uniquely separated into two equations:

$$1 - \frac{M}{S} = e^{-\omega}(S-\xi S')^2 \quad , \tag{3.23}$$

and

$$S^2 = e^{-\sigma}(\xi S')^2 \quad . \tag{3.24}$$

From equation (3.23), it is evident that, for the self-symmetry to be valid, we have the important global condition $M \leq S$. As HEW demonstrated, the existence of a self-symmetry of the second kind requires, for positive Λ, the similarity variable be given by, (cf. equations (3.15) and (3.22)),

$$\xi = \frac{t'}{r'} = \frac{e^{\lambda t}}{\lambda r} \quad , \tag{3.25}$$

where $\lambda=\sqrt{(\Lambda/3)}$.

The Misner-Sharp-Podurets form of the Einstein field equations then yield *six* independent ordinary differential equations for the dimensionless functions $S(\xi)$, $M(\xi)$, $\eta(\xi)$, $P(\xi)$, $\omega(\xi)$ and $\sigma(\xi)$. Following HEW, we see that equations (3.17) and (3.18) are easily rearranged to give

$$\xi S' = \frac{(\eta S^3-M)}{S^2(P+\eta)} , \qquad (3.26)$$

and

$$\xi M' = -\frac{P}{P+\eta}(\eta S^3-M) . \qquad (3.27)$$

Moreover, equations (3.23) and (3.24) may now be used with (3.26) to express the metric components as

$$e^{\omega/2} = \frac{(M+PS^3)}{S^2(1-M/S)^{1/2}(P+\eta)} , \qquad (3.28)$$

$$e^{\sigma/2} = \frac{\eta S^3-M}{S^3(P+\eta)} . \qquad (3.29)$$

Substituting these equations into the Bianchi identities (3.20) and (3.19) respectively, gives after a tedious calculation,

$$\xi \eta' = \frac{-3(\eta S^3-M)}{S^3} , \qquad (3.30)$$

and

$$\xi P' = \frac{[M+PS^3]^2 - 4PS^3(S-M)]}{2S^3(S-M)} , \qquad (3.31)$$

The ratio of equations (3.27) and (3.26) may be written as

$$P = -\frac{1}{S^2}\frac{dM}{dS} , \qquad (3.32)$$

which may be combined with equation (3.30) to give the integral

$$M = \frac{\eta S^3}{3} + \Delta \qquad . \qquad (3.33)$$

where we have used (3.26) to change the independent variable from ξ to S in (3.30). Here Δ is a constant of integration and again (') denotes $d/d\xi$.

Equation (3.33), expressed in dimensional parameters (cf. ansatz (3.16)), is

$$m_m = \frac{4\pi}{3}\rho_m R^3 + \frac{r\Delta}{2} \qquad , \qquad (3.34)$$

indicating that Δ is a measure of the inhomogeneity of the solution.

HEW have made a detailed study of the uniform solution, (Δ=0), in which the density of matter, and spatial sections of the manifold, are homogeneous. The assumption of uniformity admits an analytic solution of (3.26) to (3.33), given in its most general form by

$$S = K\xi \operatorname{sech}(C\ln\xi + D) \qquad , \qquad (3.35)$$

$$M = K\xi \operatorname{sech}^3(C\ln\xi + D) \qquad , \qquad (3.36)$$

$$\eta = \frac{3}{K^2\xi^2} \qquad , \qquad (3.37)$$

$$P = \frac{[3C\tanh(C\ln\xi+D)-1]}{K^2\xi^2[1-C\tanh(C\ln\xi+D)]} \qquad , \qquad (3.38)$$

$$e^{\sigma/2} = [1-C\tanh(C\ln\xi+D)] \qquad , \qquad (3.39)$$

$$e^{\omega/2} = CK\xi \operatorname{sech}(C\ln\xi+D) \qquad , \qquad (3.40)$$

where C, D and K are three further constants of integration. We will justify this analytic solution in §3.4.A. We note that D can be set equal to zero without loss of generality (through a scaling of the radial coordinate) and that K is a simple homology parameter, (a scaled solution exists involving the variables (S/K), (M/K), ($K^2\eta$), (K^2P), (e^{ω}/K^2) and e^{σ}, cf. (3.26)-(3.33)). The parameter C has a much more interesting significance, however. As $\xi\to\infty$, we have

$$S \longrightarrow 2K\xi^{1-C} \quad , \quad M \longrightarrow 8K\xi^{1-3C} \quad ,$$
$$\eta \longrightarrow \frac{3}{K^2\xi^2} \quad , \quad P \longrightarrow \frac{(3C-1)}{3(1-C)}\eta \quad . \qquad (3.41)$$

The metric then becomes

$$ds^2 = (1-C)^2dt^2 - 4K^2\xi^{2(1-C)}(C^2dr^2 + r^2d\Omega^2) \quad , \qquad (3.42)$$

and we see that $C \gtrless 1$ corresponds to 'closed' and 'open' universe models respectively. In the latter the scale factor increases without bound as $\xi\to\infty$, while in the former it passes through a maximum at $\xi=\xi_{crit}$, say, and approaches zero as $\xi\to\infty$ ($t\to\infty$ and/or $r\to0$; cf. equation (3.25)). Figures 3.1a,b demonstrate the behaviour of the scale factor, S/K, and the pressure, K^2P, for the analytic solutions of HEW for a range of values of the parameter C, illustrating each of the regions, C<1, C=1, C>1. We see from the figures that the closed solutions, (C>1), are characterised by a maximum in the scale factor, S/K, with a corresponding infinity in the pressure, K^2P, (cf. equation (3.38)), at some value of ξ, ξ_{crit} say. (This is a pressure siingularity similar to the one discussed by Barrow et al. (1986), see later). For the 'bounded' solutions, (C=1), both the scale factor and the pressure tend to finite limits as $\xi\to\infty$, ($S/K\longrightarrow2$ and $K^2P\longrightarrow1$, equations (3.35) and (3.38)). In the case of the open solutions, (C<1), the scale factor increases monotonically with the self-similar variable, ξ. However, the behaviour of the pressure is a little more interesting. If $C\leq1/3$ the pressure is

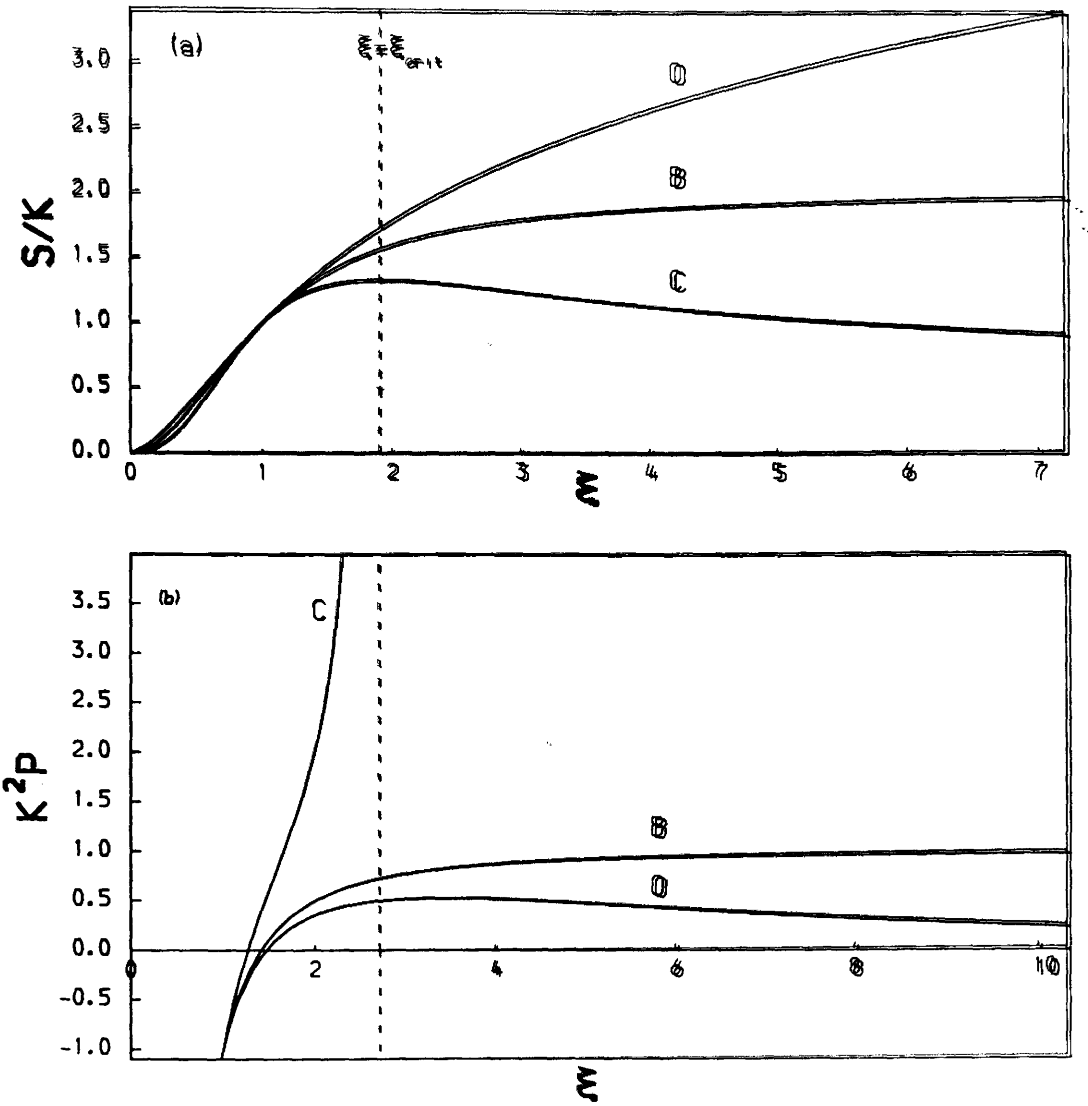


Figure 3.1 Behaviour of (a) the transverse scale factor, S/K and (b) the pressure, K^2P, for the analytic solutions of HEW. The open solutions (O) are characterised by a monotonically increasing scale factor and correspond to solutions for which $C<1$. The bounded solutions (B) have $C=1$ and both the scale factor and pressure tend to constant values as the self-similar variable, $\xi \to \infty$. The closed solutions (C) have $C>1$ and are characterised by a maximum in the transverse scale with a corresponding infinity in the pressure for some finite value of ξ, ξ_{crit} say. (Note that in all solutions the pressure is negative for small ξ).

negative and monotonically increasing for all ξ. If $(1/3)<C<1$ the pressure exhibits a maximum at a positive value of P, for some finite value of ξ, tending to zero as $\xi\to\infty$. We shall consider these solutions in more detail in §3.4.

The coordinate transformation,

$$\bar{r} = 2K\lambda^{(C-1)}r^{C} \quad , \quad \bar{t} = (1-C)t \quad ,$$

reduces (3.42) to the de-Sitter form showing that such open universes evolve into isotropic homogeneous Friedmann models at all r for sufficiently large times.

The distinction between 'open' and 'closed' models is defined solely by the behaviour of $S(\xi)$. This becomes evident when we consider the proper volume and proper circumference corresponding to a fixed time t which, with $R_s=Ke^{\lambda t}$ (the de-Sitter scale factor), are given by

$$V = 4\pi\int_0^\infty e^{\omega/2}R^2dr = \frac{4\pi K^3e^{3\lambda t}}{\lambda^3}\int_{-\infty}^{\infty}\mathrm{sech}^3X dX = \frac{2\pi^2}{\lambda^3}R_s^{\,3} \quad , \tag{3.43}$$

(where $X=C\ln\xi$) and

$$d = 2\int_0^\infty e^{\omega/2}dr = 2\pi R_s \quad . \tag{3.44}$$

Both the proper volume and the proper circumference are thus independent of the parameter C, and finite for finite t so that all solutions of this model are in fact *spatially* closed. The difference between our so-called 'open' and 'closed' models is that in the closed models at a finite time t, we can define a critical surface $r_c=e^{\lambda t}/\lambda\xi_{crit}$ such that the universe is expanding for $r>r_c$ and contracting for $r<r_c$, whereas in open models it is expanding everywhere at all times and no such critical surface exists. The critical surface described here is similar to that found by Coley and Tupper (1983) in their investigation

of a viscous magnetohydrodynamic universe. In their work the critical surface is characterised by a zero volume expansion, $\Theta \equiv u^{\alpha}{}_{;\alpha}=0$. In the case of one of our 'closed' models the volume expansion may be derived from equations (3.35), (3.39) and (3.40) by using the fact that the coordinate system is comoving, i.e. $u_{\alpha}=e^{\sigma/2}(1,0,0,0)$. It is then found that $\Theta=-3\lambda$ for $r<r_c$ and $\Theta=3\lambda$ for $r>r_c$, so that the volume expansion is discontinuous on the critical surface.

It is interesting to investigate the form of the solution in the vicinity of those surfaces $\xi=\xi_{crit}$ corresponding to maxima in S and $e^{\omega/2}$, on which the expansion reverses into a contraction. For models with C>1, ξ_{crit} must, by differentiating equation (3.35), satisfy

$$1 - C\tanh(C\ln\xi_{crit}+D) = 0 \quad .$$

This implies that at the point S'=0, $P\to\infty$ (equation (3.38)). Since S is finite at this value of ξ, the appearance of an infinite pressure is somewhat puzzling. Furthermore, the curvature invariant, defined by $\mathbb{R}=g^{ab}g^{cd}R_{abcd}$, for the metric (3.6), in the vicinity of $\xi=\xi_{crit}$, can be shown to be

$$\mathbb{R} = -\frac{12}{K^4\xi^4r^4(1-C\tanh(C\ln\xi+D)} + \frac{8\Lambda}{K^2\xi^2r^2} + O(\xi-\xi_{crit}) \quad , \qquad (3.45)$$

so that $\mathbb{R}$ is infinite at $\xi=\xi_{crit}$, i.e. on the shell of expansion reversal. This infinite curvature can be shown to be a direct consequence of the vanishing of the g_{oo} term although, as shown by HEW, the 3-curvature of the t=constant hypersurface remains finite. This 4-curvature singularity is not therefore an all-encompassing "crushing singularity", as defined by Marsden and Tipler (1980), but is, rather, analogous to the "pressure" singularity discussed by Barrow et al. (1986). As was discussed by these authors, this type of singularity prevents a Friedmann model from reaching a maximal hypersurface and therefore prevents recollapse. We reject models which display such pressure singularities since at large times, P again becomes negative and the solution therefore develops a wholly undesirable equation of state. This situation is worsened by the fact that the "surface of infinite pressure", $\xi=\xi_{crit}$, propagates along a spacelike geodesic and therefore has a proper velocity faster than the speed of light. The surfaces $\xi=\xi_{crit}$ are static limits in the sense that g_{oo} vanishes on them, so that observers at larger r (smaller ξ) values only see the critical surface cross smaller r values when they themselves are crossed.

3.3 Addition of a Matter Inhomogeneity Parameter Δ

In this section we investigate to what extent the addition of a finite Δ to a homogeneous model can "close" an otherwise "open" (C<1) solution, i.e. create the pressure singularity discussed above. By taking a starting value ξ_s on an analytic ($\Delta=0$) solution, (equations (3.35)-(3.40)), adding an amount Δ to M and integrating the basic equations (3.26) to (3.33) numerically from that point onward, we can find the minimum Δ which, when added in this way, causes $S(\xi)$ to exhibit a turning point. We denote this minimum value, i.e. the value which just closes an open solution, by Δ_c.

The required Δ_c depends not only on the parameters of the model (C, K, D) but also on the starting value ξ_s at which it is introduced. As noted above, the parameter K is a homologous scaling, so that we need only consider the parameter $\Delta_c{}^*=\Delta_c/K$. Thus, the problem reduces to finding $\Delta_c{}^*$ in the form

$$\Delta_c{}^* = \Delta_c{}^*(C;\xi_s) \qquad , \qquad (3.46)$$

such that the resulting model, integrated forward from the point of addition of the Δ_c^*, just passes through a maximum in S.

Figure 3.2 shows the form of the surface given by equation (3.46). Obviously, we find that $\Delta_c^*(1;\xi_s)=0$ (a C=1 universe is already closed). A more surprising result is that $\Delta_c^*(0;\xi_s)$ is also zero, suggesting that a universe with C=0 (for which S increases without bound) is also in some sense 'closed'. To see why this is so, note that the metric for a C=0 solution is given by

$$ds^2 = dt^2 - K^2\xi^2 r^2 d\Omega^2 \qquad , \tag{3.47}$$

at large ξ. Since g_{11} vanishes, we see that this corresponds to an expanding 2-sphere with radius equal to $(K/\lambda)e^{\lambda t}$ and with the proper distance in the r direction equal to zero. Thus, although the universe is open in the sense that the scale factor increases without bound, an r-shell crossing singularity (Barrow et al. 1986) exists where t=constant hypersurfaces are crushed together.

Since Δ_c^* is zero at both C=0 and C=1, each curve $\Delta_c^*(C;\xi_s)$ must exhibit a maximum at some value of C, which in general depends on ξ_s. We find that the required Δ_c^* increases with ξ_s for all C, i.e. that the later in time (or the closer to the origin r=0) that the mass excess is applied, the greater is the Δ_c^* required to force an open solution (which has had, for all r, more time to develop) into the parameter space corresponding to a closed solution.

These "forced" solutions, found by adding a Δ_c^* at a finite ξ_s, cannot, however, be followed back to very small ξ. This is because at S=0, M will be finite (equation (3.33)). This violates the general requirement, based on the assumed self-symmetry (HEW), that, for positive Λ models, S>M. Therefore, in order to extend our solution

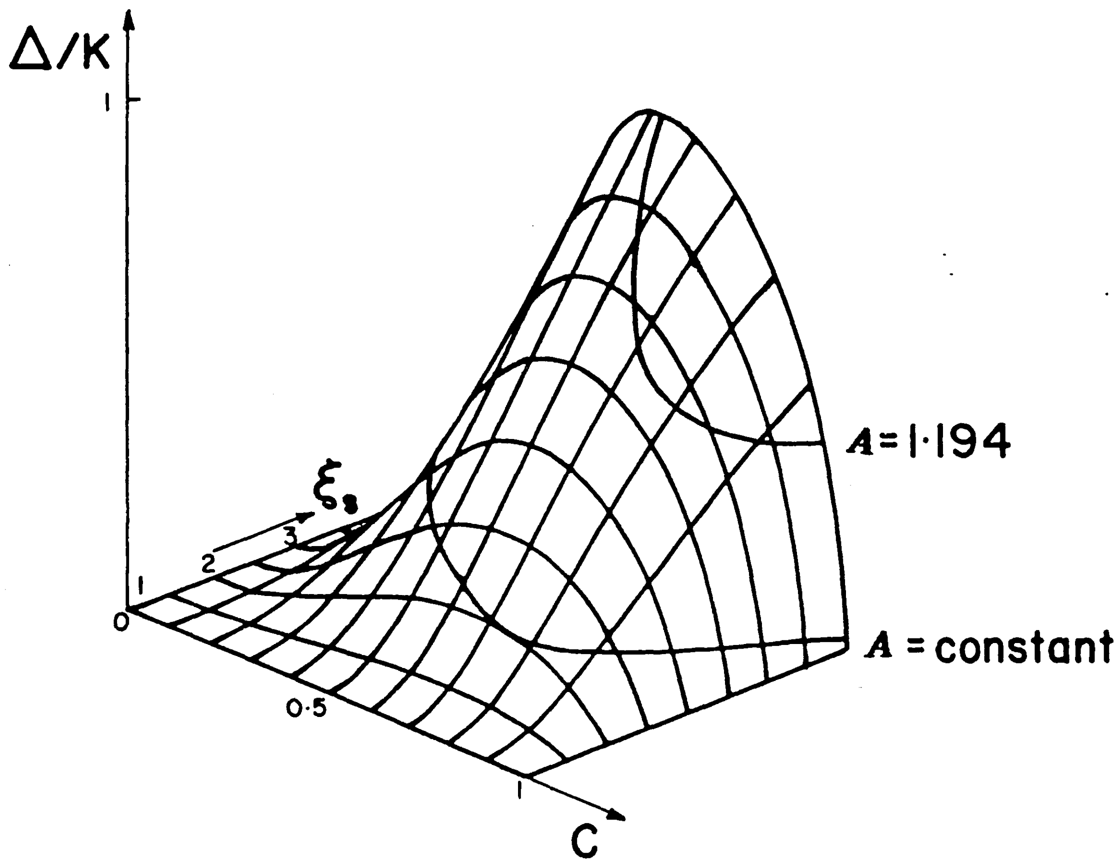


Figure 3.2 The surface $\Delta_C^*(C;\xi_s)$. Solutions for which Δ/K lies above this surface are "closed", i.e. the scale factor $S(\xi)$ goes through a maximum. Note that $\Delta_C^*(1;\xi_s)=0$ and $\Delta_C^*(0;\xi_s)=0$. The two curves which lie on the surface correspond to constant values of the parameter A, see Section 3.4. The lower of these illustrates a typical value of A in the range $1 \leq A \leq \sqrt{2}$, while the upper corresponds to the critical value, $A=1.194$, of the patched solution of Section 3.4.

backwards to early times (or to great distances) we need to patch our solution onto another solution, (one which *can* be extended back to $\xi=0$), at a point where $S>M$ still holds. Such a patch, (onto a singular solution with $S=M$ everywhere), was briefly considered by HEW; we shall discuss such a patching in detail in the next section.

3.4 Inhomogeneous ($\Delta\neq0$) Solutions

A General formulation

The addition of a Δ "suddenly" at a set of surfaces $\xi=\xi_s$ is clearly not a physical development but simply a useful way of introducing the family of solutions which exist in this model. In this section we therefore explore this family of solutions by generalising HEW's solution to a model which includes a global matter inhomogeneity, i.e. to the global inhomogeneous case where $\Delta\neq0$. To make these solutions physically meaningful we will demand that they satisfy both the weak and dominant energy conditions for a Type I matter field, (in the notation of Hawking and Ellis 1973). For such matter fields, the weak energy condition, ($T_{ab}W^aW^b\geq0$ for all timelike vectors W^a), will hold if $\eta\geq0$ and $P+\eta\geq0$. Furthermore, the dominant energy condition, $T_{oo}\geq|T_{ab}|$ for all a and b, such that the energy dominates the other components of T_{ab}, will hold if $\eta\geq|P|$.

We first define an "anisotropy parameter" A through the relation

$$A^2 = \frac{e^{\omega}}{S^2} \equiv \left[\frac{r_{\parallel}}{r_{\perp}}\right]^2 = \frac{(M+PS^3)^2}{S^6(1-M/S)(P+\eta)^2} , \tag{3.48}$$

(cf. equation (3.28)), where $r_{\parallel}$ is measure of the radial separation of two particles on a t=constant hypersurface and $r_{\perp}$ is a measure of their transverse separation. From the fundamental equations (3.26) to

(3.33), it is easily shown that $dA/d\xi$ vanishes, so that the parameter A is a constant (integral) of the solution. Further, from equations (3.35), (3.40) and (3.48) we see that in the limit $\Delta \to 0$ the integral A reduces to the parameter C of the homogeneous case. In what follows we will therefore replace the parameter C with its more general form A, and thus characterise a particular solution by the choice of this new parameter.

Rewriting equation (3.48) as an expression for the dimensionless pressure $P(\xi)$, we obtain

$$P = \frac{[3(M-\Delta)A(1-M/S)^{1/2} - M]}{S^3[1-A(1-M/S)^{1/2}]} , \qquad (3.49)$$

where we have replaced η using the integral (3.33).

Using (3.49), equation (3.32) may be written as

$$\frac{dM}{dS} = \frac{[M-3A(M-\Delta)(1-M/S)^{1/2}]}{S[1-A(1-M/S)^{1/2}]} . \qquad (3.50)$$

and equation (3.29) reduces to

$$e^{\sigma/2} = [1-A(1-M/S)^{1/2}] . \qquad (3.51)$$

Finally, the relationship of the self-similar variable ξ to the physical variables is obtained by writing equation (3.26) in the form

$$\frac{d(\ln S)}{d(\ln \xi)} = 1 - A(1-M/S)^{1/2} , \qquad (3.52)$$

(where we have again used equation (3.49)). It is convenient to use S as the independent variable. All the parameters of our solution are then determined; M numerically from (3.50), P from (3.49), η from the integral (3.33), e^{ω} from the integral (3.48) and e^{σ} from (3.51). Finally, $\xi(S)$ can be determined from equation (3.52). The evolution of the solutions is thus completely determined by our choice of the parameters A and Δ.

Having derived the relevant differential equations for the general inhomogeneous case, it is convenient, now, to re-establish the analytic (Δ=0) solution of HEW. To begin with we notice that for the similarity symmetry to hold we must have $M \leq S$, for a positive cosmological constant (cf. equation (3.23)). Thus, we have that the expression $(1-M/S)$ lies in the range $0 \leq (1-M/S) \leq 1$. Let us introduce a new variable θ, where

$$\tanh^2\theta = (1-M/S) \quad , \tag{3.53}$$

which allows us to express M in terms of the variables S and θ, viz.,

$$M = S\,\mathrm{sech}^2\theta \quad . \tag{3.54}$$

Substituting equations (3.53) and (3.54) into equation (3.50) with Δ=0, we obtain, after some tedious algebra,

$$S = B(e^{\theta/A})\,\mathrm{sech}\theta \quad , \tag{3.55}$$

where B is a constant of integration. Equation (3.52) then gives us

$$\theta = A\ln\xi + D \quad , \tag{3.56}$$

where D is a further constant of integration. Finally, we have from (3.55) that

$$S = K\xi\,\mathrm{sech}\theta \quad , \tag{3.57}$$

where $K=Be^D$ and the remaining parameters take the analytic form (3.35)-(3.40), by substitution into equations (3.54), (3.33), (3.49), (3.51) and (3.48), respectively.

Returning to the inhomogeneous case ($\Delta \neq 0$), we can divide the (A, Δ)-plane into solution types. Let us first consider the case where $\Delta<0$.

B $\underline{\Delta<0}$

It is found that $M \geq \Delta$ for all ξ. It can be seen from equations (3.33) and (3.49) that the weak energy condition will always be satisfied, i.e.

$$\eta \geq 0 \quad , \quad P+\eta \geq 0 \quad .$$

We do, however, introduce a negative dimensionless mass, $M(\xi)$, at small values of the scale factor S (equation (3.33)). This could mean that we reach a stage where the gravitational potential dominates the matter and we have a negative energy density. It is possible to produce a global solution which maintains the similarity symmetry throughout, i.e. $M \leq S$, for a positive cosmological constant, and which are characterised by the choice of the two parameters A and Δ. The negative energy density occurring in these solutions make them physically unappealing and we will not consider them further. When Δ is chosen to be greater than zero we find that the energy conditions place more significant constraints on the solutions.

C $\underline{\Delta>0}$

It is evident from equation (3.52) that all solutions for which $A<1$ can be described as open in the sense of HEW, viz., the scale factor increases monotonically. For these open solutions, we find that M is always finite and therefore we have from equations (3.51), (3.52) and (3.48) that

$$\begin{aligned} e^{\sigma} &\rightarrow (1 - A) \quad , \\ S &\rightarrow B\xi^{(1-A)} \quad , \\ e^{\omega} &\rightarrow A^2 B^2 \xi^{2(1-A)} \quad . \end{aligned} \tag{3.58}$$

The coordinate transformation,

$$\bar{t} = (1-A)t \quad , \qquad \bar{r} = B\lambda^{(A-1)/2} r^{A}$$

reduces the metric (3.6) to the form

$$ds^2 = d\bar{t}^2 - e^{2\lambda\bar{t}}(d\bar{r}^2 + \bar{r}^2 d\Omega^2) \quad , \qquad (3.59)$$

which corresponds to the Robertson-Walker de-Sitter metric. (We have replaced ξ by expression (3.25)). Thus, it is again found that the spacetime approaches a Robertson-Walker de-Sitter metric. When $A>1$, the scale factor is bounded and tends to a finite non-zero value; as $\xi\to\infty$ we find from equations (3.50) and (3.52) that

$$\begin{aligned} S &\longrightarrow \frac{3\Delta}{2}\frac{A^2}{(A^2-1)} \quad , \\ M &\longrightarrow \frac{3\Delta}{2} \quad , \end{aligned} \qquad (3.60)$$

In general, therefore, such solutions tend to an anisotropic static model.

Referring back to Figure 3.2, the condition $A=1$ would define a surface in (C, Δ/K, ξ_s) space. It is not, however, the surface indicated in that figure. To illustrate this, let us select a definite value of ξ_s. The initial values of S, P and η are then defined and the initial value of M may be written as $M_0+\Delta$, where M_0 is given by (3.36). Setting $A=1$ in (3.48), then, leads to a simple quadratic in the parameter Δ,

$$\Delta^2 + \left[2(M_0+P_0S_0{}^3)+S_0{}^5(P_0+\eta_0)^2\right]\Delta + \left[(M_0+P_0S_0{}^3)^2-S_0{}^5(S_0-M_0)(P_0+\eta_0)^2\right] = 0 \qquad (3.61)$$

where M_0, P_0, S_0, η_0 are given by equations (3.35)-(3.38). In Figure 3.3 we plot the resulting cross-section of the surface $A=1$ in the (C, Δ/K) plane. The fact that this falls below the critical surface indicates

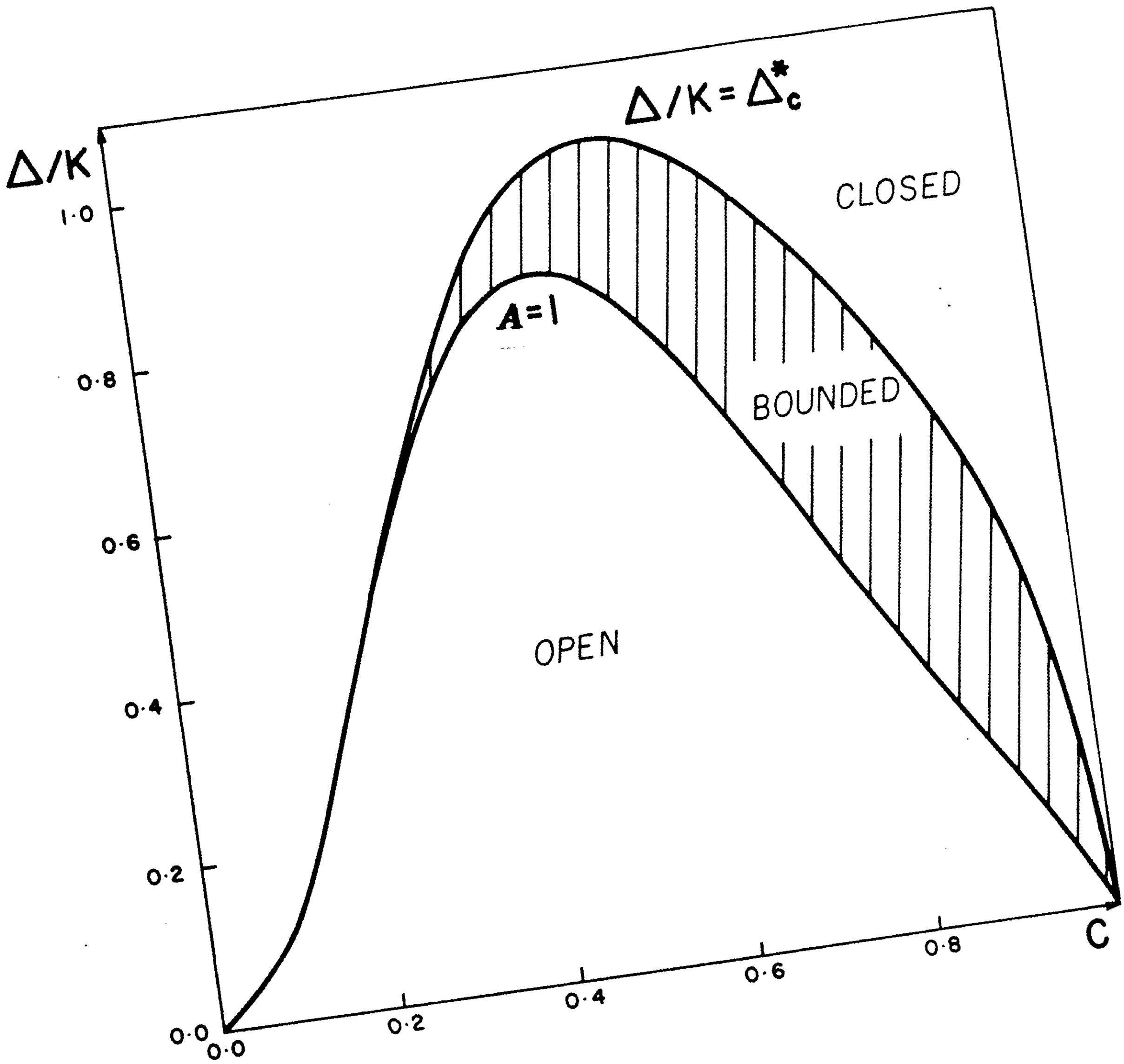


Figure 3.3 The section $\varepsilon_s=5$ of $(C,\Delta/K,\varepsilon_s)$ space showing the two limiting surfaces which divide the solution into three distinct types. The surface on which $\Delta=\Delta_c{}^*$ is <u>not</u> a surface of constant *A*.

that there are two distinct types of solution for $A>1$: (i) Those which are well behaved with the scale factor S tending to some maximum value S_{max} as $\xi\to\infty$ (these we will term *bounded*) and (ii) those which develop the type of singularity discussed in §3.2 with the scale factor reaching a maximum at a finite value of ξ, (ξ_m, say) - these we will term *closed*. For a closed solution, we find that $P\to\infty$ as $\xi\to\xi_m$ and for a bounded solution that $P\to 0$ as $\xi\to\infty$. We can, further, define a limiting case to be one in which P tends to a finite value.

Now the values of the scale factor S and the mass M, as $\xi\to\infty$, are determined by our choice of the parameters Δ and A. Therefore, instead of fixing A and varying Δ to obtain a critical solution (as was done in §3.3), it is equally valid, (and in fact more illuminating), to fix Δ and vary A. Introducing a new variable $y=M/S$ and using the integral (3.33), equation (3.50) reduces to

$$\frac{dy}{dS} = \frac{A(1-y)^{\frac{1}{2}}(2y-3\Delta/S)}{S[A(1-y)^{\frac{1}{2}}-1]} \quad . \tag{3.62}$$

In both closed and bounded solutions the numerator and denominator tend to zero as $\xi\to\infty$. In the closed solutions the denominator tends more quickly to zero, whereas in the bounded solutions it is the numerator which decreases more rapidly. Only in the critical separating case does dy/dS tend to a finite non-zero limit. If we take this critical case and integrate backwards with respect to S we get a series of critical solutions, each defined by the corresponding value of A, as shown, for various A, in Figure 3.4. The family of such curves correspond to the intersections of the A=constant surfaces with the critical surface of Figure 3.2.

Consider equation (3.62) at the point $\xi=\infty$, where both the

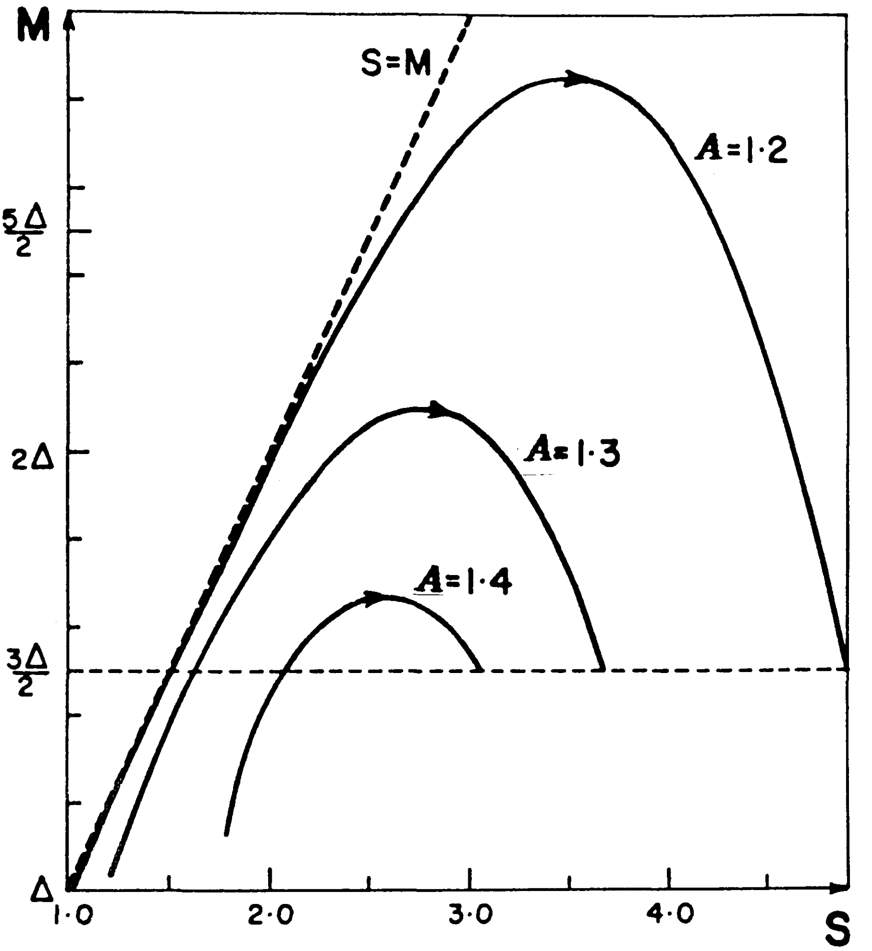


Figure 3.4 The family of critical solutions for which the scale factor S is bounded. Only those solutions which touch the line S=M can be extended indefinitely back to the origin, $\xi=0$. All others violate the symmetry condition, for a positive cosmological constant, that $S \geq M$, as $S \to 0$, see equation (3.33). The arrows indicate the direction of increasing ξ; as $\xi \to \infty$ all the solutions tend to a point on the line $M=3\Delta/2$.

numerator and denominator of the right hand side are zero. Using L'Hôpital's rule, the expression reduces to a quadratic in dy/dS at that point, i.e.

$$\left[\frac{dy}{dS}\right]^2_{\xi=\infty} + \frac{4}{SA^2}\left[\frac{dy}{dS}\right]_{\xi=\infty} + \frac{4}{A^2S^2}\,y_\infty = 0 \quad , \qquad (3.63)$$

and hence

$$\left[\frac{dy}{dS}\right]_{\xi=\infty} = \frac{-2}{SA^2} \pm \frac{2}{SA}\left[\frac{1}{A^2} - y_\infty\right]^{1/2} \quad , \qquad (3.64)$$

where y_∞ is the value of y at $\xi=\infty$. In deriving (3.64) we have used the fact that, for a critical solution, both the numerator and denominator of (3.62) tend to zero as $\xi\to\infty$, i.e.

$$\begin{aligned} 2y - \frac{3\Delta}{S} &\rightarrow 0 \quad , \\ y &\rightarrow \frac{(A^2-1)}{A^2} = y_\infty \quad . \end{aligned} \qquad (3.65)$$

Substituting (3.65) into the discriminant of (3.64) we find that real roots of this quadratic only exist for values of $A\leq\sqrt{2}$. Thus, all solutions with $A>\sqrt{2}$ are necessarily of type (ii), i.e. closed. All of the critical solutions found in §3.3 must correspond to values of A in the range $1\leq A\leq\sqrt{2}$.

Let us now consider the compatibility of the solutions with the weak and dominant energy criteria. It is evident from equations (3.33) and (3.49) that

$$P + \eta = \frac{(2M-3\Delta)}{S^3[1-A(1-M/S)^{1/2}]} \quad , \qquad (3.66)$$

and, therefore, the weak energy condition, $P+\eta\geq0$, is given by $M\geq3\Delta/2$. This is clearly violated by ordinary inhomogeneous solutions at small values of the scale factor S, (i.e. small ξ), since the assumed symmetry requires $M \leq S$. In fact, only the homogeneous solutions, ($\Delta=0$), can be extended indefinitely back to $\xi=0$. There exists, however, a singular solution $M=S=\xi$, studied by HEW, in which the equation of state is not imposed by the self-symmetry but is instead freely chosen. In this solution it can be seen that equation (3.17) requires that $P=-1/S^2$, (3.23) requires $S=\xi$ (an arbitrary multiplicative constant can be absorbed in r), (3.24) gives $e^{\sigma}=1$, (3.18) and (3.19) are identities, and only (3.20) need be solved as

$$\omega' = -\frac{4}{\xi} - \frac{2\eta'}{P+\eta} \quad , \qquad (3.67)$$

for which an equation of state for the matter must be given [$P(\eta)$]. We note that if we use $P=a_s{}^2\eta$ ($a_s{}^2$ is the square of the sound speed) a homogeneous model is obtained. With this equation of state, the solution of (3.67) is $e^{\omega}=B\xi^{\gamma}$, where B is an arbitrary constant and $\gamma=-4a_s{}^2/(1+a_s{}^2)$. The metric (3.6) thus becomes

$$ds^2 = dt^2 - B\xi^{\gamma}dr^2 - r^2\xi^2 d\Omega^2 \quad , \qquad (3.68)$$

which, with a redefinition of the radial coordinate ($dr_*{}^2=r^{-\gamma}dr^2$), reduces to a Kantowski-Sachs metric of closed form, (Kantowski and Sachs 1966).

Let us examine a patch from this singular solution to a general inhomogeneous solution on the hypersurface where M=S. In the Appendix we show that the singular solution to which we patch does exhibit a conformal symmetry, (unlike the general HEW solution), and is therefore, in the light of the recent interest in gauge theories of the early universe, an appealing candidate on physical grounds for the solutions at small ξ.

In the singular M=S solution studied by HEW, the metric coefficients and their derivatives are given by

$$\begin{array}{lll} S = \xi , & S' = 1 , & \\ e^{\sigma} = 1 , & \sigma' = 0 , & \qquad (3.69) \\ e^{\omega} = B\xi^{y} , & \omega' = y/\xi . & \end{array}$$

In our general solution at the point where S=M (=S_0, say) we deduce the following results from equation (3.48), (3.50), (3.51) and (3.52), and when necessary, their derivatives,

$$\begin{array}{lll} S = S_0 , & S' = S_0/\xi_0 , & \\ e^{\sigma} = 1 , & \sigma' = \dfrac{A^2(3\Delta-2S_0)}{S_0\xi_0} , & \qquad (3.70) \\ e^{\omega} = A^2 S_0{}^2 , & \omega' = 2/\xi_0 . & \end{array}$$

We wish to patch at $\xi=\xi_0$ subject to the continuity conditions that both the metric coefficients and their derivatives, i.e. $g_{\mu\nu}$ and $g_{\mu\nu,\lambda}$, be continuous across the patching hypersurface. If these continuity conditions hold, the patch satisfies the junction conditions of Synge (1961). Using the conditions (3.69) and (3.70) at $\xi=\xi_0$ and the continuity relations, we obtain four conditions which must be satisfied on the patching surface, viz.,

$$\begin{array}{ll} S_0 = \xi_0 , & \\ y = 2 , & \\ B = A^2 , & \qquad (3.71) \\ S_0 = 3\Delta/2 . & \end{array}$$

For a given A then this patch selects a particular solution, namely that for which the point S=M occurs for S=M=3Δ/2. The condition y=2 implies that $a_s^2=(-1/3)$ which produces an isotropic form of the singular solution, (cf. equation (3.68)). Therefore, the requirement for continuity across the patching hypersurface constrains us to patch on the line $P+\eta=0$ if the symmetry is to be maintained across the patch. This patch is more rigorous than the one discussed briefly by HEW, where the authors only matched the metric coefficients and not their derivatives. In our patch we find that the discontinuity in the metric derivatives is replaced by a more physical discontinuity, i.e. one in the density across the interface. Furthermore, the continuity conditions mean that we have no freedom in the choice of the equation of state for the singular solution.

Thus, for a given A, a unique solution is then derived apart from an arbitrary scaling due to the particular choice of Δ. Examples are shown in Figure 3.5. Due to the severe restrictions imposed by the symmetry and the continuity conditions this patch selects a unique solution of the type shown in Figure 3.4. This solution corresponds to the upper curve superimposed on the critical surface (3.46) of Figure 3.2. The value of A for which this critical solution occurs is found to be $A=1.194$, and is independent of the choice of Δ.

The dominant energy condition ($\eta \geq P$) is clearly violated by the closed solutions discussed above, since $P \to \infty$ for a finite value of ξ. In fact, as discussed by Barrow et al. (1986), a general property of solutions with pressure singularities is that they violate the dominant energy criterion and therefore are physically unacceptable. In the bounded and open cases the situation is a little more complex. We see from (3.49) that, since the largest value possible for the term $(1-M/S)^{1/2}$ is unity, the pressure is always negative for $A \leq 1/3$. If we also consider the dominant energy criterion, which can be expressed as

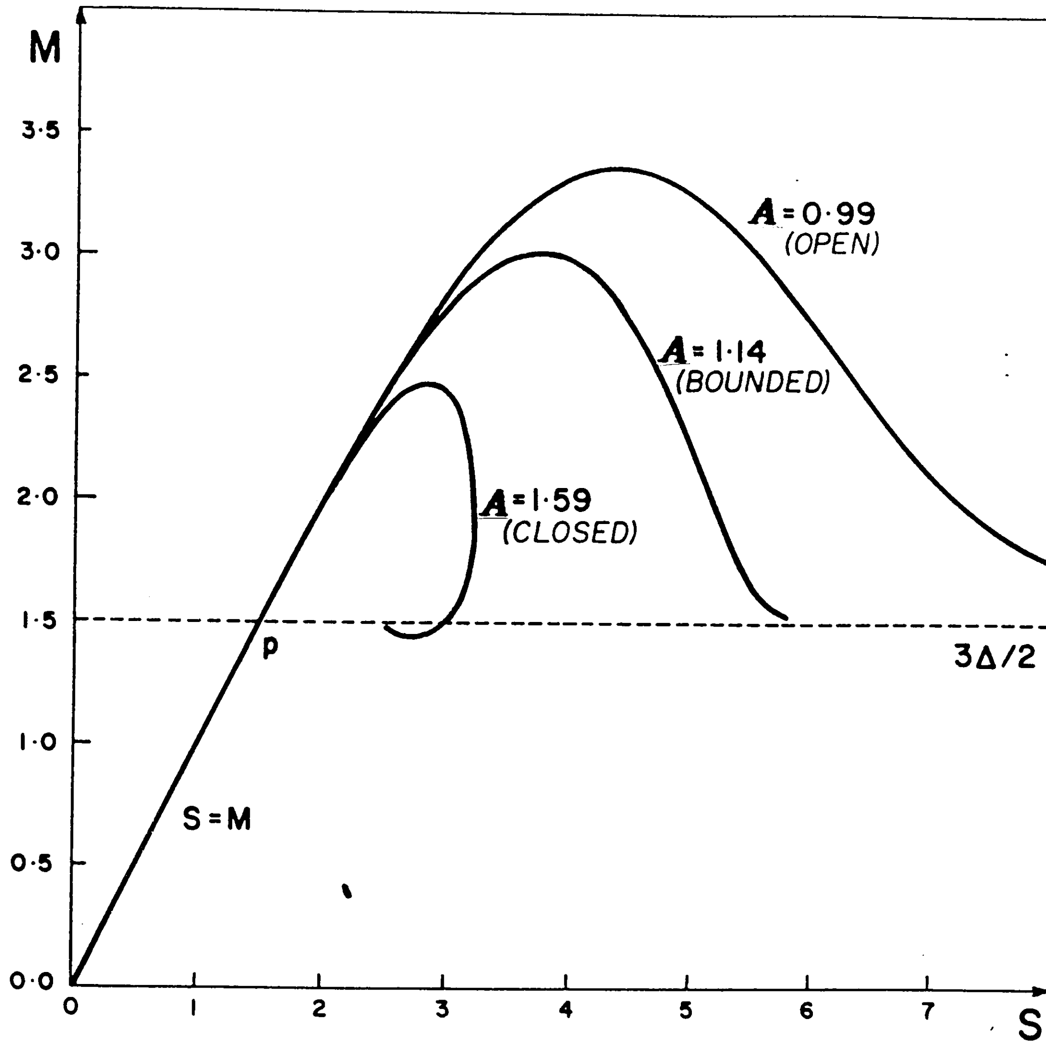


Figure 3.5 The diagram shows the form of the patched solutions for three different values of A, one from each of the regions discussed in the text. The dotted line is given by $M=3\Delta/2$ (for convenience we have chosen $\Delta=1$) and the point p is where these solutions are patched to the singular solution $S=M$ and thus continued to the origin.

η-P$\geq$0, we have, from equations (3.49) and (3.33),

$$\eta - P = \frac{(4M-3\Delta) - 6(M-\Delta)A(1-M/S)^{1/2}}{S^3[1-A(1-M/S)^{1/2}]} \quad . \tag{3.72}$$

From this expression we note that (η-P) is only (definitely) non-negative for all ξ if $A\leq2/3$, since $M\geq3\Delta/2$ for all ξ. Further, it is found numerically that solutions for which $A>2/3$ violate the dominant energy criterion at finite times although the pressure remains finite. The physically appealing solutions therefore lie in the range $(1/3)\leq A\leq(2/3)$. These restrictions can be seen most clearly in the uniform solutions, (Δ=0), of HEW, see Figures 3.1a,b, (remember that the parameter A reduces to the parameter C in the homogeneous case, Δ=0). From equation (3.38), with D set equal to zero, we see that, $P\leq0$ for all ξ when $C\leq(1/3)$. If we use equation (3.37) we obtain

$$\eta(\xi) - P(\xi) = \frac{2[2-3C\tanh(C\ln\xi)]}{K^2\xi^2[1-C\tanh(C\ln\xi)]} \quad . \tag{3.73}$$

From this equation we see that with $C>(2/3)$, $\eta-P<0$ for all finite values of ξ, thus violating the dominant energy condition. The analytic solutions of HEW therefore only satisfy the dominant and weak energy conditions between the limits $(1/3)\leq C\leq(2/3)$. However, all solutions (homogeneous or not) violate the strong energy condition, $(T_{ab}-\frac{1}{2}Tg_{ab})W^aW^b\geq0$ for all timelike W^a (Hawking and Ellis 1973), or $P+\eta\geq0$ and $3P+\eta\geq0$, due to the large negative pressures at early times.

Finally, we note that the scalar three-curvature of space in the homogeneous model is given by (cf. HEW equation (52)),

$${}^{(3)}R = \frac{6\lambda^2}{K^2e^{2\lambda t}} \quad , \tag{3.74}$$

which is positive for all time. We have, therefore, extended the asymptotically de-Sitter solutions of Jensen and Stein-Schabes (1987) to include a solution a spatial hypersurface of positive scalar curvature. Barrow (1988) pointed out that to complete the cosmic "no-hair" theorems of Jensen and Stein-Schabes (1987) and Wald (1983), we need to know the general asymptotic state of anisotropic and inhomogeneous universes with a positive curvature. Ponce de Leon (1987) considered examples of such universes which contained a positive cosmological constant and obeyed all of the energy conditions of Hawking and Ellis (1973). However, the solutions found do not tend asymptotically to de-Sitter solutions and we agree with the author that this may be due to the restriction of positive pressure. Thus we have demonstrated the existence of a new anisotropic and inhomogeneous class of solutions which exhibits an inflationary stage and which overcomes the premature recollapse problem described by Barrow (1987).

3.5 Conclusions

In this chapter we have presented a series of spherically symmetric spatially-inhomogeneous solutions of the Einstein field equations which admit a self-similar symmetry. These non-empty solutions contain a constant vacuum energy density (positive "cosmological constant") and so the similarity is necessarily of the second kind (Zel'dovich and Raizer 1967; Henriksen et al. 1983).

A class of numerical solutions has been found for an inhomogeneous matter energy density in which solutions are characterised by an anisotropy parameter A. These solutions are found to fall into two distinct categories. Solutions with $A<1$ are well behaved and tend to an inhomogeneous de-Sitter universe, in agreement with recent work by Jensen and Stein-Schabes (1987). Solutions with $A>\sqrt{2}$ develop an unphysical pressure singularity, of the type discussed by Barrow et al. (1986). When A is in the range $1\leq A\leq\sqrt{2}$ the situation is more complicated. Some of the solutions develop the pressure singularity, while others tend monotonically to anisotropic static spacetimes. We regard solutions of the latter two types to be of little relevance to the real universe and therefore we have concentrated on the asymptotically de-Sitter solutions.

Cosmologies which display this kind of asymptotic time behaviour are of great interest in light of the recent work on the cosmological "no-hair" theorems (cf. Wald 1983; Jensen and Stein-Schabes 1987). Our solutions, which have a positive scalar three-curvature, provide specific examples of inflationary-like universes which lie outwith the scope of these theorems, thus increasing the number of classes of asymptotically de-Sitter solutions.

All of the self-similar inhomogeneous solutions fail to meet the weak energy criterion of Hawking and Ellis (1973) at early times. We were thus forced to "patch" our solutions onto a singular self-symmetric solution in which the equation of state could be freely imposed, (i.e. the S=M solution of HEW). This allowed us to extend our solutions back to the origin in ξ-space. The requirements for continuity across the patching hypersurface, the continuity of the metric coefficients and their first derivatives, require the patch to be onto the isotropic form of this singular solution where the equation of state is given by $P=-\eta/3$. The continuity requirements also demand the existence of a *unique* "critical" solution, i.e. a solution which is just closed. This solution ($A=A_{crit}$) separates the solutions into those which are closed ($A>A_{crit}$) and those which are bounded ($1\leq A\leq A_{crit}$) as described above. For $A<1$ the solutions are open. In other words, there is a critical "amount of anisotropy" allowed if the solution is to satisfy the physical conditions imposed by Barrow et al. (1986) that there must not develop a pressure singularity.

These allowable patched solutions describe the transformation from a vacuum dominated universe to a non-empty asymptotically de-Sitter universe and are therefore of great physical interest. Whether or not they can be reconciled with Grand Unified Field Theories describing the early universe from the viewpoint of particle physics is at present a fascinating possibility.

4. SELF-SIMILAR IMPERFECT FLUID COSMOLOGIES

4.1 Introduction

In general relativity theory, cosmological models, stellar models and models of other astrophysical matter distributions are usually constructed under the assumption that the matter is an idealised perfect fluid. While this assumption may be a good approximation to the actual matter content of the Universe at the present epoch, effects such as viscosity, heat conduction, rotation and magnetic fields may not be negligible at earlier epochs. In fact, it is rather more probable, from a statistical point of view, that the Universe began in a far from symmetric state and evolved through some dynamical processes to become the homogeneous and isotropic cosmology we observe today. In this chapter we will investigate the problem of obtaining exact solutions to Einstein's field equations for a viscous (imperfect) fluid which displays a homothetic symmetry, i.e. exhibits a self-similarity of the first kind, (cf. Cahill and Taub, 1971, Henriksen and Wesson, 1978a, 1978b, Bicknell and Henriksen, 1978a, 1978b, and Henriksen, 1982).

The importance of treating the Universe as an imperfect fluid, at least for early epochs, is evidenced by the fact that several authors have made attempts to find exact solutions of the field equations by considering a non-ideal fluid in isotropic as well as anisotropic cosmological models, (cf. for example, Misner 1968, Weinberg 1971, Tupper 1981, Coley and Tupper 1983, 1984, 1985). An imperfect fluid is a fluid in which the processes of energy dissipation are non-negligible (Landau and Lifshitz, 1959). Energy dissipation in a moving fluid may be caused by processes such as internal friction due to any viscous stresses which may be present, thermal conductivity due to heat exchange between different parts of the fluid, rotation of the fluid (Batakis and Cohen 1975) or the presence of magnetic fields within the system (Coley and Tupper 1983). We will only consider the effect of viscous stresses on the fluid, neglecting the processes of heat conduction, rotation and magnetic fields.

4.2 Effect of Viscous Stresses on a Cosmological Fluid

The kinematics of a moving fluid can be described adequately by two fundamental equations; namely, the equation of continuity, expressing the conservation of matter, and Euler's equation, which describes the motion of a volume element of fluid, (Landau and Lifshitz 1959 and Symon 1971). For an ideal fluid, which is free from any body forces, such as gravity, these equations can be written as

$$(\partial\rho/\partial t) + \boldsymbol{\nabla}.(\rho\mathbf{v}) = 0 \qquad (4.1)$$

and

$$\partial\mathbf{v}/\partial t + (\mathbf{v}.\boldsymbol{\nabla})\mathbf{v} + \boldsymbol{\nabla}p/\rho = 0 \quad , \qquad (4.2)$$

respectively, where ρ is the fluid density, $\mathbf{v}$ is the velocity of the fluid at a given point and p is the thermodynamic pressure of the fluid. To determine the solution of these equations completely we must be able to express the pressure in terms of the independent thermodynamical quantities of the system, such as density or temperature, i.e. we must be able to specify an equation of state – $p(\rho)$, say.

If we now consider a more general fluid in which the processes of viscous friction cannot be neglected, we find that the motion of adjacent layers of fluid past each other is resisted by a shearing force which tends to reduce their relative velocity. The equations of motion, (4.2), which incorporate all the forces acting on the fluid, have to be modified appropriately to take account of this friction due to viscosity. (Note that the equation of continuity, (4.1), is equally valid for any fluid, viscous or otherwise). We do this by introducing the stress tensor, $\underline{\boldsymbol{\sigma}}$, which gives the part of the momentum flux that is not due to the direct transfer of momentum with the mass of moving fluid, and as such, consists of two components. One is due to the momentum transfer caused by the hydrostatic pressure forces and the other to the irreversible "viscous" transfer of momentum in the fluid. We can also show, using a momentum argument in component form, that $\underline{\boldsymbol{\sigma}}$ is symmetric, (Landau and Lifshitz, 1959). Equation (4.2) then generalises to,

$$\partial \rho/\partial t + \boldsymbol{\nabla}.(\rho \mathbf{v}) + \partial \mathbf{v}/\partial t + \mathbf{v}.\boldsymbol{\nabla}\mathbf{v} + \boldsymbol{\nabla}.\underline{\boldsymbol{\sigma}}/\rho = 0 \quad , \tag{4.3}$$

where we have introduced the continuity equation (4.1) into this expression. The reason for writing (4.3) in this form will become apparent when we generalise this expression to its covariant form, in §4.3. This equation, together with the equation of continuity, (4.1), determines the motion of the medium when the stress tensor, $\underline{\boldsymbol{\sigma}}$, is given. The stress at any point in the fluid may be a function of the density and temperature, of the relative positions of the elements near the point in question and perhaps also of the previous history of the medium. In a viscous fluid, the stress tensor will be expected to depend on the velocity gradients in the fluid. This is consistent with the dimensional arguments of Landau and Lifshitz, (1959). In view of the covariant representation of the stress tensor that is to follow in the next section we will assume that the relation between $\underline{\boldsymbol{\sigma}}$ and the velocity gradients must not depend on the orientation of the coordinate system. We can guarantee that this will be so by expressing the relation in a vector form that does not refer explicitly to components. The dyad

$$\nabla\mathbf{v} = \begin{bmatrix} \frac{\partial v_x}{\partial x} & \frac{\partial v_y}{\partial x} & \frac{\partial v_z}{\partial x} \\ \frac{\partial v_x}{\partial y} & \frac{\partial v_y}{\partial y} & \frac{\partial v_z}{\partial y} \\ \frac{\partial v_x}{\partial z} & \frac{\partial v_y}{\partial z} & \frac{\partial v_z}{\partial z} \end{bmatrix} \tag{4.4}$$

has as its components the nine possible derivatives of the components of $\mathbf{v}$ with respect to x, y and z. Hence we must relate $\underline{\boldsymbol{\sigma}}$ to $\nabla\mathbf{v}$. The dyad, (4.4), is not symmetric but can be separated into a symmetric and an antisymmetric part:

$$\nabla\mathbf{v} = (\nabla\mathbf{v})_s + (\nabla\mathbf{v})_a \quad , \tag{4.5a}$$

$$(\nabla\mathbf{v})_s = [\nabla\mathbf{v} + (\nabla\mathbf{v})^T]/2 \quad , \tag{4.5b}$$

$$(\nabla\mathbf{v})_a = [\nabla\mathbf{v} - (\nabla\mathbf{v})^T]/2 \quad , \tag{4.5c}$$

where $(\nabla\mathbf{v})^T$ = *transpose*($\nabla\mathbf{v}$), i.e. $(\partial v_i/\partial x^k)^T = \partial v_k/\partial x^i$. We can, further, relate the antisymmetric part above to a vector

$$\boldsymbol{\omega} = (\nabla\times\mathbf{v})/2 \quad ,$$

such that, for any vector d$\mathbf{r}$

$$(\nabla\mathbf{v})_a \cdot d\mathbf{r} = \boldsymbol{\omega}\times d\mathbf{r} \quad .$$

If d$\mathbf{r}$ is the vector from a given point, *A*, to any nearby point, *B*, we see that the tensor $(\nabla\mathbf{v})_a$ selects those parts of the velocity differences between A and B which correspond to a rotation of the fluid around A with angular velocity $\boldsymbol{\omega}$. Since no viscous forces will be associated with a pure rotation of the fluid, the viscous forces must be expressible in terms of the tensor $(\nabla\mathbf{v})_s$. $\underline{\sigma}$ is also symmetric and so we would like to express a general linear relation between $\underline{\sigma}$ and $(\nabla\mathbf{v})_s$ which is independent of the coordinate system. To do this we note that $(\nabla\mathbf{v})_s$ can be decomposed into a constant tensor and a traceless symmetric tensor in the following way:

$$(\nabla\mathbf{v})_s = (\nabla\mathbf{v})_c + (\nabla\mathbf{v})_{ts} \quad , \tag{4.6a}$$

$$(\nabla\mathbf{v})_c = {}^1/_3[\mathrm{Tr}(\nabla\mathbf{v})_s]\underline{1} = {}^1/_3(\nabla.\mathbf{v})\underline{1} \quad , \tag{4.6b}$$

$$(\nabla\mathbf{v})_{ts} = (\nabla\mathbf{v})_s - {}^1/_3(\nabla.\mathbf{v})\underline{1} \quad , \tag{4.6c}$$

where $\underline{1}$ is the identity tensor. This decomposition is independent of the coordinate system since the trace is an invariant scalar quantity. It can be shown that the tensor $(\nabla\mathbf{v})_c$ measures the rate of expansion or contraction of the fluid whereas the tensor $(\nabla\mathbf{v})_{ts}$ specifies the way in which the fluid is being sheared. We are therefore free to set

$$\underline{\sigma} = -2\eta(\nabla\mathbf{v})_{ts} - \zeta\nabla.\mathbf{v}\underline{1} \quad , \tag{4.7}$$

with a coefficient η, called the *dynamic* or *shear* viscosity, which characterises the viscous resistance to shear, and a coefficient ζ, called the *bulk* viscosity, which characterises a viscous resistance, if any, to expansion and contraction. To the viscous stress due to velocity gradients, given by (4.7), must be added a hydrostatic pressure, p, which may also be present and which depends on the density, temperature and composition of the fluid. Thus we have

$$\underline{\sigma} = p\underline{1} - \zeta\nabla.\mathbf{v}\underline{1} - 2\eta(\nabla\mathbf{v})_{ts} \tag{4.8}$$
$$= [p - \zeta\nabla.\mathbf{v}]\underline{1} - \eta[\nabla\mathbf{v} + (\nabla\mathbf{v})^T - ({}^2/_3)\nabla.\mathbf{v}\underline{1}] \quad .$$

We see from equation (4.8) that we can define an "effective pressure", Þ, by combining the hydrostatic pressure with the bulk viscosity term, viz.

$$\text{Þ} = p - \zeta\nabla.\mathbf{v} \quad . \tag{4.9}$$

Note that the effective pressure reduces to the hydrostatic pressure in the limit of vanishing bulk viscosity, $\zeta \to 0$. Before we proceed to investigate self-similar viscous solutions we shall discuss the role of viscosity, both bulk and shear, in the cosmological regime.

The role of bulk viscosity seems to be significant for the evolution of the cosmological fluid, at least during the early stages of the Universe. From the macroscopic point of view, the existence of bulk viscosity is equivalent to the existence of slow processes which restore equilibrium states, (Landau and Lifshitz, 1959). The general criterion for non-zero bulk viscosity was given by Weinberg (1971) while attempting to explain the high dimensionless entropy per baryon associated with the microwave background by taking into account the action of dissipative processes in the early Universe. Weinberg pointed out that bulk viscosity may be of importance when considering either a simple gas at the temperatures between extreme relativistic and non-relativistic limits, (the bulk viscosity being negligible in either of these limiting cases), or a fluid composed of a mixture of highly relativistic and non-relativistic particles. Explicit examples of thermodynamic systems with non-negligible bulk viscosity are given by Anderson (1969) and Israel and Vardalas (1970).

Several authors, Diósi et al. (1984), Waga et al. (1986) and Bernstein (1987), have raised the possibility that bulk viscosity can also be the driving force of the accelerated expansion associated with inflation in the early Universe. This proposal relies on the fact that the effect of bulk viscosity in an expanding universe is to decrease the value of the pressure, see equation (4.8). These authors have suggested that bulk viscosity arising around the time of a grand unified theory (GUT) phase transition can, in fact, lead to a negative pressure thereby driving inflation. However, Pacher et al. (1987) have shown that, at least in the case of weakly-interacting particles, the associated bulk viscosity cannot make the pressure negative, excluding any form of accelerated expansion, ($\partial^2 R/\partial t^2 > 0$, where R is the cosmic scale factor). For a gas of such particles the bulk viscosity arises due to the incomplete equilibrium of the relativistic and non-relativistic components and it can be shown that the pressure can never attain a negative value, see Pacher et al., (1987).

The effect of bulk viscosity in isotropic cosmological models, (the assumption of isotropy means that such models are automatically shear-free), has also been discussed by Nightingale (1973) and Heller et al. (1973). Nightingale (1973) investigated the form of the bulk coefficient derived via a relativistic Boltzmann equation and found that bulk viscosity cannot be responsible for the high dimensionless entropy per baryon of the Universe associated with the microwave background, in agreement with Weinberg (1971). Heller et al. (1973), on the other hand, considered the effect of a constant coefficient of bulk viscosity on an isotropic cosmological model and solutions were given for dust-filled universes ($p=0$) and radiative universes ($p=\rho/3$). The assumption of a constant bulk viscosity is somewhat unrealistic, since this coefficient is automatically a function of cosmic time through the dependence on temperature and pressure. However, this work together with the many investigations cited above highlight the importance of understanding the role of bulk viscosity in the evolution of the Universe.

If we now allow for the presence of anisotropies in cosmological models, then the dissipative effects of shearing motions become important during the early stages of cosmic evolution. The effect of shear can be studied independently from the bulk viscosity since we have shown that the bulk can be 'absorbed' into the pressure of the system. The presence of anisotropy, or shear, in the Universe can be attributed to two possible causes. Firstly, one cannot exclude the possibility that the Universe emerged from the Planck era with a high degree of anisotropy, dubbed *primordial* anisotropy by Barrow and Carr, (1977), in which case the shear must be fed into the initial conditions and one would expect that for sufficiently early times the shear would dominate the evolution of the Universe. The other possible "source" of anisotropy is the presence of inhomogeneities which would tend to generate shearing motions due to the tidal stresses caused by the density perturbations existing in such an inhomogeneous universe, (Liang, 1974, Barrow, 1977). For our purposes we shall assume the anisotropy to be an inherent property of the Universe and will not address its origin.

Barrow and Carr, (1977), found that the effect of the anisotropy generated by inhomogeneities is rather small when compared with the effect of primordial shear. This work, together with Carr and Barrow, (1979), considered the effect of shear on the production of primordial black holes, extending the work of Lin et al. (1976) and Bicknell and Henriksen (1978a,b). We will discuss primordial black hole production in a viscous universe in more detail in Chapter 5.

As was the case with bulk viscosity, the effect of shear on the cosmic evolution has been subject to numerous independent studies. Heckmann and Schücking (1958) found the existence of the singularity in the extremely successful Friedmann-Robertson-Walker (FRW) cosmologies very unsatisfactory and attempted to produce finitely oscillating solutions which avoided such a singular event by modifying the postulate of isotropy, thereby introducing rotation and shear. However, their model also introduced closed time-like curves making it a physically unappealing solution. Narlikar (1963) considered a similar problem within the realm of Newtonian cosmology and he too found that the singularity could not be prevented simply by the introduction of anisotropy in the form of shear and rotation.

It was not until the discovery of the microwave background radiation in 1965 by Penzias and Wilson that the first real measurements of the large scale isotropy of the Universe could be attempted. The accuracy of the observations put strict limits on the amount of anisotropy in the Universe at the present time. Misner (1968) showed that the part of the temperature anisotropy in the microwave radiation at the present epoch which is due to primordial anisotropies is extremely small. He suggested that the large anisotropies die away very rapidly due primarily to neutrino viscosity in the early Universe. It is also feasible that the shear energy may be dissipated by the action of gravitational radiation, Papadopoulos and Esposito, (1985).

Many authors, in the last twenty years or so, have introduced dissipative processes into their cosmological models in an attempt to investigate the initial singularity, (Demiański and Grischuk, 1972), to explain cosmological observations, (Saunders 1969, Batakis and Cohen 1975, Goicoechea and Sanz 1984), to study the effect of shear in inflationary universes, (Steigman and Turner 1983, Martinez-Gonzalez and Jones 1986) or to provide detailed studies of anisotropic models and dissipative processes, (Matzner and Misner 1972, Matzner 1972, Johri 1977, Sanz 1983, Papadopoulos and Sanz 1985, Banerjee et al. 1986).

Recently an interesting series of papers by Coley and Tupper, (1983, 1984 and 1985), have concluded that FRW cosmological models may represent physically viable solutions of the field equations for a viscous magnetohydrodynamic (VMHD) fluid. In general, it should be noted that different observers moving relative to each other will give different interpretations to the material content of the Universe. However, in Coley and Tupper, (1984), the interpretation of the material content as a VMHD fluid is given, not by another observer, but by the same set of hypersurface-orthogonal preferred observers who may also interpret the material content as a perfect fluid. Thus they concluded that a spacetime, such as that of the standard FRW models, can correspond to two distinguishable, viable solutions and that the interpretation of the matter distribution corresponding to this spacetime is not unique. Such an analysis suggests that the question of the eventual behaviour of the expansion of the matter content of the Universe cannot be decided merely by an accurate determination of the density of matter, visible and invisible, in the Universe. We also need to determine the qualitative nature of the matter content by investigating the effects of dissipative processes in the Universe.

4.3 General Formalism of a Viscous Cosmology

In this section we shall develop the problem of obtaining exact solutions to the field equations for an anisotropic (viscous) matter distribution. The notation will follow that of Misner, Thorne and Wheeler (1973), although we shall adopt the sign convention introduced in the previous chapters of this thesis. We will work in geometric units (G=c=1).

We will again choose a spherically symmetric spacetime with line element,

$$ds^2 = e^{\alpha(r,t)}dt^2 - e^{\beta(r,t)}dr^2 - R^2(r,t)d\Omega^2 \quad , \qquad (4.10)$$

where, for convenience of notation, the metric coefficients g_{00} and g_{11} are written as e^{α} and e^{β}, respectively, and the self-similar energy density will be denoted by ϵ, leaving the symbols σ and η free to denote the shear and dynamic viscosity, respectively.

Having derived the perfect fluid equations in Chapter 3, (§3.2), we now wish to consider the equivalent equations for an imperfect fluid matter distribution with the same non-synchronous, comoving coordinate system. The generalisation of the previous section leads to the covariant expression of the shear tensor, i.e. the traceless part of (4.7), given by

$$\sigma'_{\mu\nu} = \frac{1}{2}\left[u_{\mu;\tau}P^{\tau}{}_{\nu} + u_{\nu;\tau}P^{\tau}{}_{\mu}\right] - ({}^{1}/{}_{3})\Theta P_{\mu\nu} \quad , \qquad (4.11)$$

(cf. Misner, Thorne and Wheeler, 1973), where the prime signifies that we are only considering the shear component of the stress tensor, $\underline{\sigma}$. $\Theta\equiv u^{\mu}{}_{;\mu}$, is the volume expansion of the fluid world lines, and

$P_{\mu\nu}=g_{\mu\nu}-u_{\mu}u_{\nu}$ is a projection tensor which projects any change in the fluid four-velocity onto a hypersurface orthogonal to **u**, (Hawking and Ellis, 1973). $P_{\mu\nu}$ is the generalisation of the three-dimensional identity tensor, $\underline{\mathbf{1}}$, of equations (4.6) in four dimensions, and has only spatial components. We may regard $P_{\mu\nu}$ as the metric in the spacelike hypersurface orthogonal to **u**. In obtaining this form for the shear tensor, we have followed the method of minimal coupling which can be used to extend the concept of shear from its flat-space definition, to the general curved space-time of general relativity. We simply replace all partial derivatives by covariant derivatives and allow for the fact that we are only interested in the effects of shear on a hypersurface orthogonal to the fluid four-velocity. Thus, the four dimensional fluid shear measures the rate at which a four dimensional constant differential volume element deforms.

With the definitions above we also find that the bulk component of the stress tensor generalises, quite simply, to the form $\Theta P_{\mu\nu}$. Thus, the energy-momentum tensor for an anisotropic fluid is given by,

$$T_{\mu\nu} = (\rho+p)u_{\mu}u_{\nu} - pg_{\mu\nu} + 2\eta\sigma_{\mu\nu} + \zeta\Theta P_{\mu\nu} \quad , \tag{4.12}$$

where we have neglected all dissipative effects except those due to viscosity. η and ζ are the usual coefficients of shear and bulk viscosity, respectively. We further note that the equations of continuity and the equations of motion take the form

$$\begin{aligned} T^{o\mu}{}_{;\mu} &= 0 \qquad \text{(continuity)} \\ T^{i\mu}{}_{;\mu} &= 0 \qquad \text{(motion)} \end{aligned} \tag{4.13}$$

and we see from equation (4.12) that in the limit of special relativity, these equations reduce to the form of (4.1) and (4.3), respectively.

Now, for a spherically symmetric metric and comoving coordinates the volume expansion, Θ, can be very simply expressed as

$$\Theta = e^{-\alpha/2}\left[\frac{\beta_t}{2} + \frac{2R_t}{R}\right] \tag{4.14}$$

and the shear tensor has components

$$\begin{aligned}
\sigma^0{}_0 &= 0 \quad , \\
\sigma^1{}_1 &= e^{-\alpha/2}\beta_t/2 - ({}^1/_3)\Theta \quad , \\
\sigma^2{}_2 &= \sigma^3{}_3 = e^{-\alpha/2}R_t/R - ({}^1/_3)\Theta \quad , \\
\sigma^\lambda{}_\tau &= 0 \quad , \quad \lambda \neq \tau \quad ,
\end{aligned} \tag{4.15}$$

where we have dropped the prime for convenience. Substituting back into equation (4.12) we find that the non-zero components of the energy-momentum tensor are

$$\begin{aligned}
T^0{}_0 &= \rho \quad , \\
T^1{}_1 &= -p + 2\eta\left[e^{-\alpha/2}\frac{\beta_t}{2} - \frac{\Theta}{3}\right] + \zeta\Theta \\
T^2{}_2 &= T^3{}_3 = -p + 2\eta\left[e^{-\alpha/2}\frac{R_t}{R} - \frac{\Theta}{3}\right] + \zeta\Theta \quad .
\end{aligned} \tag{4.16}$$

Following the procedure outlined for the perfect fluid case, we can now obtain the Einstein equations in physical form for the case of a viscous fluid matter distribution. Equations (3.11) and (3.12) are left unchanged but equations (3.13) and (3.14) are modified by the presence of the non-zero viscosity in the problem. The full set of equations is , [cf. Eric Shaver, MSc. Thesis, 1986: courtesy of R.N. Henriksen, 1986, Lecture Notes.]

$$m_r = 4\pi\rho R^2 R_r \quad , \tag{4.17}$$

$$2m = R[1 + e^{-\alpha}R_t{}^2 - e^{-\beta}R_r{}^2] \quad , \tag{4.18}$$

$$m_t = -4\pi R^2 R_t(\text{Þ} + \chi) \quad , \tag{4.19}$$

$$\frac{\rho_t}{\text{Þ} + \rho} = -\left[\frac{\beta_t}{2} + \frac{2R_t}{R}\right] - \frac{\chi}{\text{Þ} + \rho}\left[\frac{\beta_t}{2} - \frac{R_t}{R}\right] \quad , \tag{4.20}$$

$$(\text{Þ} + \chi)_r + \chi\left[\frac{\alpha_r}{2} + \frac{3R_r}{R}\right] + \frac{\alpha_r(\text{Þ} + \rho)}{2} = 0 \quad , \tag{4.21}$$

where we have made use of the definitions

$$\begin{aligned} \text{Þ} &= p - \zeta\theta \quad , \\ \chi &= -\frac{2}{3}\eta e^{-\alpha/2}\left[\beta_t - \frac{2R_t}{R}\right] \quad . \end{aligned} \tag{4.22}$$

With these definitions and equation (4.14) our imperfect fluid energy-momentum tensor takes the simple form

$$T^{\mu}{}_{\nu} = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & -(\text{Þ}+\chi) & 0 & 0 \\ 0 & 0 & -(\text{Þ}-\chi/2) & 0 \\ 0 & 0 & 0 & -(\text{Þ}-\chi/2) \end{bmatrix} \quad . \tag{4.23}$$

The effect of viscosity on a cosmological model can be seen most easily in the case of a Friedmann-Robertson-Walker cosmology. For the FRW model the metric (4.10) takes the form

$$ds^2 = dt^2 - R^2\left[\frac{dr^2}{r^2(1-kr^2)} + d\Omega^2\right] \quad ,$$

with $R=rS(t)$. This model is isotropic and so is shear-free, $\chi=0$. Therefore, we need only consider the effects of bulk viscosity. For this case, the volume expansion takes the form, $\Theta=3S_t/S$. Thus equations (4.17)-(4.21) reduce to

$$
\left.
\begin{aligned}
m &= \frac{4\pi}{3}\rho r^3 S^3 \quad , \\
[S_t^2 + k]S &= \frac{8\pi\rho S^3}{3} \quad , \\
\frac{\rho_t}{p+\rho-3\zeta S_t/S} &= \frac{-3S_t}{S} \quad , \\
p &= p(t) \quad ,
\end{aligned}
\right\} \qquad (4.24)
$$

where the coefficient of bulk viscosity is regarded as a constant for the purpose of illustration. If we choose $k=0$, corresponding to the zero-curvature FRW model, and an equation of state, $p=a_s^2\rho$, we find that the Hubble parameter, $H=S_t/S$, is given by

$$H = \frac{8\pi\zeta e^{12\pi\zeta t}}{(1+a_s^2)(e^{12\pi\zeta t}+1)} \quad , \quad \text{for} \quad H < \frac{8\pi\zeta}{(1+a_s^2)} \qquad (4.25a)$$

$$H = \frac{8\pi\zeta e^{12\pi\zeta t}}{(1+a_s^2)(e^{12\pi\zeta t}-1)} \quad , \quad \text{for} \quad H > \frac{8\pi\zeta}{(1+a_s^2)} \qquad (4.25b)$$

where we have absorbed an arbitrary constant of integration. The behaviour of the Hubble parameter is shown in Figure 4.1. We see from the figure that as $t\to\infty$, $H\longrightarrow 8\pi\zeta/(1+a_s^2)$, cf. equations (4.25), i.e. the Hubble parameter tends to a constant and we have a de-Sitter expansion. If $H>8\pi\zeta/(1+a_s^2)$ then, equation (4.25b) demonstrates that for small t,

$$H = \frac{2}{3(1+a_s^2)t} \quad , \qquad (4.26)$$

which is identical to the form of the Hubble parameter for the case of a zero-curvature FRW model with a perfect fluid representation ($\zeta=0$),

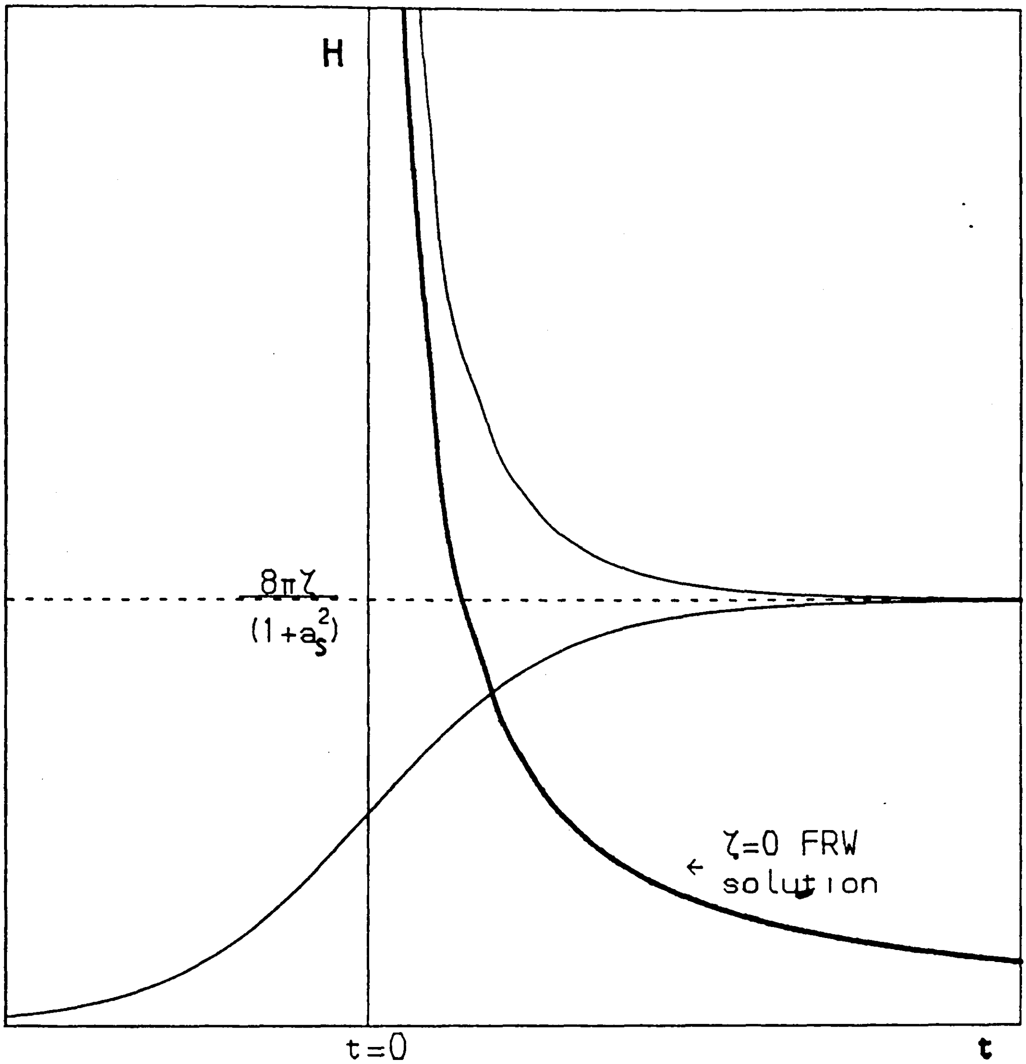


Figure 4.1 Behaviour of the Hubble parameter, $H=S_t/S$, for an isotropic spacetime with zero curvature and equation of state $p=a_s{}^2\rho$. The thick unbroken curve demonstrates the behaviour of the corresponding zero-curvature (k=0) FRW model.

also shown in Figure 4.1. Thus if the bulk viscosity is initially small, compared to the Hubble parameter, then the solution is almost indistinguishable from the perfect fluid solution at early times. If, however, the bulk viscosity is initially large, then there is a noticeable difference between the viscous and perfect fluid solution. A large bulk viscosity, which may be expected at early epochs, therefore plays an important role in the evolution of the Universe.

Equations (4.17)-(4.21) together with an equation of state specify the problem of describing a viscous fluid cosmology completely. However, as was stated in the introduction to this section, we wish to study solutions which admit a self-similar symmetry and we must, therefore, impose such a symmetry on the solution.

4.4 Self-Similar Representation of a Viscous Cosmology

In this section we shall investigate imperfect fluid cosmological models which display a homothetic symmetry. The description of such a symmetry in the cosmological regime was given by Cahill and Taub, (1971) and outlined in Chapter 2 of this thesis. In general, we shall follow the procedure of Henriksen and Wesson, (1978a) and Bicknell and Henriksen, (1978a). If our solutions are to have a homothetic symmetry, i.e. be self-similar of the first kind, no fundamental scales other than the gravitational constant, G, and the speed of light, c, can enter the problem, (Zel'dovich and Raizer, 1967). This does not prove difficult in the case of a perfect fluid cosmology with zero vacuum energy density, (cf. Henriksen and Wesson, 1978a, for example). However, by introducing viscosity we may destroy any possibility of the solution admitting a homothetic symmetry; the viscous coefficients, η and ζ, being independent dimensional quantities and therefore fundamental scales. We avoid this difficulty by choosing the form of these viscous coefficients to be functions of the characteristic scales already present in the system. The more general problem of finding a self-similar representation of imperfect fluid cosmological models when the viscous coefficients are allowed to be determined as part of the solution will be discussed in Chapter 6 of this thesis. (See also Henriksen 1987, who considers the problem of collimating a galactic nuclei jet, using viscosity to entrain the jet material, by investigating steady-state self-similar solutions in which the density and viscosity are held constant).

Both viscous coefficients have dimensions of $ML^{-1}T^{-1}$, i.e. (density x velocity x length) and must be non-negative, (Landau and Lifshitz, 1959). Therefore, we may write

$$\eta = h_s \rho |R_t| R \quad , \qquad \zeta = h_b \rho |R_t| R \quad , \tag{4.27}$$

where ρ, R and R_t are the characteristic density, velocity and scale for the cosmological model under discussion and h_s and h_b are numerical constants which are equivalent to inverse Reynolds numbers. The modulus sign reflects the fact that cosmological solutions may have expanding and contracting stages. Other possibilities for the form of the viscous coefficients are discussed in Chapter 6.

For self-similarity of the first kind to be admissible we must be able to obtain dimensionless quantities representing the physical parameters of the problem. Equations (4.17)-(4.21) have a unique dimensional representation in terms of c and G, viz., (c=G=1)

$$m = \frac{r}{2}M(\xi) \quad , \quad R = rS(\xi) \quad , \quad \alpha = \alpha(\xi) \quad , \quad \beta = \beta(\xi) \quad ,$$

$$\rho = \frac{\epsilon(\xi)}{8\pi r^2} \quad , \quad p = \frac{P(\xi)}{8\pi r^2} \quad , \quad \chi = \frac{\tau(\xi)}{8\pi r^2} \quad , \tag{4.28}$$

where $\xi = t/r$ is the dimensionless self-similar variable, Henriksen and Wesson (1978a). The Einstein field equations can thus be rewritten as ordinary differential equations in terms of these dimensionless quantities with ξ as the independent variable. First we note that the partial derivatives with respect to the coordinates t and r can be written as

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial \xi}{\partial t}\frac{d}{d\xi} = \frac{1}{r}\frac{d}{d\xi} \quad , \\ \frac{\partial}{\partial r} &= \frac{\partial \xi}{\partial r}\frac{d}{d\xi} = -\frac{\xi}{r}\frac{d}{d\xi} \quad , \end{aligned} \tag{4.29}$$

respectively. Therefore, Einstein's equations, (4.17)-(4.21), in their self-similar, dimensionless form are given by

$$M - \xi M' = \epsilon S^2(S - \xi S') \quad , \tag{4.30}$$

$$M = S[\, 1 + e^{-\alpha}S'^2 - e^{-\beta}(S-\xi S')^2 \,] \quad , \tag{4.31}$$

$$M' = -S^2 S'(P + \tau) \quad , \tag{4.32}$$

$$\frac{\epsilon'}{P+\epsilon} = -\left[\frac{\beta'}{2} + \frac{2S'}{S}\right] - \frac{\tau}{P+\epsilon}\left[\frac{\beta'}{2} - \frac{S'}{S}\right] \quad , \tag{4.33}$$

$$\alpha' = -\frac{2}{\xi^2(P+\epsilon+\tau)}\frac{d}{d\xi}[\xi^2(P+\tau)] + \frac{6\tau}{\xi S(P+\epsilon+\tau)}(S-\xi S') \quad , \tag{4.34}$$

where (') denotes $d/d\xi$, and we see that the dimensionless coefficient of shear viscosity is given by

$$\psi(\xi) = 8\pi r\eta = h_s S\epsilon|S'| \quad . \tag{4.35}$$

It is convenient to rewrite equations (4.30) and (4.32) as

$$S' = \frac{\epsilon S^3 - M}{\xi(P+\tau+\epsilon)S^2} \quad , \tag{4.36}$$

and

$$M' = -\frac{(P+\tau)(\epsilon S^3-M)}{\xi(P+\tau+\epsilon)} \quad . \tag{4.37}$$

All that remains to complete the solution is to specify an equation of state for the matter distribution. By examining the energy-momentum tensor, equation (4.23), we see that the pressure is anisotropic; due to the presence of the viscous terms. The principal pressure in the radial direction is different from the two transverse principal pressures. We thus have a different "effective" equation of state in each of these directions. For our purposes, it is most convenient to define an equation of state for one of the principal directions rather than choose some equation of state averaged over all three principal directions as is usually done in cosmological models. Thus we choose a radial equation of state, viz.,

$$P + \tau = a^2\epsilon \quad , \tag{4.38}$$

where a^2 is the square of the sound speed in units of c^2. If we also introduce a new parameter $\gamma=\epsilon S^3$, Einstein's equations then become

$$S' = \frac{S(1-M/\gamma)}{(1+a^2)\xi} \quad , \tag{4.39}$$

$$M' = -\frac{a^2}{(1+a^2)\xi}(\gamma-M) \quad , \tag{4.40}$$

$$M = S[\,1 + e^{-\alpha}S'^2 - e^{-\beta}(S-\xi S')^2\,] \quad , \tag{4.41}$$

$$\alpha' = -\frac{2a^2}{(1+a^2)}\left[\frac{2}{\xi} + \frac{\gamma'}{\gamma} - \frac{3S'}{S}\right] + \frac{6\tau S^2(S-\xi S')}{\gamma(1+a^2)\xi} \quad , \tag{4.42}$$

$$\beta' = -\frac{2}{(1+a^2)}\frac{\gamma'}{\gamma} + \frac{2(1-2a^2)}{(1+a^2)}\frac{S'}{S} + \frac{6\tau S^2 S'}{\gamma(1+a^2)} \quad , \tag{4.43}$$

with the definition

$$\tau = -\frac{2}{3}h_s\gamma|S'|e^{-\alpha/2}\left[\beta' - \frac{2S'}{S}\right]S^{-2} \quad . \tag{4.44}$$

To reduce these equations to a form suitable for integration we have to solve them simultaneously for the five unknowns, S', M', α', β', γ', in terms of the parameters, ξ, S, M, α, β, γ. Equations (4.39) and (4.40) are already in their final form, so we need only consider the remaining four equations. We begin by replacing the parameter τ in equations (4.42) and (4.43) by its expression (4.44). Thus we have, after collecting terms,

$$\alpha' = \frac{-2a^2}{(1+a^2)}\left[\frac{2}{\xi} + \frac{\gamma'}{\gamma} - \frac{3S'}{S}\right] - \frac{4h_s}{(1+a^2)\xi}(S-\xi S')|S'|e^{-\alpha/2}\left[\beta' - \frac{2S'}{S}\right] \tag{4.45}$$

$$\frac{2\gamma'}{\gamma} = \frac{2S'}{S}\left[1-2a^2+4h_sS'|S'|e^{-\alpha/2}\right] - \beta'\left[1+a^2+4h_sS'|S'|e^{-\alpha/2}\right] \quad . \tag{4.46}$$

In order to obtain a third equation, we differentiate equation (4.41) with respect to the self-similar variable ξ. This leads to the equation

$$\frac{M'}{S} - \frac{MS'}{S^2} = -e^{-\alpha}S'^2\alpha' + e^{-\beta}(S-\xi S')^2\beta' + 2\left[e^{-\alpha}S' + \xi e^{-\beta}(S-\xi S')\right]S'' \quad , \tag{4.47}$$

where M' and S' can be obtained from equations (4.39) and (4.40), respectively. In deriving equation (4.47) we have introduced the second derivative of the transverse scale factor, S'', which can be expressed in terms of 'known' quantities and the unknown γ' by differentiating equation (4.39), viz.,

$$S'' = \frac{S'^2}{S} - \frac{S'}{\xi} + \frac{S}{\xi(1+a^2)\gamma}\left[\frac{M\gamma'}{\gamma} - M'\right] \quad . \tag{4.48}$$

Thus we have three simultaneous equations, (4.45), (4.46), (4.47) with (4.48), in three unknowns, α', β', γ', which together with (4.39) and

(4.40), specify the field equations to be integrated. Solving these equations simultaneously, we find that the system of equations reduces to the following:

$$S' = \frac{S(1-M/\gamma)}{\xi(1+a^2)} \quad , \tag{4.49}$$

$$M' = \frac{-a^2}{\xi(1+a^2)}(\gamma-M) \quad , \tag{4.50}$$

$$\beta' = \frac{[(2C+AE)G + 2L + JE]}{[2F - JD - GAD + 2BG]} \quad , \tag{4.51}$$

$$\alpha' = [2C + AE + (AD-2B)B']/2 \quad , \tag{4.52}$$

$$\gamma' = -\frac{S^3}{2}[E + D\beta'] + \frac{3\gamma S'}{S} \quad , \tag{4.53}$$

where

$$A = \frac{2a^2S^3}{(1+a^2)\gamma} \quad , \tag{4.54a}$$

$$B = 4h_s|S'|e^{-\alpha/2}(S-\xi S')/\xi(1+a^2) \quad , \tag{4.54b}$$

$$C = 4[2h_s|S'|\ \frac{S'}{S}(S-\xi S')e^{-\alpha/2} - a^2]/\xi(1+a^2) \quad , \tag{4.54c}$$

$$D = [1+a^2 + 4h_s e^{-\alpha/2}|S'|S']\gamma/S^3 \quad , \tag{4.54d}$$

$$E = 4S'[1+a^2 - 2h_s e^{-\alpha/2}|S'|S']\gamma/S^4 \quad , \tag{4.54e}$$

$$F = e^{-\beta}(S-\xi S')^2 \quad , \tag{4.54f}$$

$$G = e^{-\alpha}S'^2 \quad , \tag{4.54g}$$

$$H = \frac{2[e^{-\alpha}S' + e^{-\beta}\xi(S-\xi S')]}{\gamma(1+a^2)\xi^2} \quad , \tag{4.54h}$$

$$J = \xi MS^4H/\gamma \quad , \tag{4.54i}$$

$$K = \xi(\gamma+2M)S' - \xi SM' - (\gamma-M)S \quad , \tag{4.54j}$$

$$L = \frac{M'}{S} - \frac{MS'}{S^2} - HK \quad , \qquad (4.54k)$$

The equations, (4.49)-(4.53), are too complex to solve analytically and so we utilise a finite difference technique to solve them numerically.

4.5 Numerical Solutions to the "Viscous" Field Equations

We shall now investigate the numerical solutions of the field equations (4.49)-(4.53). It was not possible to solve these equations using a truncated-step numerical procedure, such as a standard Runge-Kutta method, due to the lack of uniqueness of the solution encountered at any turning points, without carrying out the complex calculations of determining second derivatives of the physical parameters. (At the turning points the special solution S=constant is more stable and acts as an attractor for the numerical procedure). However, by adopting an Adams-Bashforth finite difference method, (see Khabaza, 1965), we were able to solve this problem, integrate through any turning points and hence obtain a complete solution.

The basic theory behind the Adams-Bashforth method is to use backward differences to predict the next value, i.e. by comparing the trend of the derivatives this method can extrapolate forwards to the next point in the solution and so on. For example, suppose we wish to solve the differential equation

$$\frac{dy}{dx} = f(x,y) \quad , \quad y=y_0 \text{ when } x=x_0 \quad ,$$

at intervals h in x. We first require a few starting values of y since the Adams-Bashforth method is not a "self-starter" and therefore needs some other method, such as Runge-Kutta, to determine the first few values of y.

Having done this we are now able to construct a table of x,y, In theory we may have to repeat steps (4.55) and (4.56) several times until the new value of f_1 agrees with its previous value. For simplicity, it was decided to dispense with the corrector phase of this procedure. This decision was vindicated by the success of the numerical integration in reproducing the analytical non-viscous solutions of Henriksen and Wesson, (1978a), hereafter HW1, and Bicknell and Henriksen, (1978a).

For a fifth order Adams-Bashforth method, using the above we find

$$f_1 = (4277f_0 - 7923f_{-1} + 9982f_{-2} - 7298f_{-3} + 2877f_{-4} - 475f_{-5})/1440 \; . \tag{4.58}$$

We then have

$$y_1 = y_0 + hf_1 \quad .$$

So our procedure is to use a fourth order Runge-Kutta numerical integration method as a starter, to produce f_{-5}, f_{-4}, ..., f_0, then an Adams-Bashforth method to obtain f_1 and repeat.

We are now in a position to solve the Einstein field equations for a self-similar imperfect fluid cosmology. Before we can continue we must specify the exact form for the equation of state. For instance, we could choose the sound speed such that the solution is adiabatic, $a^2=k\epsilon^{1-\gamma}$. However, we shall deal with the much simpler isothermal case with the square of the sound speed, a^2=constant and in particular we shall concentrate on the two extreme cases, $a^2=0$, corresponding to a "dust-like" solution, and $a^2=1$, corresponding to a "stiff" solution.

A Solutions with Equation of State, $P+\tau=0$

In the absence of any viscous terms, the equation of state for the hydrostatic pressure, p=0, corresponds to a dust solution, where the universe consists of a "pressureless-fluid" of particles. The introduction of the dissipative processes of viscosity causes us to redefine what is meant by the pressure of the fluid since the hydrostatic pressure is no longer a distinct quantity. We define the pressure in terms of the spatial components of the self-similar energy-momentum tensor, equation (4.23) in dimensionless form, so that the 'principal pressures', p_1, p_2 and p_3, of this matter distribution are given by

$$
\begin{aligned}
p_1 &= T_{11} = P + \tau \quad , \\
p_2 &= T_{22} = P - \tau/2 \quad , \\
p_3 &= T_{33} = P - \tau/2 \quad .
\end{aligned}
\tag{4.59}
$$

Our equation of state is chosen such that $p_i=a_i^2\epsilon$, (where i=1,2 or 3), and not some average $p=a^2\epsilon$. This, we feel, is more physical since we are not treating the fluid as a sum of distinct perfect fluid, bulk viscosity and shear viscosity components but as a single "viscous" fluid. In fact, as was stated in §4.4, we choose to work with the radial equation of state, $p_1=a^2\epsilon$, where for convenience we have dropped the subscript from the sound speed, and in the present discussion $a^2=0$ so that we are dealing with "viscous dust". We emphasise that by "viscous dust" we mean a fluid which obeys the equation of state, $P+\tau=0$. This is a different procedure from that of Heller et al., (1973), who chose an equation of state for the hydrostatic pressure and allowed the 'pressure' due to the non-zero bulk viscosity to be treated as a separate quantity.

The analytical, non-viscous, self-similar dust solutions of HW1 will be used to provide a check of the numerical integration. The non-viscous dust solutions are given in analytic form by

$$e^{\alpha} = 1 \quad , \quad e^{\beta} = S^2/y^2$$

$$M = y(1-\xi S'/S) \quad , \tag{4.60a}$$

$$S = \left[\frac{M}{1-M^2}\right]\sin^2 X \quad ,$$

$$\alpha_s \pm \xi = \frac{M}{(1-M^2)^{3/2}}\left[X - \frac{\sin 2X}{2}\right] \qquad \text{for } M<1 \tag{4.60b}$$

$$S = [\tfrac{3}{2}(\alpha_s \pm \xi)]^{2/3} \quad , \qquad \text{for } M=1 \tag{4.60c}$$

$$S = \left[\frac{M}{M^2-1}\right]\sinh^2 X \quad ,$$

$$\alpha_s \pm \xi = \frac{M}{(M^2-1)^{3/2}}\left[\frac{\sinh 2X}{2} - X\right] \quad . \qquad \text{for } M>1 \tag{4.60d}$$

The $\pm$ sign corresponds to an expanding or contracting solution respectively and α_s is a constant of integration expressing the 'size of the universe' at $\xi=0$. The starting conditions are chosen in such a way as to make the parameters of the viscous solutions match those of the non-viscous solutions on some surface, ξ=constant, which will be discussed separately for each solution.

We proceed then to integrate the field equations (4.49)-(4.53) numerically with $a^2=0$, using the finite difference method outlined above. As in the case of the non-viscous dust solutions, we can sub-divide the viscous dust solutions into three classes depending upon the value of the constant dimensionless mass M, see equation (4.41), i.e. $M<1$, $M=1$, $M>1$.

(i) $M=1$, (isotropic):

We shall consider this class of solution first since it is the simplest. When $M=1$ the initial conditions are found to be isotropic, (see equations (4.60)), and so are not affected by shear, $\tau=0$, and the equation of state reduces to $P=0$. Thus, this solution is represented by an isotropic imperfect fluid. For this viscous dust model the universe is ever-expanding, with the volume expansion, $\Theta=2/t$. The only difference between this solution and the HW1 solution for $M=1$ is in the interpretation of the pressure. In the current solutions the pressure is modified by the presence of a non-zero bulk viscosity. For a discussion of similar models see §4.2. The behaviour of the dimensionless scale factors, S and $e^{\beta/2}$, is shown in Figure 4.2.

This particular class of dust solution can represent an approximation to the hierarchical universe structure proposed by De Vaucouleurs, (1971), since it yields an inverse power law in density along the backward light cone for some observers, (see HW1, Figure 1.). Such a model is much less complicated than the general dust solutions of Bonnor, (1972), or Wesson, (1975), and as such further emphasises the practical use of cosmological solutions which have a self-similar symmetry. We should note that the behaviour displayed by the HW1 solutions depends critically on the existence of a parameter, α_s. In the viscous solutions discussed here it is convenient to choose α_s to be zero although in the isotropic dust case a non-zero α_s could be included quite easily. However, the viscous solution then becomes precisely the solution discussed by HW1, with the modification that the pressure must be reinterpreted since it now

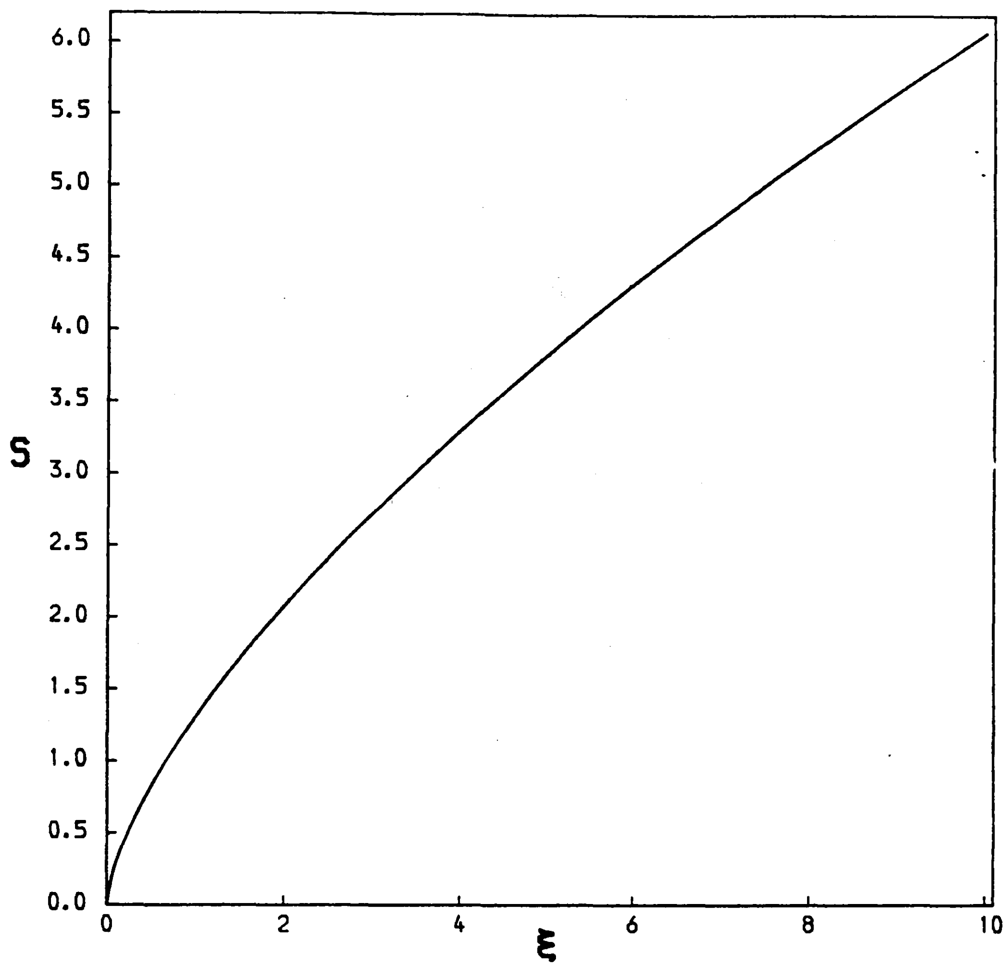


Figure 4.2 Isotropic, self-similar dust solution. The figure demonstrates the variation of the radial and transverse scale factors ($3e^{\beta/2}=S$) with the self-similar variable ξ.

contains a bulk viscosity term, and so we will not discuss it further. The convenience of choosing $\alpha_s=0$ becomes more apparent in the case of the anisotropic dust solutions. Let us now consider these more interesting classes of solution, ($M\neq1$).

(ii) $M<1$, (closed):

This class of solutions is not isotropic and therefore we have to consider the effects of a non-zero shear. The $M<1$ solutions of HW1 exhibit a maximum in the transverse scale factor, S, i.e. $S'=0$ for some finite value of ξ. When $S'=0$ we see, from equation (4.35), that the self-similar viscosity coefficient $\Psi=0$ and therefore the shear term, $\tau=0$ at this point. Thus, by choosing our initial conditions sufficiently close to this maximum in S, we minimise the effect of shear on the solution at this point. Having chosen our initial conditions the solutions are then integrated in both directions. Figures 4.3a,b show the behaviour of the transverse and radial scale factors, respectively, for different values of the shear constant h_s. It should be noted that h_s is necessarily quite small due to the limits imposed by observations of the quadropole anisotropy of the microwave background radiation at the present epoch, (cf. Fabbri et al., 1980, and Gorenstein and Smoot, 1981). However, we shall consider a wide range of values, $h_s=0\rightarrow1$, to allow for any large anisotropies which may exist at earlier epochs.

The effect of viscosity on the scale factors, S and $e^{\beta/2}$, is as we would expect. The presence of viscosity has a "slowing down" effect on any expansion such that the rate of change of either scale factor, at any given value of ξ, decreases with increasing viscosity. We see from Figure 4.3a that as the solutions contract from the initial surface, $\xi=\xi_s$ (corresponding to $S'=0$, $S=S_{max}$), towards the spatial origin ($r=0/\xi=\infty$) the transverse scale factor decreases monotonically to zero.

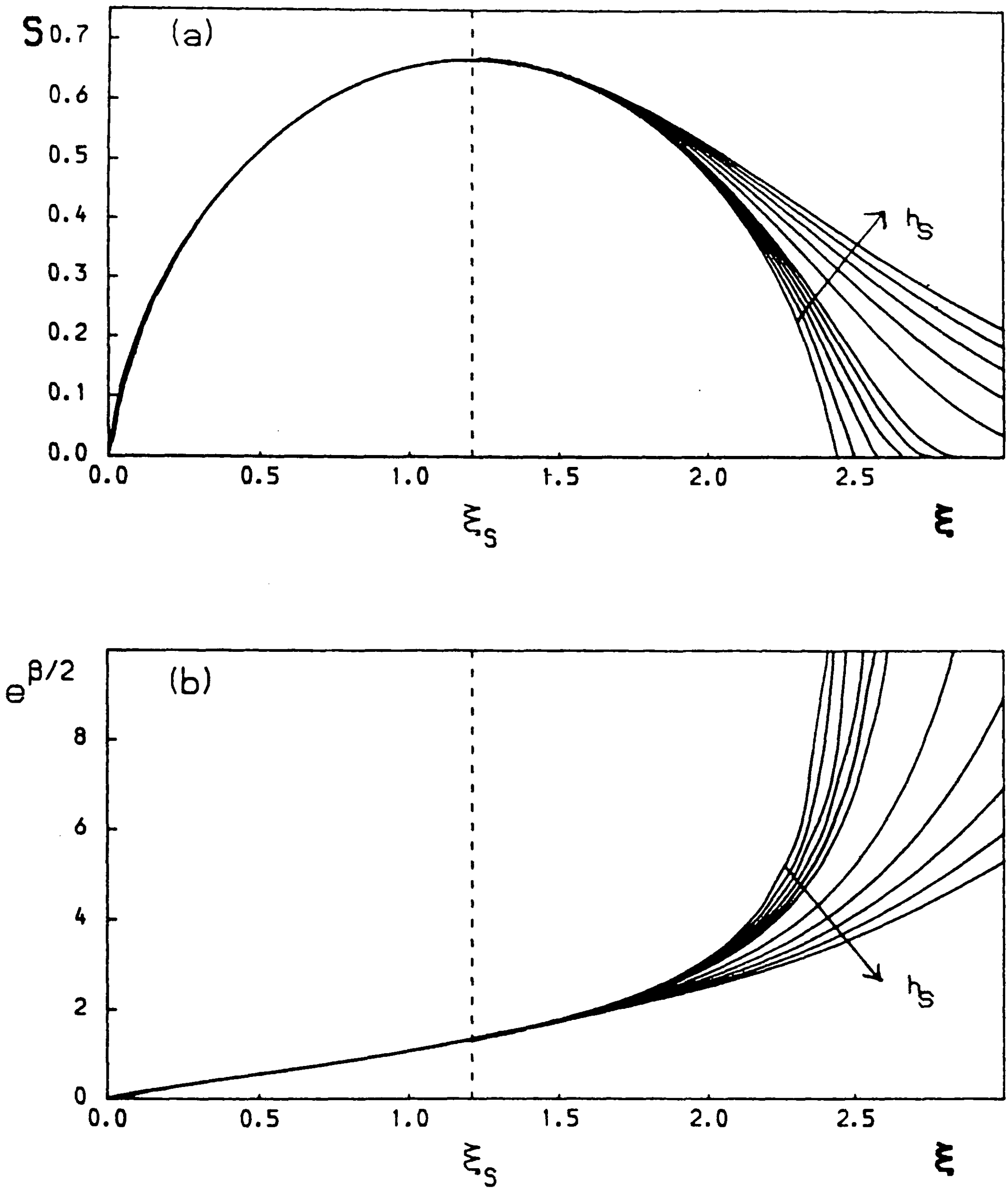


Figure 4.3 Evolution of the (a) transverse and (b) radial scale factors of a 'closed' dust solution for a range of initial viscosities, characterised by the viscous constant h_s=0,0.02,...,0.08,0.1,0.2,...,0.6. The broken line indicates the initial surface, $\xi=\xi_s$, corresponding to the maximum in the transverse scale. Throughout the figures in this chapter, the arrow will denote the direction of increasing h_s.

As the amount of initial viscosity increases the surface on which S=0 gets closer and closer to the origin, r=0. As we integrate towards the origin, ξ=0, the behaviour is markedly different. The 'contraction rate' still decreases with increasing viscosity but the effect is not so dramatic. This behaviour is not too surprising since in a self-similar solution the physical parameters, such as viscosity, may scale with the self-similar variable ξ, so that as ξ decreases from ξ_s the viscosity need not cause the solutions to diverge significantly.

The solutions are anisotropic and if we are to discuss the dissipative effects of viscosity on the expansion of the solution as a whole, we should also discuss the behaviour of the radial scale factor, $e^{\beta/2}$. It is found that all solutions exhibit a monotonically increasing radial scale with the more viscous solutions expanding more slowly, again as we would expect. This monotonic behaviour of the radial scale factor means that we have to be careful about our definition of a 'closed' solution. The M<1 dust solutions of HW1 are closed only in the sense that the transverse scale exhibits a maximum. If we consider the proper volume

$$V = 4\pi\int_0^{\infty} e^{\beta/2}R^2 dr \qquad ,$$

we find that this diverges at any finite time (for the analytic solutions) and therefore the solutions are actually open in the sense that the volume is infinite.

Now, the behaviour of the overall expansion of any solution can be most readily expressed by the evolution of the volume expansion, Θ, which is given in self-similar form by

$$\Phi(\xi) = r\Theta = e^{-\alpha/2}\left[\frac{\beta'}{2} + \frac{2S'}{S}\right] \qquad , \tag{4.61}$$

where $\Phi(\xi)$ is the dimensionless volume expansion. The behaviour of Φ for a range of h_s is shown in Figure 4.4, and again one would expect, by physical arguments, that the rate of change of Φ would decrease as h_s increased. This can be most readily expressed by the self-similar form of the Raychaudhuri equation (Raychaudhuri 1955), viz.,

$$\Phi' = -\frac{1}{3}\Phi^2 - 2\sigma^2 - \frac{1}{2}(\epsilon+3P) \quad , \qquad (4.62)$$

where σ^2 is the self-similar scalar shear, $\sigma^2=(r^2/2)\sigma_{\mu\nu}\sigma^{\mu\nu}$ and we have neglected all dissipative processes other than viscosity. We see from this equation that as the shear increases the rate of change of the volume expansion decreases. This behaviour is indeed displayed by the volume expansion, in Figure 4.4, close to the initial surface $\xi=\xi_s$. However, the imposed self-similar symmetry of the first kind forces us to choose the form of the dynamic viscosity, η, (see equation (4.35)), and although the values of h_s in the viscous solutions, allow us to compare the relative viscosity of any two solutions *initially*, we cannot extend the comparison to arbitrary values of ξ, since the physical viscosity 'evolves' in a different way for each solution. In general, therefore, it is meaningless to look for trends in behaviour as we progress from solutions with lower values of h_s to solutions with higher values. The constant h_s is merely a label and can only be treated as a measure of the physical viscosity close to the ξ-surface where it is introduced. Therefore, we should only expect to use our physical intuition to discuss comparisons between different solutions close to this *initial* surface.

Thus in these closed viscous dust solutions the introduction of a shear causes the universe initially to vary more slowly in both the transverse (contracting) and radial (expanding) directions, when

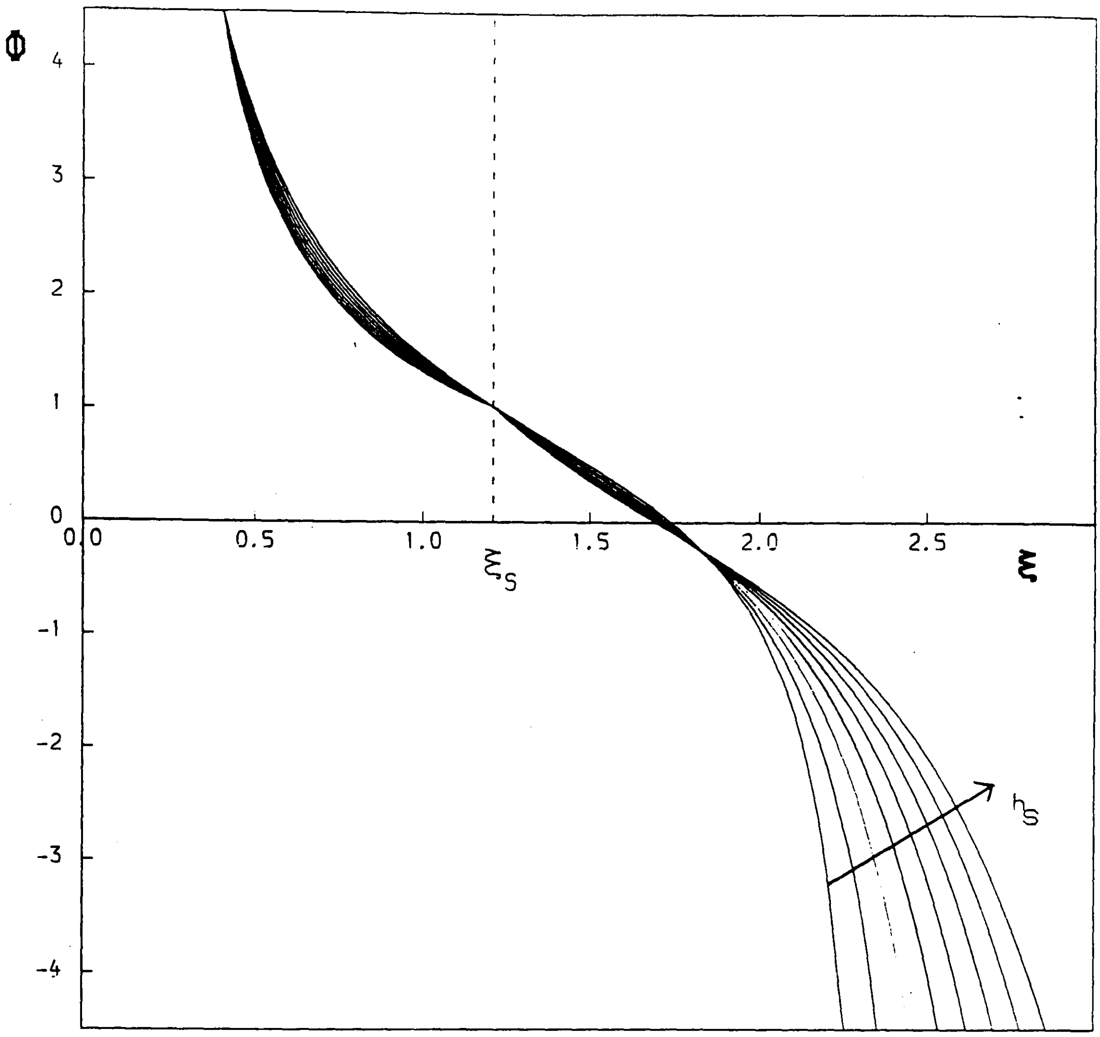


Figure 4.4 The diagram illustrates the behaviour of the dimensionless volume expansion, Φ, as a function of the variable ξ, for the solutions of Figures 4.3. Again ξ_s denotes the initial surface of the solution but now the viscosity constant takes the values, h_s=0,0.1,...,0.7.

compared to the non-viscous solutions or viscous solutions with less viscosity, characterised by the constant h_s. The overall expansion therefore also varies more slowly, initially, with increasing viscosity. However, since the physical viscosity has a functional form dependent upon the parameters describing a solution, the evolution of the scale factors becomes rather more complicated as we depart from the initial surface and before we can say anything about the effect of viscosity on the solution as a whole we must have a knowledge of how the viscosity behaves within any particular solution.

Figure 4.5 shows the evolution of the dynamic viscosity for a large range of viscous solutions. At this starting surface, $\xi=\xi_s$, the rate of expansion, S', is zero and, by virtue of the definition (4.35), the self-similar dynamic viscosity is also zero ($\psi=0$). The figure demonstrates that as the solution approaches the origin, $\xi=0$, the self-similar viscosity, ψ, tends to infinity in all of the viscous solutions. This is consistent with the behaviour of the transverse scale since $\psi\propto|S'|$. As ξ increases away from ξ_s the behaviour of the dynamic viscosity is less severe and diverges much less rapidly.

Some interesting properties of these solutions develop if we consider the volume expansion expressed in physical coordinates, i.e. $\Theta=\Phi/r$. Figure 4.6 shows the variation of Θ with the coordinate r at a given time, $t=t_0$, again for a range of values of h_s. This demonstrates that for any given value of t, there is only one value for which $\Theta=0$, i.e. there is a unique value of r, which we shall call the critical surface, $r=r_c$, such that if $r<r_c$, the matter is contracting and if $r>r_c$ the matter is expanding. In addition, the boundary defined by $r=r_c$ is increasing with t, i.e. it is moving outward as t increases. (This can easily be derived from the fact that the solution is self-similar). We

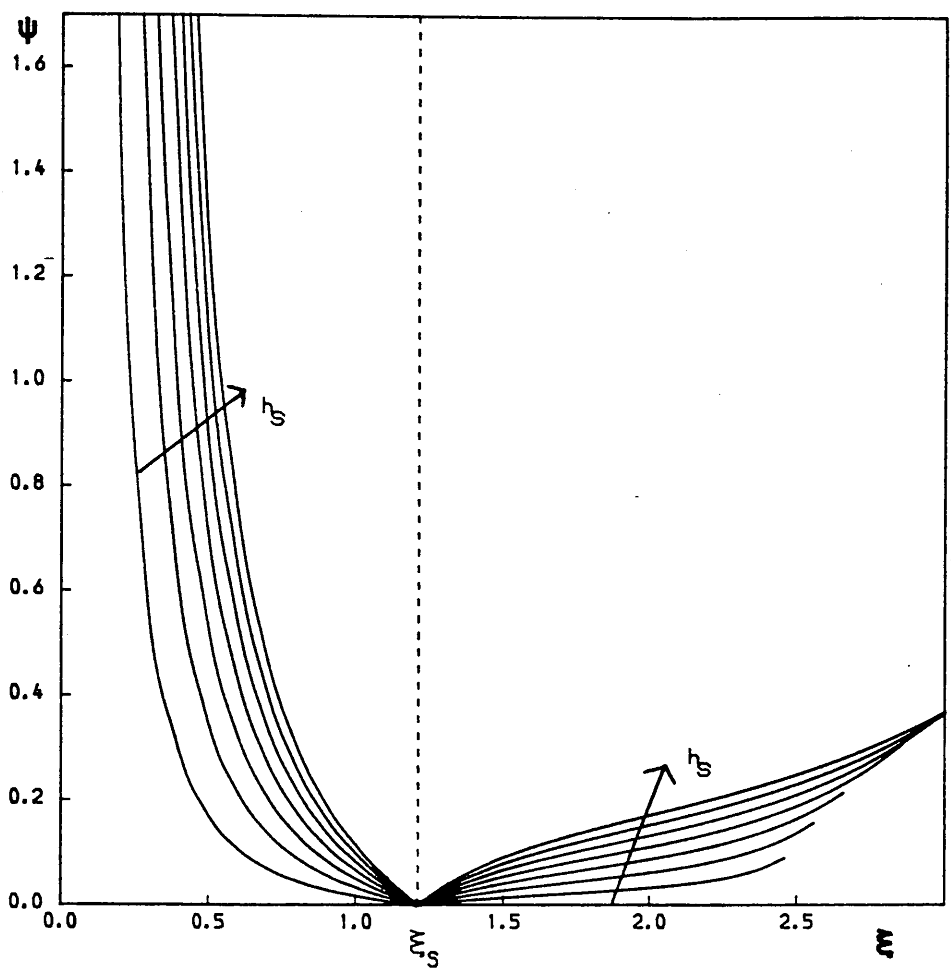


Figure 4.5 Evolution of the self-similar dynamic viscosity, ψ, in a closed viscous dust solution, demonstrating the divergence as $\xi \to 0$ in all viscous models (h_s=0,0.1,...,0.7).

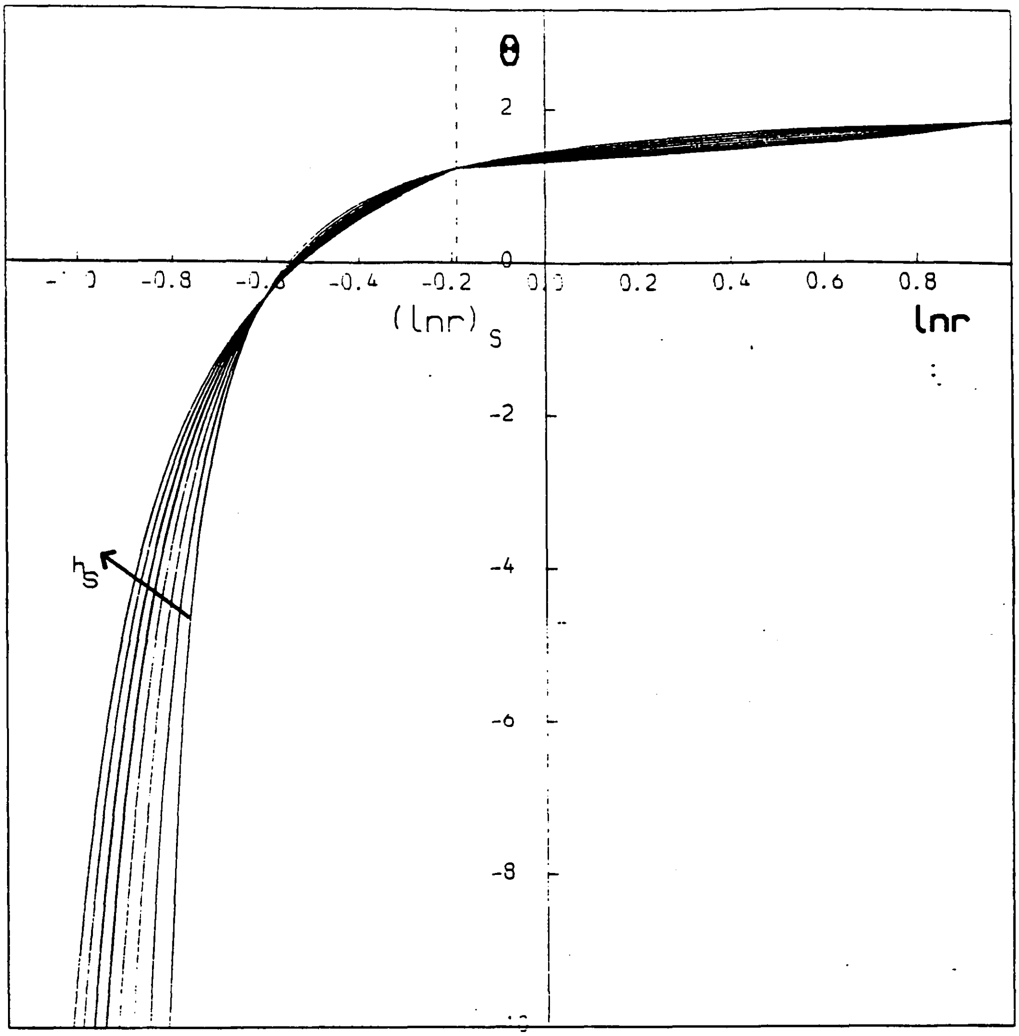


Figure 4.6 Behaviour of the physical volume expansion, Θ, with distance from the spatial origin (r=0) at a given time, $t=t_0$. The solutions are integrated away from the surface $(\ln r)_s=-\ln(\varepsilon_s)$ with t_0 scaled to unity.

also note that Θ increases monotonically from large negative values for small r, through zero at $r=r_c$, to finite positive values for large r. These properties are only slightly modified by the inclusion of shear; for a given value of t the critical surface occurs at slightly larger values of r with increasing h_s. The overall behaviour, however, remains the same with the volume expansion of the fluid world lines for all solutions, regardless of shearing rate, increasing monotonically with r. This situation is similar to that obtained by Coley and Tupper, (1983), while investigating solutions with shear, a radial magnetic field and non-negligible heat conduction, see their Figure 1b.

The properties exhibited by our model and the model of Coley and Tupper highlight the importance of anisotropic cosmological models to the study of the evolution of the Universe and, in particular, to the study of the critical density problem, (Sciama, 1971). The behaviour of the volume expansion demonstrates that we cannot decide upon the eventual fate of the Universe merely by determining the density of matter it contains. We must also determine the qualitative nature of the matter content. We concur with the belief of Coley and Tupper, (1983), that the standard treatment of the critical density problem may be too simplistic and that investigations of general anisotropic models for the early Universe must be carried out in a self-consistent manner. This latter problem will be discussed further in the concluding section of this chapter.

The last class of self-similar dust solution we need to consider is that of the anisotropic open solutions characterised by M>1.

(iii) M>1, (open):

In this class of solutions it is found that due to the chosen form of the viscosity coefficients the shear dominates the solution for small values of the self-similar variable ξ. We therefore choose our initial conditions so that the parameters of the viscous solutions match those of the non-viscous solutions at a ξ=constant surface, suitably close to the spatial origin, $r=0, \xi=\infty$. This choice of initial conditions is justified by the numerical solutions. If we integrate the viscous solutions forward towards the origin, r=0, we find that the effect of the shear term, which manifests itself as $\tau S^2 S'/\gamma$ (see equations (4.42) and (4.43)) tends to zero. This can be demonstrated by comparison with the analytic non-viscous solution and is due to the isotropisation of the solutions as $\xi\to\infty$, which reduces the shear considerably. The form of our initial conditions are therefore vindicated since the effect of shear decreases rapidly as we approach the spatial origin. Figures 4.7a,b show the variation of the transverse and radial scale factors for this class of solutions. The presence of viscosity results in the surface corresponding to $S=R/r=0$ occurring closer to the spatial origin ($\xi=\infty$) as the initial viscosity increases.

If we consider the behaviour of the self-similar volume expansion, Figure 4.8, we find that the introduction of a viscous term of the form (4.35) has a dramatic effect. Even the presence of a very small initial viscosity, h_s, causes the volume expansion to behave quite differently from the analytic solutions, close to the origin $\xi=0$. This can be directly related to the form chosen for the dynamic viscosity, Ψ (or η). It was shown earlier that as $\xi\to 0$, the viscous terms in the Einstein field equations (4.49)-(4.53) dominate the behaviour of the solution and rapidly diverge. Thus, when a small viscosity is introduced, the expansion of the fluid world lines in the solution diverges accordingly, as $\xi\to 0$.

Another interesting parameter to consider is $\gamma=\epsilon S^3$, which

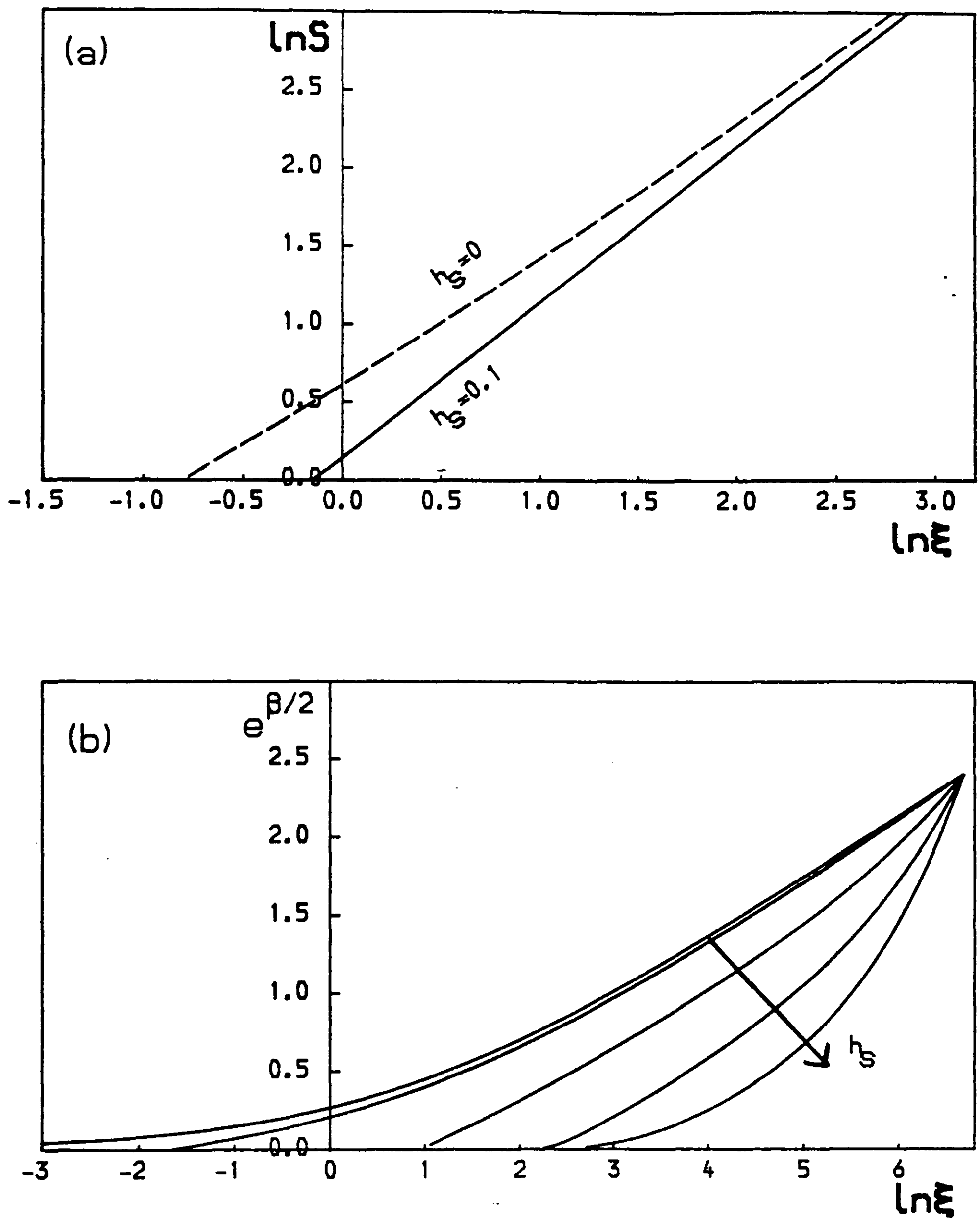


Figure 4.7 Variation of (a) the transverse and (b) the radial scale factors for the class of dust solutions, M>1. For clarity we have only shown the form of the transverse scale factor for one viscous solution (h_s=0.1). For the radial scale, the viscosity constant h_s takes the values 0, 10^{-5}, 10^{-4}, $10^{-3.5}$ and 10^{-3}.

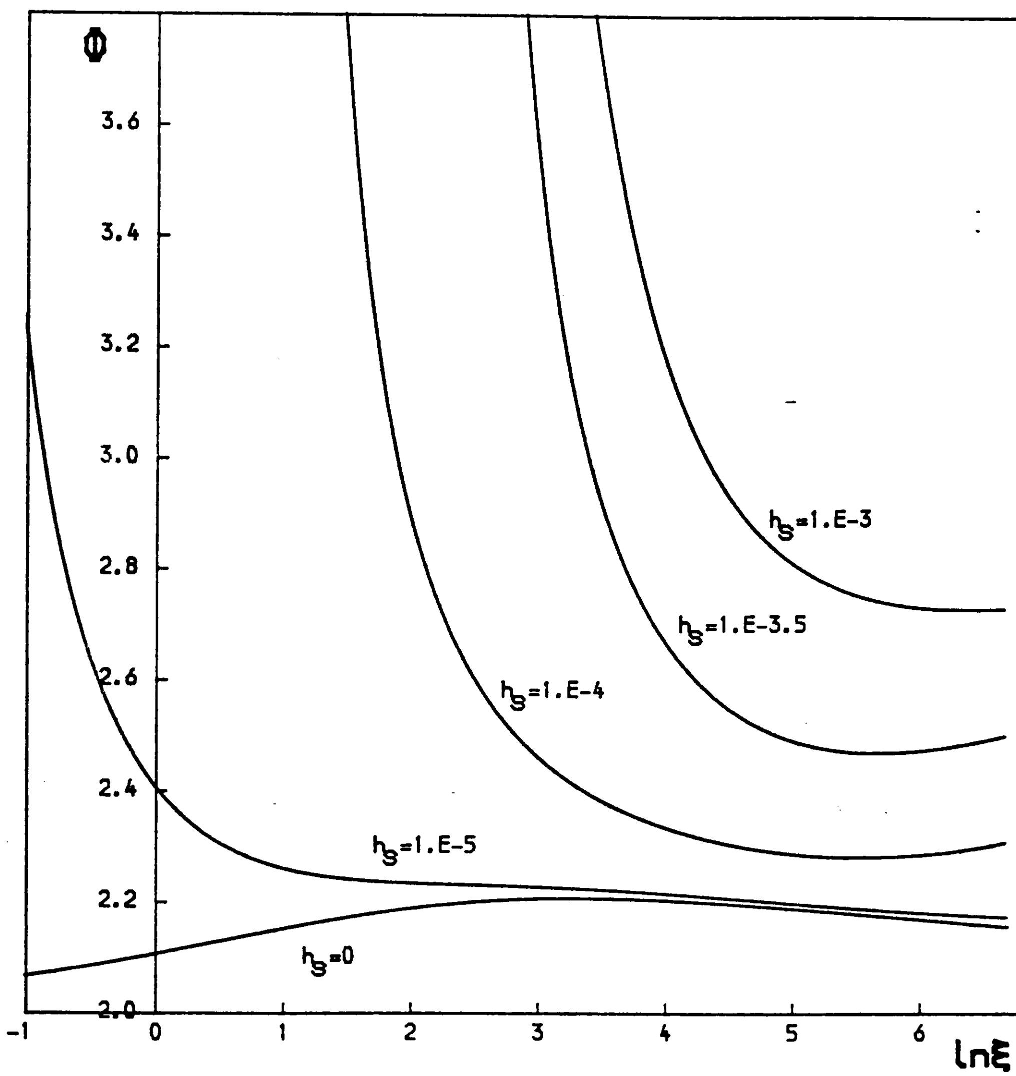


Figure 4.8 Self-similar volume expansion in an 'open' dust solution (M>1) for the range of viscous solutions shown. The right hand point of each curve corresponds to the surface on which the viscosity is introduced.

would be a measure of the mass if the solutions were uniform (see equation (3.33) with $\Delta=0$). Figure 4.9 shows the behaviour of γ for a range of initial viscosity, characterised by the constant h_s. Initially γ is very large, corresponding to the large value of the transverse scale factor S, which demonstrates that for large ξ the energy density ϵ varies more slowly than S^{-3}. In the non-viscous case, γ falls off monotonically as ξ decreases. However, the introduction of viscosity causes γ to exhibit a minimum and diverge for small ξ. This would be consistent with the energy density varying more rapidly than S^{-3} as $\xi \to 0$ ($S \to 0$). Thus, the presence of viscosity in the form (4.35) has a marked effect on the behaviour of the energy density of the solution, as ξ approaches zero.

The class of solutions for which $M>1$ is interesting in that it represents ever-expanding anisotropic cosmologies, which enable us to investigate the effects of large primordial (or induced) anisotropies on the present epoch. These investigations together with the observations of present-day anisotropies, (Gorenstein and Smoot, 1981), should help provide limits on the amount of anisotropy which may be present in the early universe whether primordial or generated by tidal motions.

Figure 4.10 shows the variation of the self-similar viscosity, ψ, with the self-similar variable, ξ. The figure demonstrates that the physical viscosity of any particular open viscous dust solution diverges rapidly as ξ becomes small. (Note the behaviour of ψ for large ξ due to the isotropisation of the cosmological fluid). For a given observer, r=constant, this behaviour is equivalent to the viscosity of the cosmological fluid decaying with time. The observed universe at the present epoch is very close to a perfect pressureless

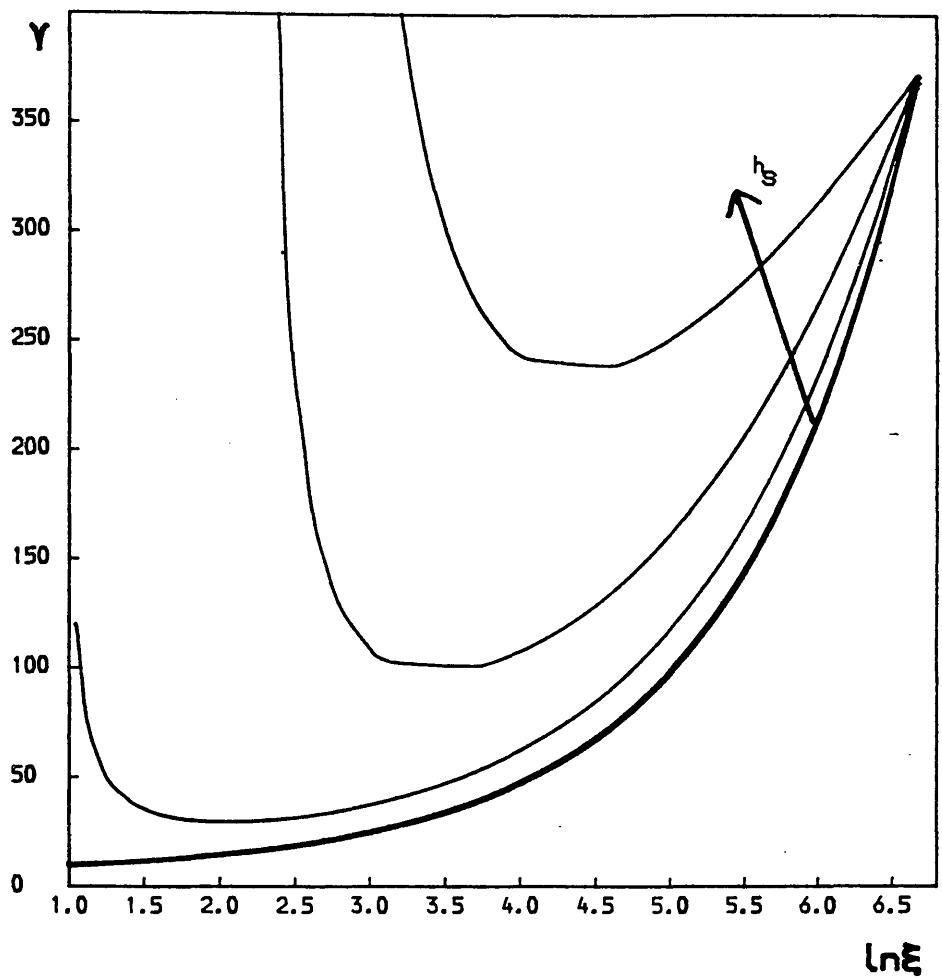


Figure 4.9 The 'uniformity parameter' γ as a function of the self-similar variable, ξ, for the values of h_s discussed in the last figure. The thick curve corresponds to the solutions $h_s=0$ and $h_s=10^{-5}$. For the range of ξ shown there is very little deviation between these two solutions. However, it should be noted that as $\xi \to 0$ the viscous solution ($h_s=10^{-5}$) does diverge in a similar fashion to the other viscous solutions shown.

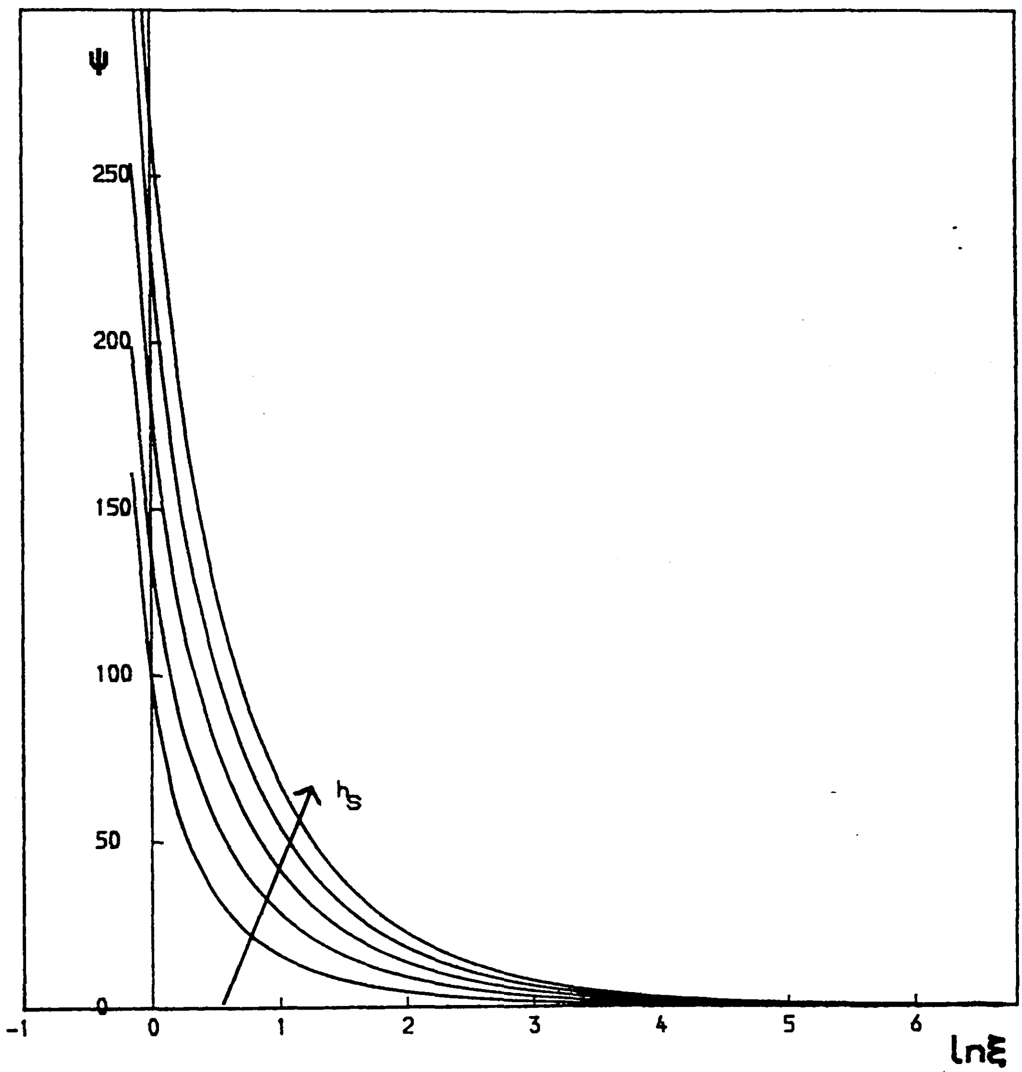


Figure 4.10 Variation of the self-similar dynamic viscosity, ψ, for a large range of viscous constants, h_s=0,0.1,...0.5, in an 'open' viscous dust model, displaying a similar behaviour to that of the viscosity in the 'closed' dust models as ξ approaches zero.

fluid and so any physically valid solution would be expected to display this type of behaviour. Thus the open (M>1) viscous dust solutions are viable cosmologies in that the anisotropic components are extremely small at the present epoch and the cosmological solutions are almost indistinguishable from an isotropic perfect fluid matter distribution.

Having covered self-similar solutions with an equation of state $P+\tau=0$ we shall now proceed to consider the other extreme case, that of a ' stiff' equation of state, $P+\tau=\epsilon$, where the speed of sound is equal to the speed of light.

B Solutions with Equation of State, $P+\tau=\epsilon$

It has been suggested by Zel'dovich (1962) that in the limit of high density in an isotropic fluid, strong interactions may cause the fluid to develop what has come to be known as a stiff equation of state, where the hydrostatic pressure equals the energy density. Harrison (1965) has criticised this work, claiming that under a more realistic treatment of the same problem the equation of state tends to that of an isotropic fluid with three degrees of freedom, $a^2=1/3$. (Harrison states that Zel'dovich's assumption of ignoring the high energy of baryons imposed by the exclusion principle is physically unrealistic). There are, however, other models which predict a stiff equation of state. For example, Walecka (1974) has refined Zel'dovich's suggestion to include a massive scalar field and extra interaction terms and, like Zel'dovich, finds that $p\rightarrow\rho$ at high density.

The problem of the equation of state in a superhigh-density region has also been discussed in review articles by Canuto, (1974, 1975). Canuto concludes that, although the available data from the study of neutron stars and various other 'experiments' are far from conclusive, the stiff equation of state seems favoured in high density regions.

There are, of course, many models of particle physics at high density which under varying assumptions do not produce a stiff equation of state. For instance, Hagedorn (1970), describes a "universal fireball" with an equation of state $p \propto \ln \rho$ as a valid thermodynamical model of the Universe at early epochs, and Collins and Perry (1975), describe a model in which particles become asymptotically free at high density resulting in a $p=\rho/3$ equation of state. Thus, there is not yet any compelling experimental evidence that the equation of state is stiff at early times, but there are several models in which it may be and so the possibility should not be dismissed. In fact, a stiff equation of state has proved useful in the study of the growth of primordial black holes in spherically symmetric similarity models of the universe, Lin et al., (1976), and Bicknell and Henriksen, (1978a,b).

We should note that a stiff equation of state will only be valid in a regime where strong interactions are possible and so we would not expect the universe to be stiff after 10^{-4}s since the density is then less than nuclear density. However, once we allow the assumption of a stiff equation of state it is possible to find solutions of the field equations which have a 'stiff' fluid as a source. For example, Wainwright et al., (1979), have produced exact inhomogeneous solutions describing an irrotational perfect fluid with a stiff equation of state and, more interestingly, McIntosh, (1978) has shown that these solutions admit a three parameter homothetic group of motions and are thus self-similar cosmologies.

Bicknell and Henriksen (1978a), hereafter BH1, set down the equations for non-viscous, self-similar cosmologies with a stiff equation of state and we will use these solutions as a base for our viscous models, i.e. to provide starting values and act as a check of the numerical procedure. The behaviour of the BH1 solutions is shown in Figure 4.11. In this diagram, $m=(a_\sigma/a_\omega)M$, $s=(a_\sigma/a_\omega)S$ and $a_\sigma{}^4=1/3$ as discussed below. The point ($m=1/4$, $s=1$) corresponds to the particle horizon of the cosmological solution. We shall refer to this diagram later.

The non-viscous dust solutions of HW1 contained the constants of integration, a_σ and a_ω, which appear in the expressions for the metric coefficients e^σ and e^ω, respectively,viz.,

$$e^\omega = a_\omega S^{-4}\eta^{-2m} \quad ,$$

$$e^\sigma = a_\sigma(\eta\xi)^{-2n} \quad , \tag{4.63}$$

where $m=1/(1+a^2)$ and $n=a^2m$. The constants, a_σ and a_ω, can be taken as unity if appropriate scale changes in the coordinates, t and r, are made. However, such arbitrariness is not permitted if the square of the sound speed is also equal to unity, (in units of $c=1$). The reason for this is that under the coordinate scalings; $t=a\hat{t}$, $r=b\hat{r}$, which are similarity preserving, the constants of integration become, (for $p=\rho$),

$$\hat{a}_\sigma = a_\sigma \quad , \quad \hat{a}_\omega = ba_\omega \quad .$$

Thus, a_ω may be given an arbitrary value by scaling the coordinate r but a_σ is independent of scale and is therefore physically significant.

BH1 were investigating similarity black hole solutions and introduced a function $V=(a_\omega/a_\sigma)^2/S^2$, which measures the velocity of the ξ=constant hypersurfaces relative to the fluid. The induced metric on such a hypersurface is given by

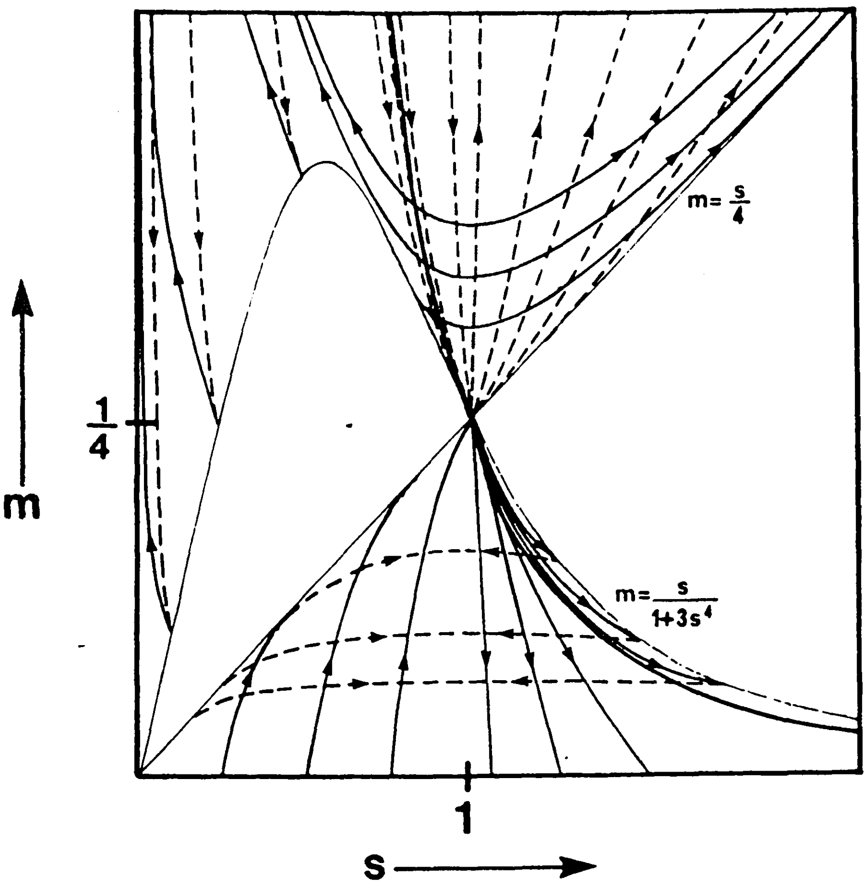


Figure 4.11 The two families of the integral curves of the analytic self-similar stiff solutions of BH1 for the case $a_\sigma{}^4=1/3$. The thin unbroken curves correspond to one family of solutions and the dashed curves the other. The boundaries $m=s/4$ and $m=s/(1+3s^4)$, separating the (m,s)-plane into forbidden and allowed regions, are indicated. The thick unbroken curve through the critical point is the Robertson-Walker curve $(1/4)s^{-3}$. The direction of increasing ξ is indicated by the arrows, (cf. Bicknell and Henriksen 1978a).

$$ds^2 = e^{\sigma}(1-V^2)dt^2 - R^2d\Omega^2 \quad .$$

Thus, if $V=1$ and e^{σ} is finite, the surface contains a null vector and is either an event horizon or a particle horizon. It transpires that there are two values of ξ for which $V=1$. For our present purposes we will only be interested in the hypersurface which corresponds to the universe particle horizon. [Our notation is slightly different from that of BH1 in that our metric coefficients are denoted e^{α} and e^{β} and not e^{σ} and e^{ω}. However, we will still retain the notation a_{σ} and a_{ω} for the constants of integration].

In what follows a_{ω} is chosen to be unity, which can be done without loss of generality, and we will concentrate on solutions with $a_{\sigma}{}^4=1/3$, since this allows us to obtain a Robertson-Walker universe as a particular solution of the field equations. This is an important point because information cannot travel beyond the particle horizon of an inhomogeneous region and we are therefore free to take the universe to be exactly Robertson-Walker outside the particle horizon. The Robertson-Walker solution is isotropic and as such will not be affected by shear. Thus we need only consider the viscous solutions within the particle horizon.

In the dust solutions, we had to specify the starting conditions on some initial ξ=constant surface. In the present solutions our initial surface is necessarily the hypersurface corresponding to the particle horizon. To obtain starting conditions for our general viscous solutions we expand away from this particle horizon, ($V=1$, $a_{\sigma}S=1$), using a standard Taylor expansion. This allows us to begin our solutions on one of the curves which "peel off" the Robertson-Walker curve, see Figure 4.11. (Since we will only be concerned with solutions for which ξ is increasing we "peel off" to the right of the

Robertson-Walker curve). It would, of course, be pointless to remain on the Robertson-Walker solution as we are interested in studying anisotropic solutions.

From BH1 we have the differential equation,

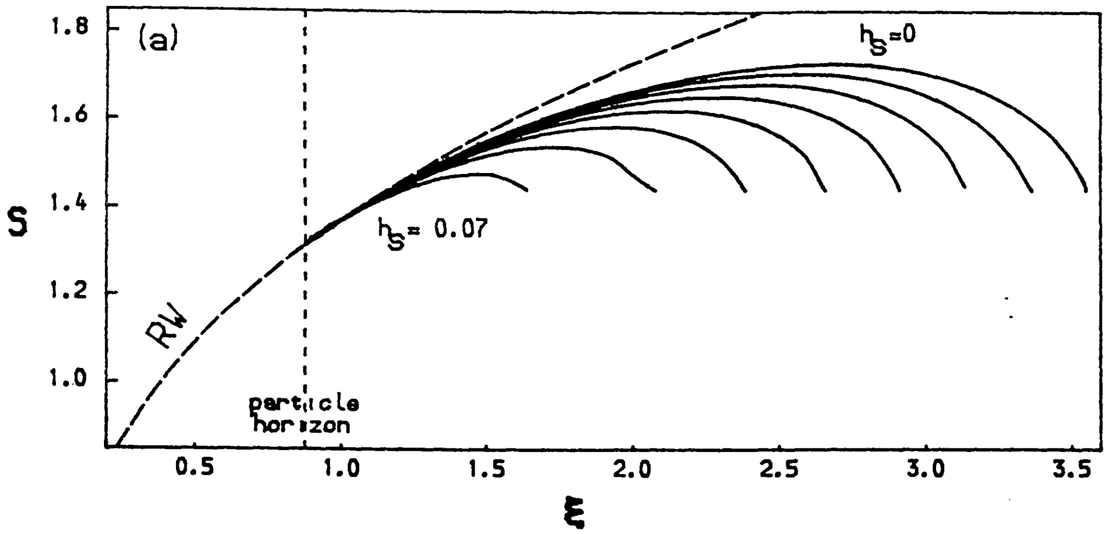
$$s^2(1-s^4)\left[\frac{dm}{ds}\right]^2 + [2ms(2a_\sigma{}^4+1+s^4) - 4a_\sigma{}^4 s^2]\frac{dm}{ds} + m^2(1-s^4) = 0 \quad , \qquad (4.64)$$

where $a_\omega=1$, $s=a_\sigma S$ and $m=a_\sigma M$. We want to expand about the critical point $s=1$ $(V=1)$, $m=a_\sigma{}^4/(a_\sigma{}^4+1)$, which with $a_\sigma{}^4=1/3$ is given in the (m,s)-plane by $s=1$, $m=1/4$. Now taking the expansion of BH1 for a trajectory which peels off from the Robertson-Walker curve we have

$$ms^3 = \frac{1}{4} - \frac{A}{3}[1-4m/s]^{3/2} + \ldots \quad , \qquad (4.65)$$

so that curves to the right of the Robertson-Walker curve correspond to $A<0$, where A is an arbitrary parameter characterising a one-parameter set of curves which pass through the critical point with the allowed slope. By specifying the constant A, we are, in effect, choosing the initial trajectory of a curve away from the critical point.

To illustrate our solutions we shall choose $A=-0.05$, although we do emphasise that the choice of A is completely arbitrary. By choosing the magnitude of A so small we ensure that initially we do not depart too far from the isotropic solution, enabling us to add the viscosity, at the surface obtained by the above expansion, $\xi=\xi_0$ say, without any detrimental effects. [Note that we also have to choose the value of either s or m on the surface $\xi=\xi_0$. Given A our procedure is then to choose a value of s, not too different from unity, use (4.65) to determine m and then use the equations of BH1 to obtain the values of the other parameters of the solution.] We are now in a position to integrate the viscous, self-similar, stiff field equations. Figures



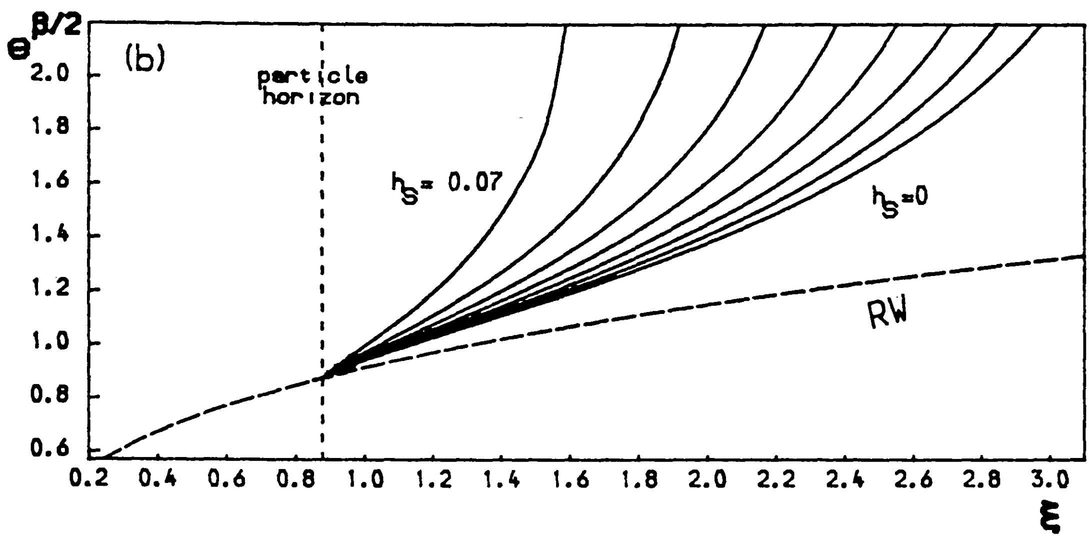


Figure 4.12 Behaviour of the scale factors in the viscous stiff solutions, showing the Robertson-Walker curves and the ξ=constant surface corresponding to the particle horizon. In these diagrams the viscosity constant has the values, h_s=0,0.01,...,0.07.

4.12a,b illustrate the variation of the transverse and radial scale factors, respectively, for a large range of viscous solutions, (h_s=0-0.07), with a given expansion parameter A=-0.05. The solutions for the scale factors , S and $e^{\beta/2}$, have been extended into the Robertson-Walker region, (outside the particle horizon), using the analytical expressions of BH1 with $a_\sigma{}^4$=1/3, i.e.

$$\begin{aligned} S^3 &= 3^{3/2}\frac{\xi}{2} \quad , \\ e^{\beta/2} &= \left(\frac{4\xi}{\sqrt{3}}\right)^{1/3} \quad . \end{aligned} \tag{4.66}$$

We cannot extend the volume expansion smoothly into the Robertson-Walker region due to the difference in expansion rates between the anisotropic and isotropic models.

The stiff solutions have the property of being transversely closed, S has a maximum, and radially open, $e^{\beta/2}$ is monotonic. As we increase the viscosity on the starting surface, we find that the more viscous solutions expand more slowly in the transverse direction and more quickly in the radial direction.

The self-similar volume expansion, Φ, is also found to increase as the viscosity increases, see Figure 4.13a. The behaviour of the volume expansion in these stiff solutions is very different from that observed in the dust solutions. In the stiff solutions, Φ is always positive and initially decreasing, as in the open dust solutions, but also exhibits a turning point. This turning point is caused by the large second derivative which develops in the radial scale factor in these stiff solutions. In all of the stiff solutions there is a value of ξ, ξ_m say, for which the rate of expansion of two comoving observers away from each other reaches a minimum. Thus, we have a self-similar model in

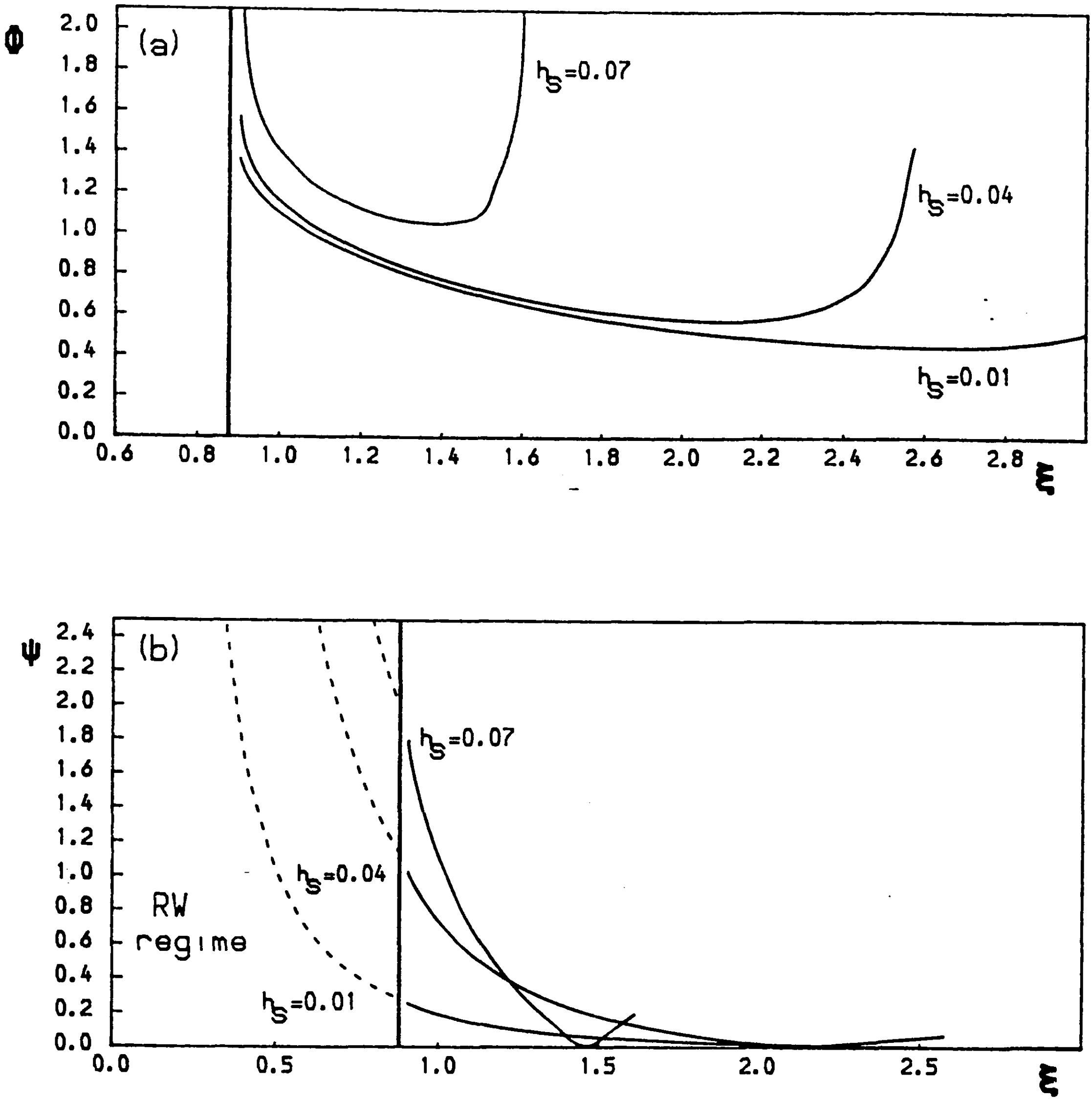


Figure 4.13 (a) Self-similar volume expansion, Φ, as a function of ξ, for the solutions h_s=0, 0.01, 0.04 and 0.07. (The h_s=0 solution is almost indistinguishable from the h_s=0.01 curve). The thick line indicates the particle horizon outside which (smaller ξ) the solution is Robertson-Walker. (b) Self-similar dynamic viscosity, Ψ, for the solutions h_s=0.01, 0.04 and 0.07. The particle horizon and the Robertson-Walker solutions are indicated by the thick line and broken curves, respectively.

which each observer sees the universe expanding for all time, (cf. closed dust solutions), and in which we can define a surface, $r_m = \xi_m$, at any given time t, where the rate of expansion of the fluid world lines is at a minimum. The spacetime itself, however, contracts in the transverse direction and the solution approaches the second hypersurface, V=1. BH1 have shown that this is not a black hole event horizon but represents a breakdown in the similarity symmetry of the problem, since on this hypersurface $e^{\alpha}(1-V^2)$ is finite. This effect also exists in the viscous solutions. In this chapter we are not concerned with the existence of black hole solutions but the generation of self-similar cosmologies and we will defer a discussion of progressing beyond the symmetry breakdown until the next chapter.

We should point out that, in the non-viscous solutions, boundaries exist which separate the (M,S)-plane into forbidden and allowed regions. In the general viscous models similar boundaries also exist but do not coincide exactly with those of BH1. The boundary obtained is the locus of all the maxima of the transverse scale factor, in any given class of solution, characterised by h_s; one maximum for each value of the arbitrary parameter A. In principle, therefore, we can determine this boundary for any viscous solution.

An interesting phenomenon occurs which was not encountered in the dust solutions. In the stiff solutions there exists a critical value of the viscosity constant, h_s, above which there are no physically meaningful solutions. If we write down the differential equation for the metric coefficient β in its most general form, cf. (4.51), viz.,

$$\beta' = \frac{[\; F(\xi_o,S_o,M_o,\gamma_o,\alpha_o,\beta_o) - G(\xi_o,S_o,M_o,\gamma_o,\alpha_o,\beta_o)h_s \;]}{[\; H(\xi_o,S_o,M_o,\gamma_o,\alpha_o,\beta_o) - K(\xi_o,S_o,M_o,\gamma_o,\alpha_o,\beta_o)h_s \;]} \quad , \qquad (4.67)$$

where F, G, H and K are all positive functions of the starting conditions and are therefore all constant, given the parameter A for any given h_s. We see immediately that there is a value of h_s, $(=H/K)$, for which an infinite gradient forms in the metric coefficient β and the solution breaks down. [For the starting conditions obtained from the Taylor expansion with $a_\sigma S=1.01$, $A=-0.05$ and $a_\sigma{}^4=1/3$, we find that an infinite gradient forms for $h_s=0.081$]. Thus, we can only consider solutions which have an initial dynamic viscosity up to a certain limit, η_{max}, the value of which depends on the initial conditions of the model.

For completeness, we also show the variation of the self-similar viscosity in Figures 4.13b and we see that the dynamic viscosity decreases from its initial value, $\eta_o(h_s)$, to zero on the surface where the transverse scale factor reaches its maximum, $(S'=0)$. As the initial viscosity increases, this surface moves outwards to larger values of the coordinate r, (smaller ξ). The dynamic viscosity begins to increase again as the universe starts to contract in the transverse direction due to the non-zero gradients in the scale factor S. This increase in the viscosity, which occurs as the transverse scale contracts towards the symmetry breaking hypersurface, V=1, causes the universe evolves into a highly anisotropic state. This could have a severe effect on the production of black holes in the early universe. We shall discuss this in more detail in the following chapter of this thesis.

4.6 Conclusions

In this chapter we discussed the role of dissipative forces in an anisotropic cosmological fluid, a topic which has been the subject of numerous investigations. We developed the Einstein field equations for a spherically symmetric geometry, in a comoving non-synchronous reference frame, which admitted a self-similar symmetry of the first kind and which described an imperfect fluid matter distribution.

The "source terms" of dissipative forces in a fluid can be of many kinds; electromagnetic caused by the presence of magnetic fields, heat conduction caused by temperature gradients, vorticity due to rotation of the fluid as a whole and viscous effects due to a resistance of the fluid to shearing or bulk motions. In this work we only considered the dissipative effects due to viscosity although this is by no means any more important than the other effects mentioned. Our choice was simply one of tractability. If we wish to consider similarity solutions of the field equations we cannot introduce any fundamental scales other than the constants, G and c. Therefore, introducing dissipative effects automatically destroys the self-similarity unless we can make some assumptions about the form that the dissipative effects take. Indeed, forcing the similarity symmetry to be of the first kind imposed certain restrictions on the form of the viscosity coefficients, both shear and bulk, such that the dynamic viscosity, η, was dependent upon the characteristic scales of the problem, (density ρ, scale factor R and velocity R_t).

Although we expect the viscosity coefficients of the anisotropic early universe to be functions of time, due to their dependence on temperature, Nightingale (1973), the exact form of the variation will, almost certainly, not be as chosen in equation (4.27). A more rigorous treatment of the problem of applying self-similarity methods to imperfect fluid cosmologies is required. As has been suggested previously in this chapter, this would entail resorting to a similarity symmetry of the second kind where the form of the self-similar variable is determined by the boundary conditions. Such an analysis was carried out with great success in the case of a perfect fluid cosmology with a non-zero cosmological constant, (cf. Henriksen et al. 1983 and Chapter 3 of this thesis). In this work, the cosmological constant acts as the 'additional' fundamental scale but can be dealt with by equating it to a vacuum energy density. However, no such solution is readily available in the case of a constant coefficient of shear viscosity. It is hoped that the work presented in this chapter, has provided some insight into the application of similarity techniques to imperfect fluid cosmologies and has provided a base from which to develop a self-consistent self-similar viscous cosmology. Such a development would be important for studies of the early universe where one expects anisotropic effects to play a major part in the evolution of the universe.

We restricted our considerations to the two extreme cases which we dubbed viscous dust, corresponding to $T^1{}_1=0$, and stiff, to $T^1{}_1=-T^0{}_0$. The most interesting solutions obtained were those of the open dust models, characterised by M>1, which are ever-expanding and therefore avoid the premature recollapse problem discussed by Barrow (1987). These models also had the desirable property that the anisotropy of the solutions is extremely small at small distances (or late times), making them viable descriptions of anisotropic universes which evolve to a state much like the observable universe at the present epoch.

Investigations of viscous solutions with different equations of state could be carried out by appealing to Bicknell and Henriksen (1978b) for the initial conditions. However, the purpose of this work was predominantly to provide the mechanism for obtaining self-similar imperfect fluid cosmologies and given the somewhat ad hoc nature of the assumptions used, we feel that the examples of the two extreme cases given are adequate.

When we come to discuss the growth of primordial black holes in the next chapter we will see that the stiff solutions have a more practical use than as an academic exercise.

5. FORMATION OF BLACK HOLES IN SELF-SIMILAR ANISOTROPIC UNIVERSES

5.1 Introduction

Black holes are normally thought of as being produced by the collapse of stars or possibly galactic nucleii. However, the existence of galaxies implies that there must have been some degree of inhomogeneity at all times in the history of the universe, and therefore one would also expect a certain number of black holes with masses from 10^{-5}g upwards to be formed in the early stages of the universe (Hawking 1971). These departures from homogeneity and isotropy could have been very large at early epochs and, even if they were small on average there would be occasional regions in which they were very large. One would therefore expect at least a few regions to become sufficiently compressed to overcome pressure forces and the velocity of expansion and collapse to a black hole. Such black holes are referred to as primordial.

The earliest time at which one can hope to apply classical general relativity is the Planck time $\simeq 10^{-43}$s. A black hole formed at this time would have an initial mass of about 10^{-5}g and radius 10^{-33}cm. For comparison, a black hole formed at the time of Helium formation when the temperature was 10^{9}K would have a mass of about 10^{7} solar masses.

One would expect that once the primordial black holes are formed they would grow by accreting nearby matter. The first estimate of the rate of accretion of matter onto a black hole in the early universe was made by Zel'dovich and Novikov (1967). They considered the accretion as a quasi-stationary process where the velocity of matter crossing the horizon (r_g=2M) is of order of the velocity of light. Further, they concluded that if the black hole was small compared to the particle horizon at the time of formation there would not be much accretion. On the other hand, if the black hole was of the order of the size of the particle horizon at the time of formation, the mass of the black hole would increase directly with time, i.e. $M \propto t$. In other words, the black hole would grow at the same rate as the particle horizon. Observations indicate that the universe is homogeneous on large scales and this suggests that the black hole could not continue to grow at this rate until the present epoch but would terminate at some much earlier time, say the end of the radiation era when there was no more radiation pressure to force matter into the black hole. If this were the case, the black holes would grow to a mass of 10^{15} to 10^{17} solar masses, the mass within the particle horizon at the end of the radiation-dominated era. The observational evidence, from the study of tidal motions in the Virgo Cluster (Van den Bergh 1969) and of fluctuations of the microwave background on small angular scales (Boynton and Partridge 1973), suggests that no such giant black holes exist in the universe at the present time.

Faced with this lack of observational evidence, Carr and Hawking (1974) argued that the assumption of a quasi-stationary accretion mechanism breaks down in the critical case of a black hole whose size is of the same order as the particle horizon. In this situation, the expansion of the universe has to be taken into account, a factor not included by Zel'dovich and Novikov (1967). Carr and Hawking (1974) proceeded to demonstrate that, in the $p=(^1/_3)\rho$ situation of a spherically symmetric universe, there is no solution to the field equations in which a black hole formed by purely local processes grows as fast as the particle horizon. (If the rate of accretion is insufficient in the spherical case to make a black hole grow at the same rate as the universe, it is reasonable to assume that it would also be insufficient in the more general non-spherical case since departures from sphericity would tend to decrease the rate of accretion). This negative result is proved by considering the properties of spherically symmetric similarity solutions of Einstein's equations. (A similarity solution is one in which all length scales increase with time at the same rate and is what would be required to represent a black hole growing as fast as the universe). If one also requires that the universe is exactly Robertson-Walker outside the local homogeneity, it can be shown that the same conclusion applies for any equation of state of the form $p=a^2\rho$, with a^2 positive and strictly less than unity, (Bicknell and Henriksen 1978b). This means that any black hole formed at an early epoch must soon be considerably smaller than the universe.

Hacyan (1979) proposed a model for primordial black hole growth in the early universe which consisted of a Vaidya sphere of radially ingoing photons expanding into a spatially flat, radiation dominated ($a^2=1/3$) Friedmann background. He found that the black hole grows in proportion to the horizon mass of the background. This result has been generalised to the case of arbitrary equation of state $p=a^2\rho$, cf. Cameron Reed and Henriksen (1980). These authors demonstrated that the general relativistic boundary conditions demand the Vaidya metric used by Hacyan (1979) to be self-similar. However, they also concluded that because the transition from a Friedmann gas to ingoing photons seems implausible, except possibly when $a^2=1$, the self-similar behaviour is probably unlikely to continue past the end of the hadron era. This suggests a solar mass limit for self-similar growth in accordance with earlier work, (Bicknell and Henriksen 1978a,b).

The proof of the non-existence of a black hole similarity solution only applies for a^2 strictly less than 1. However, there is still the possibility that $a^2=1$, corresponding to a stiff equation of state. (For a discussion of such an equation of state see Chapter 4 of this thesis). In fact, Lin et al. (1976) demonstrated that the Einstein equations do permit a similarity solution in this situation. This means that, if the universe ever did have a stiff equation of state, black holes which formed then might have grown as fast as the universe until the stiff era ended.

In discussing the cosmological consequences of black holes forming in a stiff era Lin et al. stress that it is somewhat doubtful whether primordial black holes can form naturally at all when $a^2=1$ since the Jeans length is then effectively the particle horizon size, i.e. the regions which can form black holes are necessarily nearly separate universes. However, if we do assume that the universe was stiff from the end of the Planck era to some time t_* we can conclude that there could be no primordial black holes smaller than $10^{15}(t_*/10^{-23}s)g$. The choice of this particular scaling will become more apparent below. We have already noted in the last chapter that the universe could not remain stiff after 10^{-4}s since strong interactions are unimportant then. This provides a very loose upper limit on the mass of the primordial black holes, of order of 10 solar masses, if we assume that there was not much accretion after the stiff era.

An important consequence of having $t_*>10^{-23}$s would be that no primordial black holes would remain small enough to evaporate within the lifetime of the universe (10^{17}s). Under the assumption that a primordial black hole does not accrete very much matter, a black hole of mass m will emit particles according to Hawking (1974, 1975) like a blackbody of temperature $10^{26}(m/1\text{g})^{-1}$K because of quantum effects. Thus primordial black holes with original mass smaller than about 10^{15}g would have evaporated by the present epoch.

Observations of the γ-ray background radiation indicate that black holes of around 10^{15}g must have an average mass density of less than 10^{-8} times the critical density required to close the universe (Chapline 1975, Carr 1976, Page and Hawking 1976). However, the γ-ray background limitation only applies if black holes of $\approx 10^{15}$g exist at the present epoch. If $t_* > 10^{-23}$s then we cannot rely upon this observation. The only upper limit one can place on the density of the primordial black holes then comes from measurements of the universe's deceleration parameter (Sandage 1972), which indicate that the total density of the universe cannot much exceed the critical density. Therefore, having a stiff equation of state before 10^{-23}s leaves open the possibility that primordial black holes have a critical density.

In this chapter we shall discuss black hole similarity solutions in a stiff universe and investigate the effects of anisotropy, in the form of viscous shear, on the formation black holes in the early universe.

5.2 Black Hole Similarity Solutions

In order to determine whether a black hole could accrete a significant amount of matter, Carr and Hawking (1974) introduced the use of similarity solutions of the Einstein field equations. Such solutions can describe black holes whose event horizon expands at a rate comparable to that of the universe particle horizon. Carr and Hawking considered equations of state $p=0$ and $p=(^1/_3)\rho$ for the primordial matter and came to a negative conclusion, namely that a primordial black hole cannot expand as quickly as the universe. Lin et al. (1976) have studied the case of a 'stiff' early universe wherein $p=\rho$ and have concluded that a black hole can accrete as rapidly as the particle horizon.

In light of a rather simpler formulation of the self-similar equations due to Cahill and Taub (1971) and Henriksen and Wesson (1978a), Bicknell and Henriksen (1978a) (hereafter BH1) have reconsidered the problem of black hole similarity solutions in a universe with a stiff equation of state. Their analysis supports the conclusion of Lin et al. (1976) indirectly, but differs substantially in the detailed description of accretion flow.

As in the case of the viscous similarity solutions we shall again work with a metric expressed in a comoving, non-synchronous reference frame by

$$\begin{aligned} ds^2 &= e^{\alpha}dt^2 - e^{\beta}dr^2 - R^2 d\Omega^2 \quad , \\ d\Omega^2 &= d\theta^2 + \sin^2\theta d\phi^2 \quad . \end{aligned} \tag{5.1}$$

Matter moves along the t-lines and the variables α, β and R are functions of r and t.

A similarity solution of the first kind is one in which the metric admits a homothetic Killing vector v^{μ}, that is,

$$\mathcal{L}_v g_{\mu\nu} = 2g_{\mu\nu} \quad , \tag{5.2}$$

where the left hand side is the Lie derivative of the metric tensor in the direction of v^{μ}. Lengths increase at the same rate along the orbits of the vector field v^{μ}, and this corresponds to the notion of a similarity solution that one encounters in hydrodynamics.

We shall use the form of the field equations for an imperfect fluid matter distribution given by equations (4.49)-(4.53) and for the moment we shall assume that the matter distribution is that of a perfect fluid. Consequently we shall put $h_s=0$.

In order to discuss black hole similarity solutions we consider the function

$$V = \xi^{-1} e^{(\beta-\alpha)/2} \quad , \tag{5.3}$$

introduced by Cahill and Taub (1971) and used by Carr and Hawking (1974). V represents the velocity of the ξ=constant hypersurfaces relative to the flow lines of the matter. These surfaces, which have equation $r=t/\xi_0$ (ξ_0=a constant), represent a family of spheres expanding through the matter. The induced metric on such a hypersurface is

$$ds^2 = e^{\alpha}(1-V^2)dt^2 - R^2 d\Omega^2 \quad . \tag{5.4}$$

Hence when $V>1$ the surfaces are space-like and when $V<1$ they are 'time-like'. If $V=1$ and e^{α} is finite, the ξ=constant hypersurface contains a null vector and is either an event horizon or a particle horizon. However, if e^{α} becomes infinite at $V=1$, one is required to calculate the limit $e^{\alpha}(1-V^2)$ as $V\to1$, in order to determine the nature of the hypersurface. The required behaviour of V, which has been demonstrated by Carr and Hawking (1974), is indicated in Figure 5.1. There are two values, ξ_1 and ξ_2 ($\xi_2>\xi_1$), at which $V=1$. The inner surface $\xi=\xi_1$ can be regarded as the particle horizon, i.e. it describes the outward propagation of light rays emitted from $r=0$, at the beginning of the universe, $t=0$. For $\xi_1<\xi<\xi_2$ the surfaces of constant

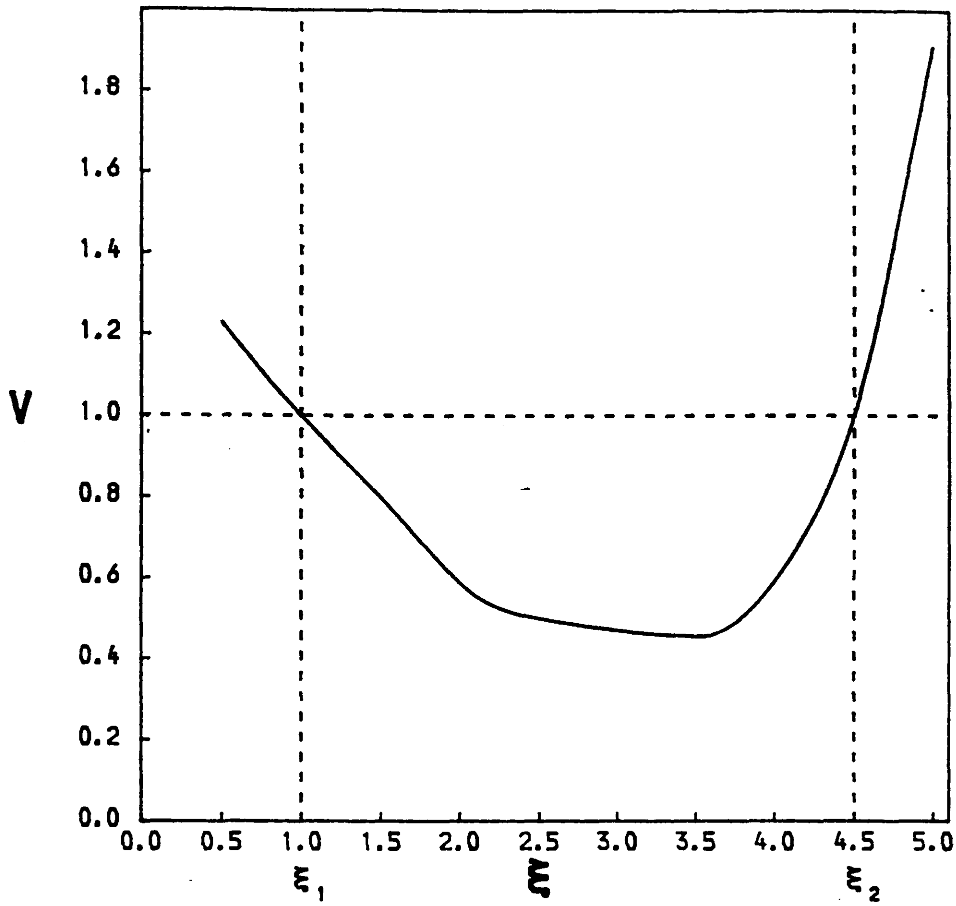


Figure 5.1 Necessary behaviour of the function $V=\xi^{-1}e^{(\beta-\alpha)/2}$ in order that a similarity solution represent a black hole in an expanding universe. The surfaces $\xi=\xi_1$ and $\xi=\xi_2$ are possible locations of the particle and event horizons, respectively.

ξ are time-like and it would therefore be possible for an observer in a rocket to remain in this region. However, should the rocket cross the null surface $\xi=\xi_2$, the surfaces of constant ξ would become space-like and it would inevitably hit the singularity, $V=\infty$, at $\xi=\xi_\infty$. This shows that the surface $\xi=\xi_2$ is the event horizon of the black hole.

When considering a perfect fluid matter distribution, $h_s=0$, we find that equations (4.52) and (4.53) yield integrals, which for a stiff equation of state ($a^2=1$), are given by

$$
\begin{aligned}
e^{-\alpha} &= \frac{1}{a_\sigma{}^4}\frac{\xi^2 y}{S^3} , \\
e^{-\beta} &= \frac{1}{a_\omega{}^4} S y ,
\end{aligned}
\tag{5.5}
$$

where a_σ and a_ω are constants of integration. We showed in Chapter 4 that for a stiff universe the constant a_ω was arbitrary but that a_σ was physically significant.

Equations (5.5) imply that, in the perfect fluid situation, the function V is given by

$$
V = \xi^{-1} e^{(\beta-\alpha)/2} = \left[\frac{a_\omega}{a_\sigma}\right]^2 \frac{1}{S^2} , \tag{5.6}
$$

so that, for a black hole solution, S is required to go through a_ω/a_σ, rise to a maximum, and then decrease through a_ω/a_σ again. Accordingly, we shall follow BH1 and look for solutions of the field equations which exhibit this behaviour.

These solutions were discussed in detail and extended to the non-ideal fluid case in Chapter 4. [It is easy to see that in the more general viscous situation, $h_s\neq 0$, no such integrals (5.5) exist and we cannot make use of the definition (5.6). However, in the viscous solutions the transverse scale factor does exhibit a maximum and has a similar behaviour to that of the non-viscous solutions]. In Chapter 4 we generated viscous self-similar solutions and found that the similarity symmetry broke down on the hypersurface corresponding to $\xi=\xi_2$. To discuss black hole solutions we have to understand the nature of this hypersurface in more detail.

BH1 found that this hypersurface, $\xi=\xi_2$, was not, in fact, an event horizon even though it corresponded to V=1. The reason for this is that the limit $e^{\alpha}(1-V^2)$, as $\xi\to\xi_2$, is always greater than zero. Thus the induced metric (5.4) on the hypersurface has signature (1,-1,-1) and the surface is time-like.

Although e^{α} is infinite on $\xi=\xi_2$, all the tetrad components of the Riemann-Christoffel tensor are finite. Hence the infinity in both e^{α} and e^{β} reflects a pathology in the coordinate system rather than a singularity in spacetime. BH1 conclude that the physical difficulties incurred by these solutions are associated with the extreme equation of state which makes it impossible for material flowing onto a black hole to become supersonic contrary to what is expected from accretion analysis, (Novikov and Thorne 1973). In view of this apparent impasse, wherein the only $p=\rho$ solution heading towards a black hole event horizon must break down at $\xi=\xi_2$, BH1 considered a patch to another solution which does continue to an event horizon and which preserves the self-similar symmetry. This result is in contrast to that of Lin et al. (1976) who claim that the second V=1 surface is an event horizon and the $p=\rho$ solutions can be continued beyond it. BH1 suggested that the incorrect conclusion arrived at by Lin et al. was due to those authors misinterpreting the behaviour of the self-similar energy density, ϵ.

It is found that in all of the stiff solutions, viscous or otherwise, the energy density, ϵ, vanishes as the solutions approach the hypersurface $\xi=\xi_2$, see Figure 5.2. This suggests that the external solution should be patched to an inner null fluid which might consist of photons, neutrinos or gravitons. The photons may arise naturally in a matter-antimatter symmetric early universe (Omnes 1969, Alfvén 1971), while the latter two cases represent the conversion of matter into ingoing neutrinos or gravitational radiation by an accreting black hole.

BH1 demonstrated that such a patch was possible and that the relevant spherically symmetric solution of the field equations describing infalling radiation is the advanced time form of the Vaidya (1951, 1953) solution and found that this null fluid does contain a black hole event horizon. Thus, with the continued self-similarity of the solution this means that BH1 have managed to embed a black hole which grows with the background universe. The question we would like to consider is: How does the presence of anisotropy in the form of shear affect this analysis?

5.3 Effect of Anisotropy on the Formation of Primordial Black Holes

The conditions under which black holes can form in the early universe have been discussed by a number of authors, beginning with Hawking (1971). There are various exotic cosmologies in which prolific black hole formation would appear to be inevitable - e.g. models in which the early universe is cold (Carr 1977) or tepid (Carr and Rees 1977). However, within the context of conventional cosmological models, it is clear that black holes could have formed prolifically only if the universe's initial density fluctuations had a very special form. Carr

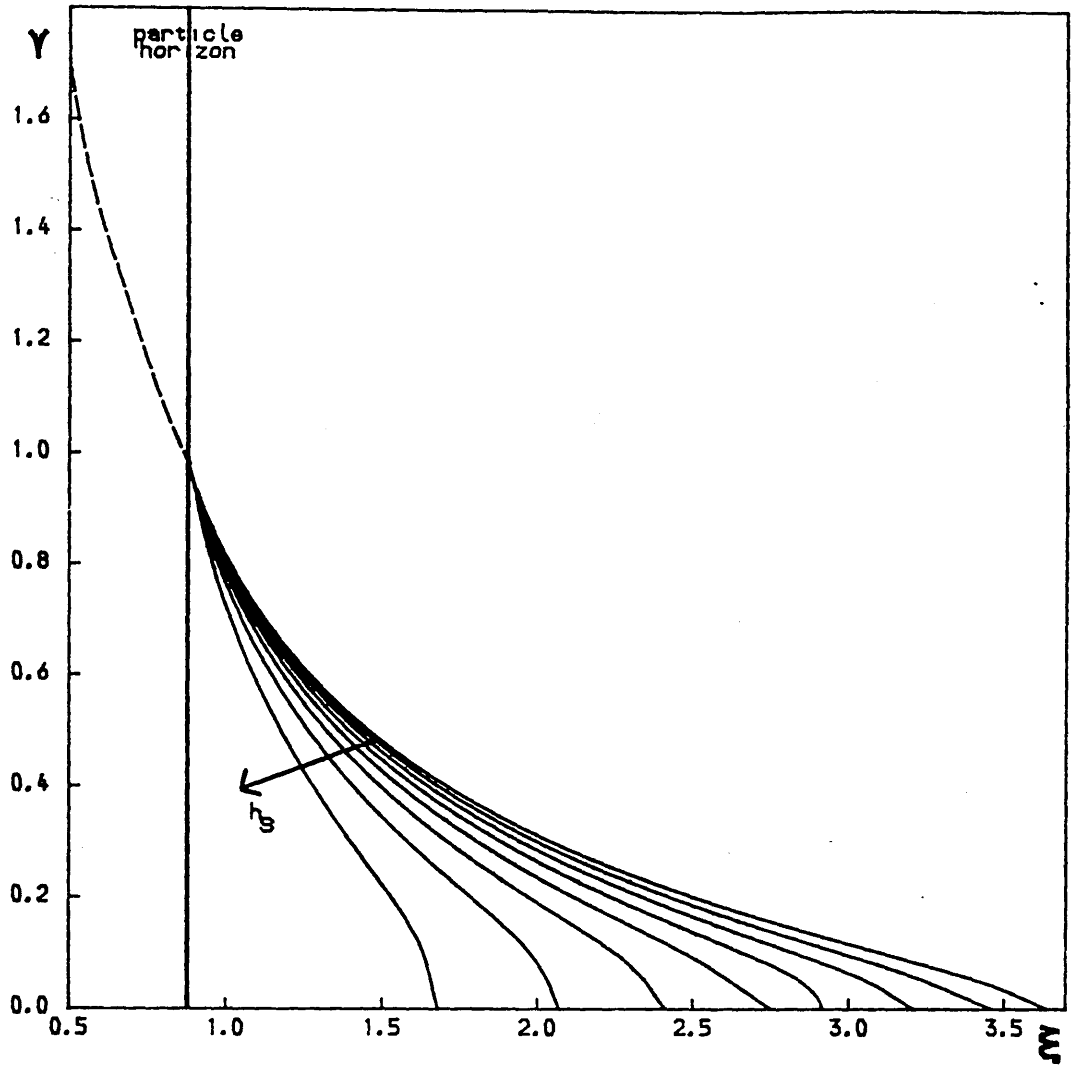


Figure 5.2 Evolution of the uniformity parameter, $\gamma = \epsilon S^3$, as the solution moves from the particle horizon to the second surface V=1 for the viscous stiff solutions, h_s=0,0.01,...,0.07. The broken curve indicates the Robertson-Walker behaviour outside the particle horizon.

(1975) has shown that in an isotropic universe, primordial black holes of mass M could have formed only if the fluctuations on a mass scale, M, had a carefully chosen amplitude, δ, when that scale first fell within the particle horizon.

Carr (1975) demonstrated that the black hole density at the present epoch, Ω_B, can be significant only if δ lies in a narrow range around 0.04. If δ was lower than 0.04 then Ω_B would be negligibly small and if δ was larger than 0.04 then Ω_B would be inconsistent with measurements of the cosmological deceleration parameter (Sandage and Hardy 1973). We can, therefore, argue that the early universe could not have been completely chaotic ($\delta \simeq 1$), thus the linear approximation, on which the considerations of Carr (1975) depend, is permitted. This argument is not rigourous, however, since extra qualitative features may enter the picture in the chaotic situation which may render any extrapolation, from the present-day Ω_B to Ω_B in the early universe, inappropriate.

In particular, one might enquire as to the effects of anisotropy (shear) in the early universe. Barrow and Carr (1977) reconsidered the work of Carr (1975) and discussed the effect of anisotropy on primordial black hole formation. They described two forms of anisotropy, primordial and induced (see Chapter 4), and found that the effect of induced shear is small but that primordial shear could have an important inhibiting effect upon black hole formation. In the situation of primordial shear, the anisotropy must dominate the dynamics of the universe at early enough times. Thus, a region which binds when the density of the universe is dominated by shear must collapse, together with an appreciable part of the shear energy it contains. This makes collapse more difficult since the shear provides an extra pressure against which gravity must battle. The assumption that the universe is initially shear-dominated is, in fact, more in the spirit of chaotic cosmologies than is the assumption that it starts off with large inhomogeneities but isotropic (Misner 1968).

If we demand that the solutions have a homothetic symmetry, we find, cf. Chapter 4, that the dynamic viscosity and hence the contribution of the shear to the matter distribution, must take a certain form. We thus treat the anisotropy in a somewhat different manner to Barrow and Carr (1977). In Chapter 4 we also defined an equation of state which was intrinsic to the anisotropic solutions, viz.,

$$P + \tau = a^2\epsilon \qquad , \tag{5.7}$$

where P, ϵ are the dimensionless effective pressure and energy density, respectively, a is the constant sound speed and τ is the self-similar shear parameter defined by equation (4.44). [This is equivalent to choosing an equation of state,

$$T^1{}_1 = -a^2 T^0{}_0 \qquad , \tag{5.8}$$

where $T^{\mu}{}_{\nu}$ is the energy-momentum tensor].

For our present purposes we are interested in black holes which grow as fast as the particle horizon and therefore require a stiff equation of state, $a^2=1$. Barrow and Carr (1977) have shown that the stiffness may be provided by anisotropy rather than strong interactions, which is the case in the conventional cosmological scenarios, (it might, of course, be provided by both effects). The equation of state (5.8) with $a^2=1$ shows that in the current analysis the stiffness is a property of the matter distribution as a whole, anisotropy included, with no *explicit* assumptions being made about the origin of such an equation of state. Our analysis will proceed along similar lines to that of BH1.

We saw in Chapter 4 that the viscous solutions displayed a similar behaviour to the non-viscous solutions of BH1 in that the transverse scale factor, S, reaches a maximum on some hypersurface, $\xi=\xi_m$, (ξ_m varies with each solution). When S contracts towards the V=1 hypersurface the solution encounters the problems described in the last section, i.e. the metric coefficients, e^{α} and e^{β}, tend to infinity as $V\to 1$ while the tetrad components of the Riemann-Christoffel tensor remain finite. Figure 5.3 illustrates the infinity in e^{α} as $V\to 1$ for the viscous solution, $h_s=0.02$, and the non-viscous solution, $h_s=0$. (We have continued the solutions into the Robertson-Walker regime). The figure shows that in both solutions the metric coefficient e^{α} rapidly tends to infinity as the hypersurface V=1 is approached. The hypersurface, V=1, occurs further out from the origin, r=0, in the viscous solutions than in the non-viscous solutions. This is due to the presence of shear causing the transverse scale to reach its maximum 'sooner', by slowing down the expansion. In light of this coordinate breakdown, we therefore have to attempt to patch the viscous models to some other solution which preserves the self-similarity and may contain a black hole.

Figure 5.2 shows that in all the viscous solutions, characterised by the viscosity constant h_s, the energy density, ϵ, vanishes as the solutions approach the V=1 hypersurface. This suggests that we should again try matching the external solution to an inner null fluid. In order to achieve this matching, we must find a common admissible set of coordinates for the $P+\tau=\epsilon$ fluid and for the null fluid, (see, e.g., Synge 1961). We shall follow the procedure of BH1, with $a_{\sigma}{}^4=(1/3)$ and $a_{\omega}=1$.

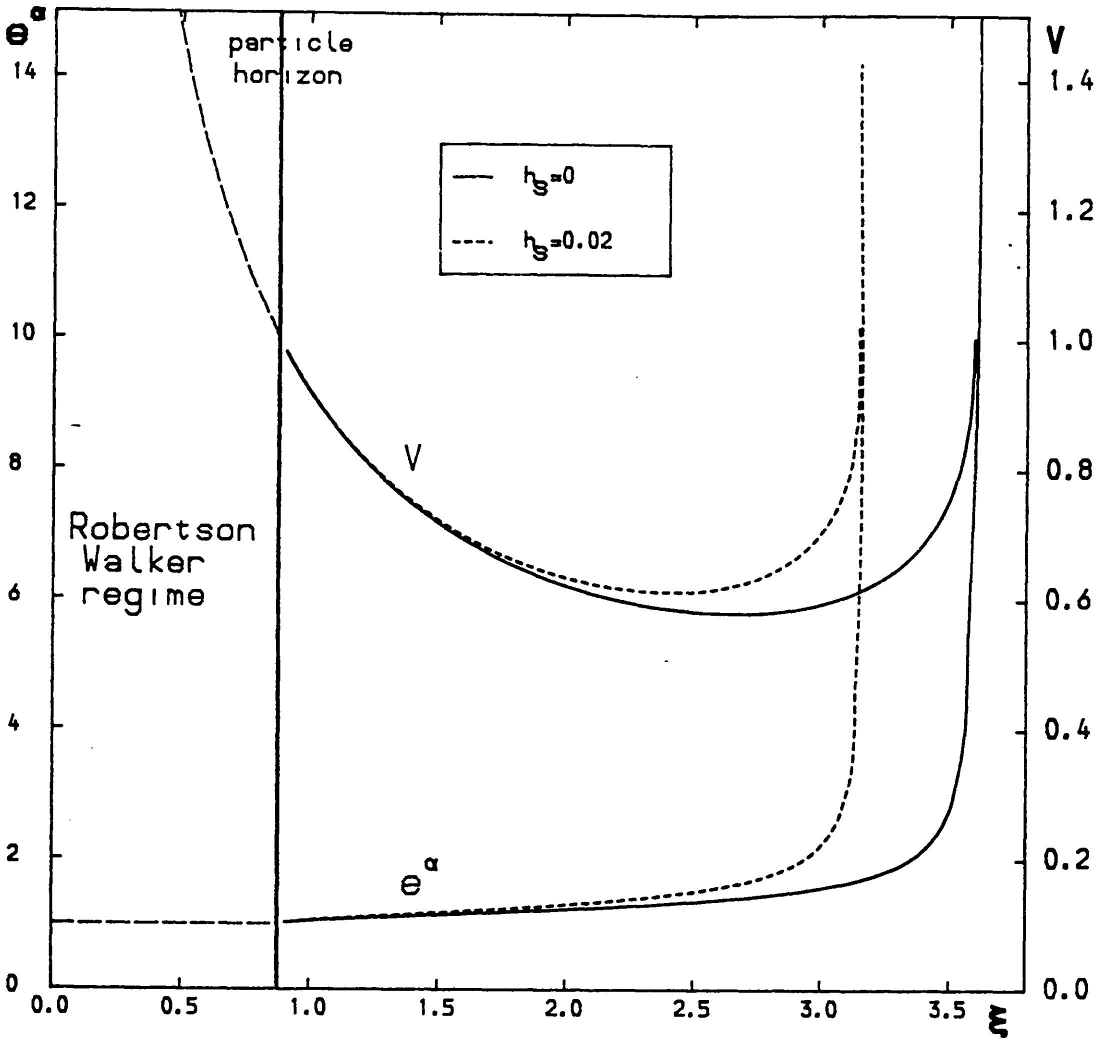


Figure 5.3 Variation of the velocity function V and the metric component e^{α}, demonstrating the coordinate breakdown encountered as the solution approaches the surface V=1. The particle horizon and Robertson-Walker regime are also indicated.

We begin by defining an advanced time coordinate, w, given by,

$$w = \lambda r \exp\left[\int_{\xi_2}^{\xi} \frac{d\xi}{\xi(V+1)}\right] \quad , \qquad (5.9)$$

where λ is an arbitrary constant and the function w is constant along ingoing null geodesics. Expressing the metric in terms of the new coordinate w and the transverse scale factor $R=rS(\xi)$, we obtain

$$ds^2 = g_{00}dw^2 + 2g_{01}dwdR - R^2d\Omega^2 \quad , \qquad (5.10)$$

where it can be demonstrated that,

$$g_{00} = \frac{(V+1)^2\xi^2r^2}{w^2} \frac{\left[e^{\alpha}(S-\xi S')^2 - e^{\beta}S'^2\right]}{\left[\xi VS' - (S-\xi S')\right]^2} \quad , \qquad (5.11)$$

and

$$g_{01} = -\frac{\xi r(V+1)}{w} \frac{\left[e^{\alpha}(S-\xi S')\xi V - e^{\beta}S'\right]}{\left[\xi VS' - (S-\xi S')\right]^2} \quad . \qquad (5.12)$$

Here S, ξ, α and β are to be regarded as implicit functions of w and R. If we now consider a fluid trajectory (dr=0) in this new coordinate system (w,R), we can show without too much difficulty that

$$\frac{dw}{dR} = \frac{w}{\xi(V+1)rS'} = \frac{2w}{(V+1)r}\frac{\gamma}{(\gamma-M)S} \quad .$$

Hence as $V \to 1$, ($w \to \lambda r$, $M \to M_2$, $S \to S_2$ and $\gamma \to 0$), $dw/dR \longrightarrow 0$ and the fluid trajectories become null geodesics, justifying the statements made above.

The metric (5.10) is only valid up until the hypersurface V=1, that is, the hypersurface

$$w = \frac{\lambda R}{S_2} \quad , \qquad (5.13)$$

as $w=\lambda r$ and $R=r/S_2$ at $\xi=\xi_2$. We must therefore match to an ingoing null fluid across this hypersurface.

As was stated in the last section, the relevant spherically symmetric solution of the field equations describing infalling radiation is the advanced time form of the Vaidya (1951, 1953) solution,

$$ds^2 = [1-2m(w)/R]dw^2 - 2dwdR - R^2d\Omega^2 \quad , \qquad (5.14)$$

where m is a function of the advanced time coordinate w. The only non-zero component of the Ricci tensor is

$$R_{00} = \frac{-1}{R^2}\frac{dm}{dw} \quad , \qquad (5.15)$$

and the null vector $W_\mu = (w_0,0,0,0)$ representing the infalling radiation is given by

$$w_0 = \frac{c^4}{8\pi G}\frac{1}{R^2}\frac{dm}{dw} \quad . \qquad (5.16)$$

By constraining the metric to admit a self-similar symmetry, i.e. to satisfy equation (5.2), we find that m(w) must be of the form

$$m(w) = bw \quad , \qquad \text{(b constant)} \quad .$$

Thus we have the self-similar metric

$$ds^2 = [1-2bw/R]dw^2 - 2dwdR - R^2d\Omega^2 \quad , \qquad (5.17)$$

which is to be matched to the $P+\tau=\epsilon$ metric across the hypersurface $w=\lambda R/S_2$.

To accomplish, this we first determine the limits of g_{00} and g_{01} as $V\to 1$. Thus we must find the asymptotic limits of the functions:

$e^{\alpha}(S-\xi S')^2$, $e^{\beta}S'^2$, $\xi S'$, $e^{\alpha}(S-\xi S')$ and $e^{\beta}S'$, cf. expressions (5.11) and (5.12) for g_{oo} and g_{o1}, respectively. Using the differential equation for the transverse scale factor S, equation (4.39), and the definition of V, equation (5.3), we obtain,

$$\lim_{V\to 1}(g_{oo}) = \frac{4}{\lambda^2 M_2} \lim_{V\to 1}(\gamma e^{\beta}) \quad ,$$
$$\lim_{V\to 1}(g_{o1}) = -\frac{2}{\lambda M_2 S_2} \lim_{V\to 1}(\gamma e^{\beta}) \quad , \tag{5.18}$$

where M_2 and S_2 are the values of M and S at V=1, $(\xi=\xi_2)$.

If we now consider the energy equation (4.41), substituting $e^{\alpha}=e^{\beta}/\xi^2V^2$, we find that

$$\lim_{V\to 1}(\gamma e^{\beta}) = \frac{S_2{}^3 M_2}{(S_2-M_2)} \quad . \tag{5.19}$$

Consequently, we note that equation (5.18) now becomes

$$\lim_{V\to 1}(g_{oo}) = \frac{4S_2{}^3}{\lambda^2(S_2-M_2)} \quad ,$$
$$\lim_{V\to 1}(g_{o1}) = -\frac{2S_2{}^2}{\lambda(S_2-M_2)} \quad . \tag{5.20}$$

Hence if we choose $\lambda=2S_2{}^2/(S_2-M_2)$, g_{o1} is clearly continuous by (5.17) and the value of g_{oo} is

$$g_{oo} = 1 - M_2/S_2 \quad . \tag{5.21}$$

Comparison with the Vaidya value of g_{oo} on $w=\lambda R/S_2$, namely

$$g_{oo} = 1 - \frac{4S_2 b}{(S_2-M_2)} \quad ,$$

shows that if

$$b = \frac{M_2(S_2 - M_2)}{4S_2^{\,2}} \quad , \tag{5.22}$$

g_{00} is continuous. Note that on V=1 all the radiation starts with w>0; and since it falls along lines of constant w, we consider only the region of spacetime (5.10) for which w>0. We should also notice from equation (4.35), for the self-similar dynamic viscosity, that as V→1 (γ→0),

$$\lim_{V\to 1}(\Psi) = \frac{h_s}{S_2^{\,2}} \lim_{V\to 1}(\gamma|S'|) = \frac{h_s M_2}{2\mathcal{E}_2 S_2} \quad , \tag{5.23}$$

which is non-zero for all viscous solutions, $h_s \neq 0$. Thus we have to invoke some sort of phase change on the surface V=1 where by converting the matter of the external solution to ingoing neutrinos or gravitational radiation the viscous properties are "lost".

The induced metric on hypersurfaces of constant Z=w/R is

$$ds^2 = [Z^2(1-2bZ) - 2Z]dR^2 - R^2 d\Omega^2 \quad .$$

with b given by (5.22). Now, defining

$$F(Z) = 2bZ^2 - Z + 2 \quad , \tag{5.24}$$

we see that Z=constant is timelike if F(Z)<0, null if F(Z)=0 and spacelike if F(Z)>0. There are apparently two null hypersurfaces:

$$Z = \frac{1 \pm (1-16b)^{1/2}}{4b} \quad . \tag{5.25}$$

However, the smaller of the two values lies outside the Vaidya coordinate patch. Thus the Z-surfaces are timelike until

$$Z = \frac{1 + [1 - 4M_2(S_2-M_2)/S_2^{\,2}]^{1/2}}{M_2(S_2-M_2)/S_2^{\,2}} \quad , \tag{5.26}$$

where we have substituted for b, at which surface we encounter an event horizon. With further increase in Z we encounter the surface (g_{00}=0),

$$Z = \frac{2S_2{}^2}{M_2(S_2-M_2)} \qquad ,$$

which is an apparent horizon (cf. Hawking and Ellis 1973). As Z becomes infinite, that is, as R→0, the infalling radiation approaches the black hole singularity.

A rough numerical analysis shows that the hypersurface corresponding to the event horizon, (5.26), occurs at larger and larger values of Z as the initial viscosity, h_s, of the external solution increases. Thus the more viscosity present initially, the closer to the origin, R=0, is the event horizon.

We should point out that the above analysis is not rigorous, in the sense that we have only shown that the metric coefficients are continuous across the patching hypersurface. To satisfy the junction conditions of Synge (1961) we must also require continuity of the first derivatives of the metric.

From equations (4.42) and (4.43), we see that if the derivatives of the metric are to be continuous, the terms containing the shear parameter, τ, must vanish in the limit as V→1. Unfortunately, we find from the differential equations that the terms due to the presence of viscosity are proportional to $1/\gamma$ as V→1 (γ→0). Therefore, we cannot satisfy the junction conditions of Synge (1961) if we attempt to patch from an external viscous fluid to an inner non-viscous, null fluid.

5.4 Conclusions

In this chapter we have investigated the formation of black holes in an early universe with a high degree of anisotropy, (in the form of shear), and a stiff equation of state. The anisotropy was introduced in such a fashion as to maintain the self-similar symmetry of the solutions, i.e. the dynamic viscosity was chosen to have a functional form which was dependent on the characteristic scales of the problem. It was demonstrated that the anisotropic similarity solutions display a similar behaviour to that of the non-viscous similarity solutions found by Bicknell and Henriksen (1978a). The solutions contain two hypersurfaces on which the velocity of ξ=constant surfaces relative to the fluid equals the speed of light.

The most important correspondence between these solutions was the development of a coordinate singularity on the outer hypersurface V=1, (the inner V=1 surface corresponds to the universe particle horizon). We were thus forced to extend the spacetime beyond this hypersurface by patching to another solution which preserved the self-similar symmetry. In light of the vanishing energy density on this surface, we attempted to match the external solution to an inner null fluid.

The appropriate solution of the field equations describing infalling radiation was found to be the advanced time form of the Vaidya solution (1951, 1953) which, for simplicity, we took to represent a perfect fluid cosmology, cf. BH1. However, it was found that such a patch, from a viscous external solution to an inner (non-viscous), null fluid could not satisfy the junction conditions of Synge (1961), whereby the only allowable discontinuities are in the second (or higher) derivatives of the metric. It was found that the first derivatives of the metric could not be continuous when patching to a non-viscous spacetime from a spacetime with a non-zero viscosity. We are thus faced with the situation that black hole similarity solutions with a stiff equation of state only exist in a non-viscous (perfect fluid) universe.

A possible solution to this apparent restriction may be to patch the external solution obtained in §5.3 to an inner null fluid which is not described by a perfect fluid matter distribution. The radiation-like imperfect fluid cosmologies of Coley and Tupper (1985) provide a description of such a fluid. Indeed, they also show that if the velocity four-vector is chosen to be "tilting" in such a way that it has a spacelike component in the radial direction, their solutions admit a homothetic Killing vector and are therefore self-similar of the first kind. However, the solutions described in this work correspond to a conformally flat representation of the FRW metric and as such it is unlikely that they would produce black hole solutions, cf. Carr and Hawking (1974).

In light of the comments made in the last chapter regarding the rather ad hoc assumptions about the form of the dynamic viscosity in self-similar solutions (first kind), we should not rule out the possibility that black hole similarity solutions may exist in an anisotropic early universe. If the viscosity was modelled in a self-consistent manner and was introduced as an additional fundamental scale any similarity solutions found would necessarily be of the second kind. Such an analysis should be performed before we can make any conclusive statements about the growth of primordial black holes in an anisotropic early universe.

6. FUTURE WORK

6.1 Review

The aim of this thesis has been to investigate the possibilities of obtaining cosmological solutions of the Einstein field equations which admit a self-similar symmetry. It was demonstrated in Chapter 2 that there are two different kinds of self-similar solutions. Self-similarity of the first kind possesses the property that the similarity exponent, which defines the appropriate self-similar variable for the solution, is determined by dimensional considerations or from the conservation laws. As far as cosmological problems are concerned, if the only dimensional constants present in the model are the constant of gravitation, G, and the speed of light, c, then the solution may admit a self-similar motion of the first kind. However, if the cosmological model contains any additional dimensional constants, then the independent scale lengths introduced destroy the simple self-similar symmetry and the similarity solution, if it exists, is necessarily of the second kind. In self-similar problems of the second kind, the similarity exponent cannot be found from dimensional considerations or from the conservation laws without solving the equations.

Two distinct classes of cosmological models were considered: (i) an anisotropic, perfect-fluid solution to the Einstein field equations with a non-zero cosmological constant ($\Lambda \neq 0$) and (ii) an anisotropic viscous-fluid solution with $\Lambda = 0$. In situation (i), the presence of the additional scale length, Λ, destroys the possibility of obtaining a simple self-similar solution. However, as was discussed in Chapter 3, it was found by Henriksen et al. (1983) that by identifying the cosmological constant with the energy density of the vacuum, ρ_V, and treating ρ_V as a strict constant then the equations admit a self-similar symmetry of the second kind with a suitable form of the self-similar variable, ξ. This class of self-similar solutions was found to contain asymptotically ($\xi\to\infty$) de-Sitter solutions which, therefore, extend the cosmological "no-hair" theorems of Wald (1983).

The viscous-fluid solutions were considered in Chapter 4 and it was shown that with an appropriate form for the viscosity coefficients the solutions could admit a self-similarity of the first kind. Two classes of solutions were considered for which the equation of state was given by $T^1{}_1=0$ (viscous dust) and $T^1{}_1=-T^0{}_0$ (stiff), respectively. The stiff viscous-fluid solutions were then considered as a class of black hole similarity solutions in Chapter 5 and the effect of anisotropy on these solutions was discussed.

In the Appendix, the geometric symmetry corresponding to a self-similarity of the second kind was considered. It was found that, although a similarity of the first kind could be identified with a homothetic (or conformal) motion (see Chapter 2), self-similar symmetry of the second kind, in general, had no such simple analogue. (It should be noted, however, that a special class of these self-similar solutions did, in fact, admit a conformal symmetry). The existence of a more complex geometrical analogue to self-symmetry of the second kind has yet to be investigated.

6.2 Asymptotic Behaviour of Monotonic Self-Similar Solutions

Starobinskii (1983) demonstrated that if the energy-momentum tensor of the matter contains a positive cosmological term, $T^k{}_i=\rho_V\delta^k{}_i$, where $\rho_V>0$ is the energy density of the vacuum, then it is possible to construct the asymptotic structure of inhomogeneous cosmological expansion. This analysis is valid whether the cosmological constant is a true constant (ρ_V=constant) or is only an effective constant ($\rho_V\approx$constant over a certain time interval). The second case is important for applications in which the universe passed through some quasi-de-Sitter stage during the early stages of its evolution.

Starobinskii considered an empty matter distribution so that the only contribution to the curvature of the spacetime was made by the cosmological vacuum term. He obtained an asymptotic form as $t\to\infty$, given by

$$\begin{aligned} ds^2 &= dt^2 - \gamma_{\alpha\beta}dx^\alpha dx^\beta , \\ \gamma_{\alpha\beta} &= e^{2Ht}a_{\alpha\beta} + b_{\alpha\beta} + e^{-Ht}c_{\alpha\beta} + \ldots , \end{aligned} \tag{6.1}$$

where $a_{\alpha\beta}$, $b_{\alpha\beta}$, $c_{\alpha\beta}$ are functions of three spatial coordinates and $H^2=8\pi G\rho_V/3$. This solution is similar to the quasi-isotropic solution of Lifshitz and Khalantnikov (1963) but differs from it by the large number of physically different, arbitrary functions and by the fact that it is an expansion near $t=\infty$, rather than at $t=0$. The solution (6.1) contains four physical arbitrary functions of three coordinates and, therefore, is a general solution since four is the maximum number of arbitrary functions possible in this case. This solution is also stable relative to perturbations that are not too large. Two physically arbitrary functions are contained in $a_{\alpha\beta}$ and two in $c_{\alpha\beta}$. (Three functions in $a_{\alpha\beta}$ can be eliminated by three transformations of spatial coordinates, not including time, and a fourth function is eliminated by a transformation which preserves the synchronism of the coordinate system).

Starobinskii found that rapid local isotropisation with expansion is a typical phenomenon in the presence of a cosmological constant,

$\Lambda=3H^2$ and so the spacetime inside a constant physical volume rapidly approaches a de-Sitter spacetime, and the initial conditions are forgotten. Thus, the cosmological constant is capable of eliminating all types of inhomogeneities over very large scales. After the decay of the effective cosmological constant and the end of the quasi-de-Sitter state (6.1), perturbations begin to grow once again. However, if the phase (6.1) lasted for a sufficiently long time ($\approx$60-70 Hubble times, in practice) then homogeneity and isotropy of the observed part of the universe would not have had sufficient time to break down by the present epoch.

Lifshitz and Khalantnikov (1963) stated that a criterion for the generality of a solution is the number of arbitrary functions of the spatial coordinates it contains. It should be noted that among the arbitrary functions contained in any cosmological solution there are those whose arbitrariness is connected with the arbitrariness of the reference frame. (The greatest possible number of arbitrary functions in an arbitrary reference frame is 20). More important is the number of "physically arbitrary" functions which cannot be reduced by any choice of reference frame. The solutions of Starobinskii (1983) contain the maximum number of physically arbitrary functions of three coordinates and are therefore completely general. It would be interesting to relate the asymptotic form of the 'open' (asymptotically de-Sitter) solutions found in Chapter 3 to the general asymptotic series of Starobinskii. The immediate difficulty, although not insurmountable, is the non-synchronicity of the self-similar solutions, which prevents any direct comparison with those of Starobinskii. To first order the asymptotic solutions do tend to a synchronous system and indeed to a de-Sitter phase. However, to answer the question regarding the generality of the solutions, a higher order analysis should be carried out.

A further use of the asymptotically de-Sitter solutions of Chapter 3 might be to provide a means by which to progress smoothly from an early universe model to a later universe model. The presence of a large vacuum energy density in these solutions requires that they are, necessarily, early universe solutions. However, it is possible that these self-similar solutions of the second kind could be patched, via a phase transition in which the cosmological term disappears, to a self-similar solution of the first kind, which could be used to describe the later epochs of the universe. The possibilities of such a smooth transition between these different stages of evolution was briefly discussed by Wesson (1986b), for the special case of the singular solution of Henriksen et al. (1983) (cf. Chapter 3). However, it was found by Alexander and Green (1988), that in the more general case, the severe restrictions imposed by the symmetry and the continuity conditions make a patch possible only at a fixed time, $t=t_0$. Further, the lack of freedom in the choice of the equation of state, caused by the symmetry being of the second kind, prevents us from obtaining a solution which is self-similar *and* which satisfies the continuity conditions across the whole $t=t_0$ hypersurface.

It is possible to patch to a solution which does not have the restriction of self-similar symmetry, again via a phase change to effectively remove the large (unobserved) cosmological constant. Such a possibility is made more interesting by the fact that the spatial inhomogeneity inherent to the general solutions of Chapter 3, may allow separate regions of the 'patched solution' to collapse, as seeds for galaxy formation, while the general background spacetime expands (a situation not possible with similarity solutions since all length scales expand at the same rate). Thus, the existence of a patch from an early universe ($\Lambda \gg 0$) self-similar solution to a spatially-inhomogeneous, $\Lambda \simeq 0$, solution is worth a more detailed investigation.

6.3 Applications of Self-Similar Symmetry of the Second Kind

In Chapter 2 of this thesis, it was found that in self-similar problems of the second kind the initial conditions of the problem contain a dimensional parameter with the units of mass but lack a *unique* parameter which contains only the units of length and time, i.e. either no such parameter exists in the problem or there is more than one such parameter. In this case, no unique dimensionless (self-similar) variable can be formed from the initial conditions, but has to be solved for as part of the solution.

It was demonstrated in Chapter 3 that the 'additional' scale, Λ, could be absorbed into the equations as a vacuum energy density and the similarity variable could be obtained. However, in order to describe the imperfect fluid solutions of Chapter 4 as self-similar motions the 'additional' scales, i.e. the viscosity coefficients of bulk and shear, had to be of a certain, somewhat restrictive, form, (cf. equation (4.27)). A much more realistic procedure would be to treat the viscosity coefficients as constant and allow the equations to determine the self-similar symmetry. Such a symmetry would necessarily be of the second kind.

The dimensional constants present in a cosmological model with a viscous matter distribution are, the gravitational constant G (units $M^{-1}T^{-2}L^{3}$), the speed of light c (LT^{-1}) and the coefficients of shear and bulk viscosity η, ζ (both with units $ML^{-1}T^{-1}$). The gravitational constant G (or one of the viscosity coefficients) immediately gives us a constant which contains the units of mass. However, three independent dimensionless variables can be formed from the remaining three constants and the independent variables r and t, namely

$$\xi_1 = \frac{r}{ct} \quad , \quad \xi_2 = \frac{G}{c^3}r\eta \quad , \quad \xi_3 = \frac{G}{c^3}r\zeta \quad . \qquad (6.2)$$

Therefore, we cannot define a unique similarity exponent from dimensional considerations alone, and so we cannot obtain a self-similar solution of the first kind.

We could consider the effects of bulk and shear viscosity separately. For instance, if we were to choose an isotropic model so that the solutions were shear-free then we could produce a self-similar solution of the first kind, i.e. the M=1 dust solutions of Chapter 4. In that chapter we chose a specific form for the viscosity coefficients so that we could treat the equations as self-similar of the first kind. However, this choice was fairly arbitrary, with the only restrictions being that the viscous coefficients must remain positive throughout the solution and that their chosen form did not involve the introduction of any additional scales. On purely dimensional grounds it would also be possible to consider the forms

$$\left[\frac{\sqrt{G}}{c^2} \right]\eta = h_s\rho^{1/2} \quad , \qquad (6.3a)$$

and

$$\left[\frac{G}{c^3} \right]\eta = h_s/R \quad , \qquad (6.3b)$$

where ρ is the characteristic mass density of the fluid and R is the characteristic scale size for the cosmological model under consideration. Both of these encounter the difficulty discussed in Chapter 4, i.e. as $R\to 0$, $\eta\to\infty$. However, (6.3a) is a more realistic choice in the sense that the viscous coefficients depend only on the conditions of the fluid. Such forms for the viscous coefficients will be considered at a later date. For the moment, we are interested in a self-consistent viscous model in which we do not make any ad hoc assumptions about the variation of the viscosity coefficients. [Note that, since the bulk term can be absorbed into the pressure, we shall concentrate on the shear viscosity and assume that the bulk coefficient , $\zeta=0$].

One possibility would be to try a similar procedure to that used for the cosmological constant in Chapter 3, i.e. equate the shear term to some sort of 'viscous pressure', p_s (with a corresponding 'viscous energy density', ρ_s) and write the total pressure and energy density as

$$\begin{aligned} p &= p_m + p_s \quad , \\ \rho &= \rho_m + \rho_s \quad , \end{aligned} \tag{6.4}$$

where p_m, ρ_m are the normal matter components (there are no vacuum terms present). An equation of state for the viscous terms is required and the simplest form would be $\rho_s=a_s{}^2p_s$, where a_s is assumed to be a dimensionless constant.

The Einstein equations, $G_{\mu\nu}=-8\pi T_{\mu\nu}$ with $T_{\mu\nu}$ given by equation (4.12), would then be of the form

$$G_{\mu\nu} = -8\pi[(\rho+p)u_\mu u_\nu - pg_{\mu\nu}] \quad , \tag{6.5}$$

but with the viscous and normal matter terms given their separate equations of state. Because of the form of the shear tensor (4.15), the

(0,0) component of (6.5) implies that $a_s{}^2=0$ ($\rho_s=0$). Unfortunately, the other components of this equation give contradictory results for the viscous pressure, p_s, except in the case when $\sigma_{\mu\nu}=0$ and the solution is isotropic. Thus, it seems clear that the shear viscosity cannot be incorporated into the equations in the same way as the cosmological constant.

A more promising line of investigation may be the study of non-self-similar viscous solutions which become self-similar in the limit as $t\rightarrow\infty$. Such a behaviour is likely to be more common in nature than the exactly self-similar solutions. A characteristic of asymptotically self-similar solutions is that the system 'forgets' the initial conditions of the problem at some stage of its evolution. This type of behaviour is common in the "no-hair" theorems, of Hawking and Moss (1982) and Wald (1983), and the inflationary models, of Guth (1981) and Linde (1982), discussed in Chapters 2 and 3. These models were developed in order to explain, among other things, the large degree of isotropy and homogeneity observed in the present-day universe. However, the series of papers by Coley and Tupper (1983, 1984, 1985) have questioned this observed isotropy and have shown that an anisotropic matter distribution could also be a viable description for the matter content of a Friedmann-Robertson-Walker spacetime. Thus, it may be possible that an initial highly-anisotropic universe could pass through some inflationary epoch but still retain a significant amount of its pre-inflationary anisotropy. Martinez-Gonzalez and Jones (1986) considered the role of primordial shear in two inflationary scenarios (Linde inflation and GUT inflation). They found that in the case of Linde inflation the universe becomes truly isotropic but in the case of GUT inflation the initial anisotropy reduces the GUT era coherence length and it becomes more difficult to form the present universe from a single 'isotropic' bubble. Such a situation may allow the system essentially to 'forget' its initial conditions and it would then be possible to treat the problem as an asymptotically self-similar solution. By investigating the asymptotic limits, either analytically or numerically, of this or some other non-self-similar anisotropic model, which evolves to a self-similar regime, we could, in principle, determine the similarity exponent and therefore determine a self-similar variable which would define the symmetry. We could then treat the model as a similarity solution of the second kind.

Similarity solutions of the second kind could also be useful in the study of Kerr black holes. The rotation inherent to such a system acts as an extra degree of freedom and therefore introduces an additional dimensional parameter which would destroy the possibility of any self-symmetry of the first kind. However from the above discussion, we see that it may be possible to discuss this system in terms of a similarity solution of the second kind. Such a solution, if it exists, may provide some interesting developments in the physics of rotating cosmological systems.

Finally, similarity solutions (of first or second kind) may also prove useful in the study of power-law singularities in cosmological models, (cf. Wainwright 1984). Power-law behaviour immediately lends itself to a self-similar analysis. The application of such an analysis to systems which display a power-law behaviour may be worth a more detailed investigation (cf. Ori and Piran 1987).

6.4 Geometric Interpretation of Self-Similar Symmetry

In the Appendix it has been demonstrated that the general self-similar solutions of Henriksen et al. (1983) do not have a simple geometrical analogue, except in the special case of the singular solution (M=S). The main difficulty was shown to be the fact that the infinitesimal transformations corresponding to a conformal motion do not preserve the physical symmetry of self-similarity of the second kind. It was found that the generator used by Henriksen et al. introduced an acceleration, which transformed the reference frame to that of an accelerated observer, thereby destroying the conformal symmetry. One way to remove this difficulty would be to transform back to the reference frame of a comoving observer using a Lorentz transformation (or 'clock synchronisation' as it was called) after each infinitesimal point transformation. However, such a combination of transformations (i.e. the infinitesimal point transformation followed by a clock synchronisation) should be expressible in terms of the vanishing Lie derivative of some geometric object.

In Chapter 2 we discussed briefly some of the various types of geometric symmetries that could be imposed on a spacetime (see also Katzin et al. 1969). However, a more interesting possibility was contained in the recent publication of Ludwig (1987) in which he discussed conformal rescalings coupled with Lorentz transformations. The relevance of this work to the present problem was discussed in the Appendix. It remains for us to translate the findings of Ludwig into a form suitable to facilitate the determination of the geometric equivalent to the general self-similar symmetry of the second kind displayed in Chapter 3.

APPENDIX: Conformal Motions and Self-Similarity

In this Appendix we investigate the global symmetry properties of the solutions discussed in Chapter 3 of this thesis, wherein the physical symmetry was that of a similarity symmetry of the second kind. Henriksen et al. (1983), HEW, claim that the generator of the Lie group symmetry corresponding to their self-similar motion may be taken as

$$v^a = e^{\lambda t}(1/\lambda, r, 0, 0) \tag{A.1}$$

in the positive Λ case, where $\lambda = \sqrt{(\Lambda/3)}$, and that the Lie derivative of the metric is

$$\mathcal{L}_{\mathbf{v}} g_{ab} = 2e^{\lambda t} g_{ab} \quad , \tag{A.2}$$

whereby the symmetry is seen to be conformal (Yano 1955). However, the (0,1) term of the Lie derivative is

$$\mathcal{L}_{\mathbf{v}} g_{01} = \lambda e^{\lambda t} r g_{11} \neq 0 \quad , \tag{A.3}$$

so that the generator, (A.1), in fact does *not* give a conformal symmetry for this spacetime. Indeed, any symmetry of this spacetime which preserves the self-similar variable ξ cannot have a conformal symmetry as defined by (A.2), since the generator must be of the form (A.1) multiplied by some scalar function.

However, we can define a *partial* conformal symmetry (e.g., Tomita 1981), by considering the subspace dt=0, i.e. we can define a symmetry between t=constant hypersurfaces. If we have two spatial vectors on the $t=t_0$ hypersurface and transform them to $t=t_1$ then the angle between the two vectors is conserved, thus providing a partial conformal symmetry. Similar considerations apply on subspaces dr=0.

If we now consider the isotropic form of HEW's singular solution, where the metric is given by equation (3.68) with B=1 and γ=2, we find that we *do* have a conformal symmetry. We wish to show that for this isotropic singular solution the Lie derivative of the metric is of the form

$$\mathcal{L}_{\mathbf{v}} g_{ab} = 2f g_{ab} \quad . \tag{A.4}$$

The equations to be satisfied, for a conformal symmetry of any metric to exist, are

$$\sigma_t \alpha + \sigma_r \beta + 2\alpha_t = 2f \quad , \tag{A.5a}$$

$$\omega_t \alpha + \omega_r \beta + 2\beta_r = 2f \quad , \tag{A.5b}$$

$$R_t \alpha + R_r \beta = fR \quad , \tag{A.5c}$$

$$e^{\sigma}\alpha_r - e^{\omega}\beta_t = 0 \quad , \tag{A.5d}$$

where we have written the generator of the conformal symmetry as,

$$v^a = (\alpha, \beta, 0, 0) \quad , \tag{A.6}$$

with α,β and f unknown functions of the coordinates r and t to be obtained from equations (A.5). Since $\sigma=\sigma(\xi)$, $\omega=\omega(\xi)$, $R=rS(\xi)$ and $\xi=e^{\lambda t}/\lambda r$ (see Chapter 3) the first three equations of (A.5) can be further reduced to

$$\frac{e^{\lambda t}}{r}\sigma' \left[\alpha - \frac{\beta}{\lambda r}\right] + 2\alpha_t = 2f \quad , \tag{A.7a}$$

$$\frac{e^{\lambda t}}{r}\omega' \left[\alpha - \frac{\beta}{\lambda r}\right] + 2\beta_r = 2f \quad , \tag{A.7b}$$

$$e^{\lambda t} S' \left[\alpha - \frac{\beta}{\lambda r}\right] + \beta S = frS \quad , \tag{A.7c}$$

where (') denotes $d/d\xi$. For the isotropic singular solution we have that

$$S = \xi \quad , \quad e^{\sigma} = 1 \quad , \quad e^{\omega} = \xi^2$$
$$P = -1/S^2 = -\eta/3 \tag{A.8}$$

and therefore,

equation (A.7a) $\Rightarrow$ $\alpha_t = f$, (A.9a)

equation (A.5d) $\Rightarrow$ $\alpha_r = \dfrac{e^{2\lambda t}}{\lambda^2 r^2}\beta_t$, (A.9b)

equation (A.7c) $\Rightarrow$ $\lambda\alpha = f$, (A.9c)

then equation (A.7b) $\Rightarrow$ $\beta = rg(t)$. (A.9d)

From equations (A.9a) and (A.9c) we obtain

$$\alpha = e^{\lambda t}h(r) \quad , \tag{A.10}$$

and (A.9b) then gives

$$\lambda r\frac{dh}{dr} = \frac{e^{\lambda t}}{\lambda}\frac{dg}{dt} = K \quad , \tag{A.11}$$

where K is an arbitrary constant. Thus, we obtain the general solutions for h and g (by integrating (A.11)),

$$h(r) = \frac{K}{\lambda}\ln(Dr) \quad ,$$
$$g(t) = -K(e^{-\lambda t}+E) \quad , \tag{A.12}$$

where D and E are two constants of integration, and so

$$\alpha = \frac{K}{\lambda}\ln(Dr)e^{\lambda t} \quad ,$$
$$\beta = -Kr(e^{-\lambda t}+E) \quad , \tag{A.13}$$
$$f = K\ln(Dr)e^{\lambda t} \quad .$$

Thus, the singular solution of HEW admits a conformal Killing vector, which (with K=1, D=λ and E=0) is

$$v^a = [(e^{\lambda t}/\lambda)\ln(\lambda r), -re^{-\lambda t}, 0, 0] \qquad (A.14)$$

acting as the generator of the Lie group. The Lie derivative of this isotropic metric with the generator (A.14) is, then,

$$\mathcal{L}_{\mathbf{v}} g_{ab} = 2 \ln(\lambda r) e^{\lambda t} g_{ab} \quad . \qquad (A.15)$$

The lack of a conformal symmetry in the HEW solutions somewhat limits their physical applicability. However, the isotropic form of their singular solution, although restrictive, *is* conformal and may therefore have some connection with gauge theories describing the state of the early universe. The fact that in all of these solutions we have a self-similar symmetry of the physical variables with no corresponding conformal symmetry of the metric is puzzling. This seems to suggest that self-similarity of the second kind is more complicated than previously believed with no simple conformal analogue.

The generator (A.1) corresponds to the linear transformation

$$\bar{x}^{\mu} = x^{\mu} + v^{\mu} d\alpha \quad , \qquad (A.16)$$

where $\bar{x}^{\mu}$ and x^{μ} are points on a ξ=constant surface, (**v** is ξ-preserving), and α is the transformation parameter. This introduces an acceleration, (or rotation), as can be seen from equation (A.3). Thus our reference frame becomes that of an accelerated observer and we no longer have a conformal symmetry on these surfaces of constant ξ. However, if for *each* point, $\bar{x}^{\mu}$, we make a further transformation ("clock synchronisation")* to the metric of a comoving observer we may still be able to define a global symmetry corresponding to a similarity symmetry of the second kind which combines a series of infinitesimal

* R.N. Henriksen, private communication

linear transformations with a coordinate rotation at each step.

A recent publication by Ludwig (1987) discusses the decomposition of a general element of the group $GL(2,C)\otimes G\tilde{L}(2,C)$, the direct product of the two-dimensional complex linear group with itself, into a "standard" conformal rescaling and a "standard" Lorentzian transformation. The analysis is basically as follows:

Any 2x2 matrix,

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad ,$$

with $\theta^{-1} = ad-bc \neq 0$ may be written as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a\theta & b \\ c\theta & d \end{bmatrix}\begin{bmatrix} \theta^{-1} & 0 \\ 0 & 1 \end{bmatrix} \quad . \tag{A.17}$$

The matrix,

$$\begin{bmatrix} \theta^{-1} & 0 \\ 0 & 1 \end{bmatrix} \quad ,$$

has the same determinant as M, (the second matrix in the decomposition is unimodular).

A general element $(M,\tilde{M})$ of $GL(2,C)\otimes G\tilde{L}(2,C)$ may be decomposed as follows,

$$(M,\tilde{M}) = \left\{ I, \begin{bmatrix} \theta^{-1}\tilde{\theta}^{-1} & 0 \\ 0 & 1 \end{bmatrix} \right\}\left\{ \theta^{1/2}\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} \tilde{a}\tilde{\theta}\theta^{1/2} & \tilde{b}\tilde{\theta}\theta^{1/2} \\ \tilde{c}\theta^{-1/2} & \tilde{d}\theta^{-1/2} \end{bmatrix} \right\}$$
$$\times \left(\theta^{-1/2}I, \theta^{1/2}I \right) \quad , \tag{A.18}$$

where I is the unit matrix, i.e. $(M,\tilde{M})$ can be decomposed into a pure spin transformation, followed by a standard Lorentzian transformation, followed by what Ludwig (1987) called a basic right conformal rescaling. [In the 'real' case the tilde reduces to the complex conjugate, but in general, tilded and tilde-free quantities are independent complex variables].

$\langle M,\tilde{M}\rangle$ could equally well be decomposed into a product of a pure spin transformation, a standard Lorentzian transformation and a basic *left* conformal rescaling. Other variations are also possible. A more symmetric choice would be,

$$\langle M,\tilde{M}\rangle = \left\{ \begin{bmatrix} \theta^{-1} & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \tilde{\theta}^{-1} & 0 \\ 0 & 1 \end{bmatrix} \right\} \left\{ \begin{bmatrix} a\theta & b\theta \\ c & d \end{bmatrix}, \begin{bmatrix} \tilde{a}\tilde{\theta} & \tilde{b}\tilde{\theta} \\ \tilde{c} & \tilde{d} \end{bmatrix} \right\} . \qquad (A.19)$$

This is a standard Lorentzian transformation followed by a basic conformal rescaling. It is this last decomposition which may prove useful in determining some non-trivial geometric object, (Yano 1955), which would define the global symmetry of a spacetime corresponding to a similarity motion of the second kind.

REFERENCES

Aarseth, S. J., Gott, J. R. and Turner, M. S., 1979. *Ap. J.*, **228**, 664.

Abell, G. O., 1958. *Ap. J. Suppl.*, **3**, 213.

Abers, E. S. and Lee, B., 1973. *Phys. Rep.*, **9C**, 1.

Albrecht, A. and Steinhardt, P. J., 1982. *Phys. Rev. Lett.*, **48**, 1220.

Albrecht, A. and Turok, N., 1985. *Phys. Rev. Lett.*, **54**, 1868.

Alexander, D. and Green, R. M., 1988, in *The Post-Recombination Universe*, eds. N. Kaiser and A. N. Lasenby, (Kluwer Academic, Dordrecht).

Alfvén, H., 1971. *Phys. Today*, **24**, 28

Allen, J. E. and Andrews, J. G., 1970. *J. Plasma Physics*, **4**, 187.

Alpher, R. A., Follin, J. W. and Herman, R. C., 1953. *Phys. Rev.*, **92**, 1347.

Anderson, J. L., 1969, in M. Carmeli, S. Fickler and L. Witten (eds.), *Relativity - Proceedings of the Relativity Conference in the Midwest*, New York.

Banerjee, A., Duttachoudhury, S. B. and Sanyal, A. K., 1986. *Gen. Rel. Grav.*, **18**, 461.

Barenblatt, G. I. and Zel'dovich Ya. B., 1972. *Ann. Rev. Fluid Mech.*, **4**, 285.

Barrow, J. D., 1976. *Mon. Not. R. astr. Soc.*, **175**, 359.

——————, 1977. *Nature*, **267**, 117.

——————, 1987. *Phys. Lett.*, **187B**, 12.

——————, 1988. *Nucl. Phys.*, **B296**, 697.

Barrow, J. D. and Carr, B. J., 1977. *Mon. Not. R. astr. Soc.*, **182**, 537.

Barrow, J. D., Galloway, G. J. and Tipler, F. J., 1986. *Mon. Not. R. astr. Soc.*, **223**, 835.

Barrow, J. D. and Stein-Schabes, J. A., 1984. *Phys. Lett.*, **103A**, 315.

Batakis, N. and Cohen, J. M., 1975. *Phys. Rev. D*, **12**, 1544.

Bernstein, J., 1987. *The bulk-viscosity-driven inflationary model: re-visited*, Rockefeller Univ. Report, RU 87/B_1/188.

Bertschinger, E., 1983. *Ap. J.*, **268**, 17.

Bianchi, L., 1897. *Mem. Soc. It. Della. Sc.*, **11**, 267.

Bicknell, G. V. and Henriksen, R. N., 1978a. *Ap. J.*, **219**, 1043.

——————————————, 1978b. *Ap. J.*, **225**, 237.

Bluman, G. W. and Cole, J. D., 1974. *Similarity Methods for Differential Equations*, (Springer, N.Y.).

Bond, J. and Szalay, A., 1983. *Ap. J.*, **276**, 443.

Bondi, C. M. and Bondi, H., 1949. *Mon. Not. R. astr. Soc.*, **109**, 62.

Bonnor, W. B., 1972. *Mon. Not. R. astr. Soc.*, **159**, 261.

Boynton, P. E. and Partridge, R. B., 1973. *Ap. J.*, **181**, 243.

Brandenberger, R. H., 1987. *Inter. J. Mod. Phys. A*, **2**, 77.

Brans, C. and Dicke, R. H., 1961. *Phys. Rev.*, **124**, 925.

Brout, R., Englert, F. and Gunzig, E., 1978. *Ann. Phys.*, **115**, 78.

Brown, J. C. and Emslie, A. G., 1988, preprint.

Cahill, M. E. and Taub, A. H., 1971. *Commun. math. Phys.*, **21**, 1.

Cameron Reed, B. and Henriksen, R. N., 1980. *Ap. J.*, **236**, 338.

Cantwell, B. J., 1981. *Ann. Rev. Fluid Mech.*, **13**, 457.

Canuto, V., 1974. *Ann. Rev. Astr. Astrophys.*, **12**, 167.

——————, 1975. *Ann. Rev. Astr. Astrophys.*, **13**, 335.

Carr, B. J., 1975. *Ap. J.*, **201**, 1.

——————, 1976. *Ap. J.*, **206**, 8.

——————, 1977. *Mon. Not. R. atr. Soc.*, **181**, 293.

Carr, B. J. and Barrow, J. D., 1979. *Gen. Rel. Grav.*, **11**, 383.

Carr, B. J. and Hawking, S. W., 1974. *Mon. Not. R. astr. Soc.*, **168**, 399.

Carr, B. J. and Rees, M. J., 1977. *Astr. Astrophys.*, **61**, 705.

Chapline, G. F., 1975. *Nature*, **253**, 251.

Cheng, T. P. and Li, L. F., 1984. *Gauge Theory of Elementary Particle Physics*, (Oxford University Press, Oxford).

Coleman, S. and Weinberg, E., 1973. *Phys. Rev.*, **D7**, 1888.

Coley, A. A. and Tupper, B. O. J., 1983. *Ap. J.*, **271**, 1.

——————————————, 1984. *Ap. J.*, **280**, 26.

——————————————, 1985. *Ap. J.*, **288**, 418.

Collins, J. and Perry, M. J., 1975. *Phys. Rev. Lett.*, **34**, 1353.

Collinson, C. D. and French, D. C., 1967. *J. Math. Phys.*, **8**, 701.

Dekel, A. and Silk, J., 1986. *Ap. J.*, **303**, 39.

Demiański, M. and Grischuk, L. P., 1972. *Commun. math. Phys.*, **25**, 233.

Denavit, J., 1979. *Phys. Fluids*, **22**, 1384.

de Lapparent, V., Geller, M. and Huchra, J., 1986. *Ap. J. Lett.*, **302**, L1.

de Sitter, W., 1917. *Proc. Kon. Ned. Akad. Wet.*, **19**, 1217.

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

Dicke, R. H., Peebles, P. J. E., Roll, P. G. and Wilkinson, D. T., 1965. *Ap. J.*, **142**, 414.

Diósi, L., Keszthelyi, B., Lukács, B. and Paál, G., 1984. *Acta Phys. Pol.*, **B15**, 909.

Dirac, P. A. M., 1938. *Proc. Roy. Soc.*, **A165**, 199.

Duff, M. J., Nilsson, B. E. W. and Pope, C. N., 1986. *Phys. Rep.*, **130**, 1.

Dyer, C. C., 1979. *Mon. Not. R. astr. Soc.*, **189**, 189.

Eardley, D. M., 1974. *Commun. math. Phys.*, **37**, 287.

Efstathiou, G., 1983, in *Early Evolution of the Universe and Its Present Structure*, ed. G.O. Abell & G. Chincarini, (Reidel, Dordrecht).

Einstein, A., 1905. *Ann. Phys. (Germany)*, **17**, 891.

——————, 1915. *Preuss. Akad. Wiss. Berlin Sitzber*, 799.

Ellis, G. F. R. and MacCallum, M. A. H., 1969. *Comm. math. Phys.*, **12**, 108.

Fabbri, R., Guidi, I., Melchiorri, F. and Natale, V., 1980. *Phys. Rev. Lett.*, **44**, 23.

Fillmore, J. A. and Goldreich, P., 1984, *Ap. J.*, **281**, 1.

Friedmann, A., 1922. *Z. Phys.*, **10**, 377.

Gaffet, B. and Fukue, J., 1983. *Publ. Astron. Soc. Japan*, **35**, 365.

Gamow, G., 1948. *Nature*, **162**, 680.

Gamow, G., 1953. *Dan. Mat. Fys. Medd.*, **27**, N^{O}10.

Gibbons, G. W., Hawking, S. W. and Siklos, S. T. C., 1983. *The Very Early Universe*, (Cambridge University Press, Cambridge)

Giovanelli, R. and Haynes, M. P., 1982. *Astron. J.*, **87**, 1355.

Goicoechea, L. J. and Sanz, J. L., 1984. *Ap. J.*, **286**, 392.

Gorenstein, M. V. and Smoot, G. F., 1981. *Ap. J.*, **244**, 361.

Gőtz, G., 1988. *Phys. Lett.*, **128A**, 129.

Guth, A. H., 1981. *Phys. Rev. D*, **23**, 347.

Hacyan, S., 1979. *Ap. J.*, **229**, 42.

Hagedorn, R., 1970. *Astr. Astrophys.*, **5**, 184.

Harrison, E. R., 1965. *Ap. J.*, **142**, 1643.

Harrison, E. R., 1973. *Ann. Rev. Astr. Astrophys.*, **11**, 155.

Hawking, S. W., 1971. *Mon. Not. R. astr. Soc.*, **152**, 75.

———————, 1974. *Nature*, **248**, 30.

———————, 1975. *Commun. math. Phys.*, **43**, 199.

Hawking, S. W. and Ellis, G. F. R., 1973. *The Large Scale Structure of Space-Time*, (Cambridge University Press, Cambridge).

Hawking, S. W. and Moss, I. G., 1982. *Phys. Lett.*, **110B**, 35.

Heckmann, O. and Schűcking, E., 1958, in *Gravitation*, ed. L. Witten, (John Wiley).

Helgason, S., 1962. *Differential Geometry and Symmetric Spaces*, (Academic Press, N.Y.).

Heller, M., Klimek, Z. and Suszycki, L., 1973. *Astr. Sp. Sc.*, **20**, 205.

Henriksen, R. N., 1982. *Phys. Lett.*, **119B**, 85.

———————, 1987. *Ap. J.*, **314**, 33.

Henriksen, R. N., Emslie, A. G. and Wesson, P. S., 1983. *Phys. Rev. D.*, **27**, 1219.

Henriksen, R. N. and Turner, B. E., 1984. *Ap. J.*, **287**, 200.

Henriksen, R. N. and Wesson, P. S., 1978a. *Astr. Sp. Sc.*, **53**, 429.

———————————————, 1978b. *Astr. Sp. Sc.*, **53**, 445.

Hubble, E. P., 1929. *Proc. Nat. Acad. Sci. U.S.*, **15**, 169.

Hughston, L. P. and Sommers, P., 1973. *Commun. math. Phys.*, **32**, 147.

Israel, W. and Vardalas, J. N., 1970. *Nuovo Cimento Letters*, **4**, 887.

Jeans, J. H., 1902. *Phil. Trans. Roy. Soc.*, **199A**, 49.

Jensen, L. G. and Stein-Schabes, J. A., 1987. *Phys. Rev. D*, **35**, 1146.

Johri, V. B., 1977. *Mon. Not. R. astr. Soc.*, **178**, 395.

Kaiser, N., 1985, in *Inner Space/Outer Space*, eds. E. W. Kolb and M. S. Turner, (University of Chicago Press, Chicago).

Kaluza, T., 1921. *Preuss. Akad. Wiss.*, 966.

Kantowski, R. and Sachs, R. K., 1966. *J. Math. Phys.*, **7**, 443.

Kasner, E., 1921. *Am. J. Math.*, **43**, 217.

Katzin, G. H., Levine, J. and Davis, W. R., 1969. *J. Math. Phys.*, **10**, 617.

Khabaza, I. M., 1965. *Numerical Analysis*, (Pergamon Press, London).

Kibble, T. W. B., 1976. *J. Phys.*, **A9**, 1387.

Killing, W., 1892. *J.F.D.R.U.A. Math.*, **109**, 121.

King, A. R. and Ellis, G. F. R., 1973. *Commun. math. Phys.*, **31**, 209.

Kinnersley, W., 1969. *J. Math. Phys.*, **10**, 1195.

Kirschner, R. F., Oemler, A., Schechter, P. L. and Shectman, S. A., 1983, in *Early Evolution of the Universe and Its Present Structure*, IAU Symp. **104**, 197.

Klein, O., 1926a. *Z. Phys.*, **37**, 895.

———, 1926b. *Nature*, **118**, 516.

Kormendy, J. and Knapp, G. R., 1987. *Dark Matter in the Universe*, IAU Symp. **117**, (Reidel, Dordrecht).

Landau, L. D. and Lifshitz, E. M., 1959. *Fluid Mechanics*, transl. J. B. Sykes and W. H. Reid (Pergamon Press, London).

Lemaitre, A. G., 1927. *Annales de la Société Scientifique de Bruxelles*, **XLVIIA**, 49. English translation, 1931. *Mon. Not. R. astr. Soc.*, **91**, 483.

Liang, E. P. T., 1974. *Phys. Lett.*, **51A**, 141.

Lifshitz, E. M. and Khalantnikov, I. M., 1963. *Soviet Phys. Usp.*, **6**, 495.

Lin, D. N. C., Carr, B. J. and Fall, S. M., 1976. *Mon. Not. R. astr. Soc.*, **168**, 399.

Linde, A. D., 1979. *Rep. Prog. Phys.*, **42**, 389.

————————, 1982. *Phys. Lett.*, **108B**, 389.

Liubimov, G. A., 1956. *Theoretical Hydromechanics*, **7**, (Oborongiz, Moscow).

Loh, E. D., 1987, in *Observational Cosmology*, IAU Symp. **124**, ed. A. Hewitt, G. Burbridge and L. Z. Fang, (Reidel, Dordrecht).

Ludwig, G., 1987. *Gen. Rel. Grav.*, **19**, 1.

Mach, E., 1912. *Die Mechanik in Ihrer Entwicklung Historisch-Kritisch Dargestellt*; English translation by T. J. McCormack, 1960. *The Science of Mechanics*, Open Court, La Salle, Ill.

Marsden, J. E. and Tipler, F. J., 1980. *Phys. Rep.*, **66**, 109.

Martinez-Gonzalez, E. and Jones, B. J. T., 1986. *Phys. Lett.*, **167B**, 37.

Matzner, R. A. and Misner, C. W., 1972. *Ap. J.*, **171**, 415.

Matzner, R. A., 1972. *Ap. J.*, **171**, 433.

McCrea, W. H., 1951. *Proc. R. Soc.*, **A206**, 562.

McIntosh, C. B. G., 1975. *Phys. Lett.*, **50A**, 429.

————————, 1976. *Gen. Rel. Grav.*, **7**, 199.

————————, 1978. *Phys. Lett.*, **69A**, 1.

McIntosh, C. B. G., 1980, in *Gravitational radiation, collapsed objects and exact solutions*, ed. C. Edwards, (Springer-Verlag, N.Y.).

Melott, A., 1983. *Ap. J.*, **264**, 59.

Misner, C. W., 1964, in DeWitt and DeWitt, *Differential Geometry and Differential Topology*.

Misner, C. W., 1967. *Nature*, **214**, 40.

————————, 1968. *Ap. J.*, **151**, 431.

Misner, C. W. and Sharp, D. H., 1964. *Phys. Rev.*, **136**, B571.

Misner, C. W., Thorne, K. S. and Wheeler, J. A., 1973. *Gravitation*, (Freeman, New York).

Munier, A., Burgan, J. R., Feix, M. and Fijalkow, E., 1980. *Ap. J.*, **236**, 970.

Narlikar, J. V., 1963. *Mon. Not. R. astr. Soc.*, **126**, 203.

Nightingale, J. D., 1973. *Ap. J.*, **185**, 105.

Novikov, I. D. and Thorne, K. S., 1973. in *Black Holes*, ed. C. De Witt and B. De Witt (New York, Gordon & Breach).

Omnes, R. L., 1969. *Phys. Rev. Lett.*, **23**, 38.

Ori, A. and Piran, T., 1987. preprint (submitted to *Phys. Rev. Lett.*).

Ostriker, J. P. and Cowie, L. L., 1981. *Ap. J. (Letters)*, **243**, L127.

Pacher, T., Stein-Schabes, J. A. and Turner, M. S., 1987. *Can bulk viscosity drive inflation?*, Fermilab-Pub.-87/69-A.

Page, D. N. and Hawking S. W., 1976. *Ap. J.*, **206**, 1.

Pakula, R. and Sigel, R., 1985. *Phys. Fluids*, **28**, 232.

Papadopoulos, D. and Esposito, F. P., 1985. *Ap. J.*, **292**, 330.

Papadopoulos, D. and Sanz, J. L., 1985. *Nuovo Cimento Letters*, **42**, 215.

Peccei, R. and Quinn, H., 1977. *Phys. Rev. Lett.*, **38**, 1440.

Peebles, P. J. E., 1965. *Ap. J.*, **142**, 1317.

——————, 1980. *The Large-Scale Structure of the Universe*, (Princeton University Press, Princeton).

Penzias, A. A. and Wilson, R. W., 1965. *Ap. J.*, **142**, 419.

Petrov, A. Z., 1969. *Einstein Spaces*, (Pergamon, Oxford).

Podurets, M. A., 1964. *Astron. Zh.*, **41**, 28 [*Sov. Astron.*, **8**, 19].

Ponce de Leon, J., 1987. *Phys. Lett.*, **126A**, 75.

Preskill, J. P., 1979. *Phys. Rev. Lett.*, **43**, 1365.

Raychaudhuri, A. K., 1955. *Phys. Rev.*, **98**, 1123.

Reynolds, O., 1883. *Phil. Trans.*, t.clxxxiv, 935.

Rubin, V. C., Ford, W. K. and Thonnard, N., 1980. *Ap. J.*, **238**, 471.

Ryan, M. P. and Shepley, L. C., 1975. *Homogeneous Relativistic Cosmologies*, (Princeton University Press, Princeton).

Saarinen, S., Dekel, A. and Carr, B. J., 1987. *Nature*, **325**, 598.

Sandage, A. R., 1972. *Ap. J.*, **178**, 1.

Sandage, A. R. and Hardy, E., 1973. *Ap. J.*, **183**, 743.

Sanz, J. L., 1983. *Astr. Astrophys.*, **120**, 109.

Saunders, P. T., 1969. *Mon. Not. R. astr. Soc.*, **142**, 213.

Shaver, E., 1986. MSc. Thesis, Queen's University, Ontario, Canada.

Schrődinger, E., 1926. *Ann. Physik*, **79**, 361, 489.

——————————————, **81**, 109.

Schutz, B. F., 1980. *Geometrical methods of mathematical physics*, (Cambridge University Press, Cambridge).

Schwarz, J. H., 1985. *Superstrings*, (World Scientific, Singapore).

Sciama, D. W., 1971, in *General Relativity and Cosmology*, ed. R. K. Sachs, (New York, Academic Press).

Sedov, L. I., 1959. *Similarity and Dimensional Methods in Mechanics*, (Academic Press, N.Y.).

Seldner, M., Siebers, B., Groth, E. J. and Peebles, P. J., 1977. *Ap. J.*, **82**, 249.

Shapley, H. and Ames, A., 1932. *Harvard Ann.*, **88**, N^{o}2.

Shellard, E. P. S., 1987. *Nuc. Phys. B*, **283**, 624.

Smalley, L. L., 1974. *Phys. Rev.*, **D9**, 1635.

Stanyukovich, K. P., 1955. *Unsteady Motion of Continuous Media*, (Gostekhizdat, Moscow), English translation by M. Holt, 1960,

(Academic Press, N.Y.).

Starobinskii, A. A., 1983. *Sov. Phys. JETP Lett.*, **37**, 66.

Steigman, G. and Turner, M. S., 1983. *Phys. Lett.*, **128B**, 295.

Symon, K. R., 1971. *Mechanics*, (Addison-Wesley, London).

Synge, J. L., 1961. *Relativity: The General Theory*, (Amsterdam, North Holland).

Szekeres, P., 1975. *Comm. math. Phys.*, **41**, 55.

Taub, A. H., 1951. *Ann. Math.*, **53**, 472.

Tomita, K., 1981. *Suppl. Prog. Theor. Phys.*, **70**, 286.

Turner, M. S., 1987, in *Dark Matter in the Universe*, IAU Symp. **117**, ed. J. Kormendy and G. R. Knapp, (Reidel, Dordrecht).

Turok, N., 1987. *Two Lectures on the Cosmic String Theory of Galaxy Formation*, Imperial/TP/86-87/23.

Turok, N. and Brandenberger, R. H., 1985. *Cosmic Strings and the Formation of Galaxies and Clusters of Galaxies*, UCSB/TH-7/1985.

Tupper, B. O. J., 1981. *J. Math. Phys.*, **22**, 2666.

Vaidya, P. C., 1951. *Proc. Indian Acad. Sci.*, **A33**, 264.

——————, 1953. *Nature*, **171**, 260.

Van den Bergh, S., 1969. *Nature*, **224**, 891.

Varey, R. H. and Sander, K. F., 1969. *Brit. J. Appl. Phys.*, **2**, 541.

Vilenkin, A., 1981. *Phys. Rev. Lett.*, **46**, 1169.

Vilenkin, A., 1985. *Phys. Rep.*, **121**, 263.

Waga, I., Falcão, R. C. and Chanda, R., 1986. *Phys. Rev. D*, **33**, 1839.

Wainwright, J., 1984. *Gen. Rel. Grav.*, **16**, 657.

Wainwright, J., Ince, W. C. W. and Marshman, B. J., 1979. *Gen. Rel. Grav.*, **10**, 259.

Wald, R., 1983. *Phys. Rev. D*, **28**, 2118.

Walecka, J. D., 1974. *Ann. Phys.*, **83**, 491.

Weinberg, S., 1971. *Ap. J.*, **168**, 175.

Weizsäcker, C. F. von, 1951. *Ap. J.*, **114**, 165.

Wesson, P. S., 1975. *Astr. Sp. Sc.*, **32**, 273, 305, 315.

——————, 1979. *Ap. J.*, **228**, 647.

——————, 1984. *Gen. Rel. Grav.*, **16**, 193.

——————, 1986a. *Astron. Astrophys.*, **166**, 1.

——————, 1986b. *Phys. Rev. D*, **34**, 3925.

White, S. D. M., Davis, M. and Frenk, C. S., 1984. *Mon. Not. R. astr. Soc.*, **209**, 15.

White, S. D. M., Davis, M., Efstathiou, G. and Frenk, C. S., 1987. *Ap. J.*, **313**, 505.

Witten, E., 1981. *Nucl. Phys. B*, **186**, 412.

Yano, K., 1955. *The Theory of Lie Derivatives and their Applications*, (North-Holland, Amsterdam).

Yano, K. and Bochner, S., 1953. *Curvature and Betti Numbers*, (Princeton University Press, Princeton).

Zel'dovich, Ya. B., 1962. *Soviet Phys. J.E.T.P.*, **14**, 1143.

——————, 1970. *Astr. Astrophys.*, **5**, 84.

——————, 1980. *Mon. Not. R. astr. Soc.*, **192**, 663.

Zel'dovich, Ya. B. and Khlopov, M. Yu., 1979. *Phys. Lett.*, **79B**, 239.

Zel'dovich, Ya. B. and Novikov, I. D., 1967. *Sov. Astr. AJ.*, **10**, 602.

Zel'dovich, Ya. B. and Novikov, I. D., 1971. *Relativistic Astrophysics Vol. 1*, (University of Chicago Press, Chicago).

Zel'dovich, Ya. B. and Raizer, Yu. P., 1967. *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena*, Volumes 1 & 2, (Academic Press, N.Y.).