LINEAR LIBRARY
C01 0068 1077

OBSERVATIONAL SPACE-TIMES

S.D. NEL

A dissertation submitted to the University of Cape Town in partial fulfilment of the requirements for the degree Ph.D. in Applied Mathematics

University of Cape Town

1980

The University of Cape Town has been given the right to reproduce this thesis in whole or in part. Copyright is held by the author.

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

OBSERVATIONAL SPACE-TIMES

S.D. NEL

Ph.D. Dissertation
University of Cape Town, 1980

Supervisor: G.F.R. Ellis

Alternatives to the standard Friedman-Robertson-Walker models of the universe are investigated. After discussing the philosophical and observational status of the space-time symmetry assumptions on which these models are based, the consequences of weakening these assumptions are examined. We present a systematic method of obtaining explicit solutions of Einstein's field equations for space-times that are spatially homogeneous (but anisotropic), and derive a number of (apparently) new solutions. As an example of a model that is inhomogeneous, we study the properties of a static, spherically symmetric space-time. The remainder of the thesis is devoted to an entirely different approach to cosmology. Instead of making *a priori* assumptions about the space-time geometry, we attempt to deduce the large-scale structure of the universe by using astronomical observations made at one point in space-time as initial data for the field equations. Under the assumption that the universe may be modelled by a pressure-free perfect fluid, it is proved that observations of distant galaxies provide sufficient information to determine completely the local structure of the cosmology on a part of the past light cone of the observer. This integration is performed explicitly for polynomial initial data. In addition, it is indicated how the solution may be propagated towards the interior of the light cone. Since the required initial data is not currently available, we suggest that a series of assumptions be made about the observations. Each set of assumptions gives rise to an *observational space-time*, i.e. to an equivalence class of models in which the general relativistically predicted observational relations are precisely those assumed.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Prof. G.F.R. Ellis, for suggesting this research problem and for his guidance and encouragement. I am indebted to Prof. R. Penrose, Dr. S.T.C. Siklos and, in particular, R. Maartens and Dr. K.P. Tod for many helpful discussions. I would also like to thank the A.E.C.I. Corporation (South Africa) and the Rhodes Scholarship Trust for financial support during the course of this work.

S.D. Nel

1980

CONTENTS

1. STATEMENT OF THE PROBLEM

§1.1. Introduction

The work presented in this thesis forms part of a programme undertaken in collaboration with G.F.R. Ellis and R. Maartens, the primary aim of which is to examine in detail how cosmological observations may be used to determine the large scale structure of space-time (see, e.g.,refs. [1-5]). In order to place this work in context as part of the ongoing cosmological enterprise, it is necessary to review briefly the main objectives of cosmology, the basic assumptions underlying most modern attempts to achieve these objectives, and the observational as well as philosophical status of these assumptions.

§1.2. The aim of cosmology

The main objectives of cosmology are, broadly speaking, three-fold:

(i) It seeks to determine the large scale structure of the universe "from a fusion of the results of astronomical observations with knowledge derived from local physical experiments " ([6]). Within the class of astronomical observations that may be regarded as characterising the universe ([7]), one may distinguish six observations that are of particular importance: (a) the systematic redshifts of galaxies, (b) the magnitude-redshift relation, (c) the distortion-redshift relation, (d) the number counts-redshift relation, (e) the apparent proper motions of galaxies and (f) the observed background radiation; in particular, the near-isotropy of its dominant component, the micro-wave radiation. (These observations will be discussed in detail in

Chapter 4.) A large part of this thesis will be concerned with investigating to what extent these observations determine the structure of the universe (see Chapter 5). For the present purposes, however, we merely note that any viable cosmological theory must, at the very least, be consistent with these observations.

(ii) It aims to account for cosmologically significant physical processes that are observed or thought to have taken place at various stages during the evolution of the universe. Specifically, it needs to explain: (a) the origin of the observed background radiation, as well as the near-black body spectrum and near-isotropy of the micro-wave component (see, e.g., [8],[9]), (b) the origin of high-energy extragalactic cosmic rays (provided their origin as extragalactic is finally established; see, e.g., [10]), (c) observed light element abundances (by providing detailed models of primordial element formation; see, e.g., [11],[12]), (d) galaxy formation (see, e.g., [13]).

Such explanations should be based on physical theories that are again, at least, compatible with the locally determined laws of physics.

(iii) One may summarise the two aims of cosmology listed thus far by saying that they attempt to answer the questions, "What is there?" and "How did it come to be?" An essential part of any complete answer to these two questions would be a discussion of the nature of the initial or boundary conditions of the universe, both from a "Machian" point of view, and as boundary conditions for whatever differential equations are assumed to govern large scale dynamical evolution ([14]). The final aim of cosmology is then to explain the answers to the previous two questions. Thus it asks, "Why is there what there is, and why did it come to be in just this way?"

In order to be more specific, suppose for the moment that overwhelming cosmological evidence (both from direct observation and indirect arguments) eventually forces us to conclude that the initial stages of the universe were close to being both spatially homogeneous and isotropic, i.e. the Friedmann-Robertson-Walker (FRW) models (see, e.g., [15],[16]) are taken to be reasonably accurate, even at early times. It is clear that the initial conditions in such a universe would have to be very special ([17]): exact homogeneity would rule out galaxy formation, while large perturbations would collapse to form black holes before the stars could evolve. Thus strictly constrained density perturbations need to occur in the same way in all places in this universe, even though the existence of horizons rules out causal communication between most of the places involved.

The problem is then: is there some physico-philosophical principle which will single out these exceptional initial conditions from the "vast manifold" ([18]) of initial data compatible with physical laws and observations? And, of course, the same question may be asked whatever the initial conditions were in our actual universe. It is this riddle that Carter's "Anthropic Principle" ([19]), Penrose's "Entropy Hypothesis" ([18]) and Misner's "Chaotic Cosmology" programme ([20]) are ultimately designed to solve, albeit in completely different ways. (One may interpret Misner's idea as an attempt to explain away the problem.) The same question arises - in a "space-like" sense - in models in which there is no initial singularity, such as the spherically symmetric static universe described in [21] (included as Appendix III).

§1.3. Basic assumptions

In order to pursue the objectives outlined above, it is necessary to make a number of assumptions. Some of these are so common to almost all modern cosmological theories that they have acquired the status of *fundamental principles:*

(i) "Whenever normal physical laws can be applied, they correctly predict the structure of the universe." Thus formulated by Ellis ([6]), this "Local Predictability" assumption embodies both the belief that locally determined laws are valid at all points in space-time, as well as the principle (formalised mathematically by requiring that space-time be inextendible) that these laws should be applied for as long as is possible. Although it may be argued ([22], [23]) that there is some observational evidence to support this hypothesis, its status is essentially that of an unverifiable philosophical assumption which is invoked mainly to exclude "exotic" physical processes or ad hoc models of the kind discussed in, e.g., [6].

Throughout this thesis, standard astrophysical theory and elementary particle and nuclear physics will be accepted as accurate (but not necessarily perfect) descriptions of the phenomena to which they are applicable. Similarly, we shall usually assume Einstein's General Theory of Relativity to be an adequate description for dynamical purposes; however, in view of the large number of competing theories of gravity (see, e.g., [24]), it is sometimes useful to take a cosmographic approach, in which no particular set of field equations is assumed (in this connection, see [2],[16]).

(ii) On a suitably large scale, the universe may be modelled sufficiently accurately by the usual space-time manifold picture, with tensor fields or distributions representing the matter and radiation content.

This rather vague statement covers a whole series of interrelated assumptions concerning the main constituents of an adequate (even this concept needs clarification) cosmological model, and the domain of validity of such a model.

(a) The first ingredient of all models considered in the literature is a space-time (M,g), where M is a connected, 4-dimensional (and throughout this thesis) Hausdorff manifold, and g is a Lorentz metric on M (see, e.g., [25]). The metric determines the results of space-time measurements and, through its derivatives, endows M with an affine connection. With the exception of Chapter 2, we shall work throughout with a metric of signature (+---).

The use of a manifold picture involves a "smoothing" approximation, independently (at least until the field equations are assumed) of whether one adopts, in addition, the fluid approximation discussed below. Thus it involves the assumption that even if a fully quantised theory of gravity eventually necessitates a radically different representation of space-time - such as that envisaged by, e.g., Twistor Theory (see, e.g., [26]) - this breakdown of the classical manifold picture would occur at cosmologically insignificant scales of the order of 10^{-13}cm. What *is* important, particularly as it affects the problem of initial conditions, is the possible breakdown of the classical description at a cosmological singularity; this seems to be one region where quantum gravity would in any case be required to provide the final answers.

Singularities such as black or white holes present a slightly different problem. Clearly these can be incorporated in the classical framework simply by excluding them from the set of points taken to constitute the space-time manifold, but this has the disadvantage of leaving unresolved the causal difficulties associated with the presence of these holes in space-time. It is therefore often convenient to assume (for cosmological purposes) that such holes do not exist, i.e. to adopt a "hole-free" assumption of the kind used in, e.g.,[27]. (Of course, when individual observations are actually interpreted, the possible effects of black holes need to be considered quite carefully - see §4.1. for a further discussion of this point.

(b) Secondly, a mathematical description of the matter and radiation content of the universe is required. It is an obvious, but apparently sometimes overlooked fact that the particular description adopted depends crucially on the ordering of matter in the universe. We shall assume that the ordering of matter into stars, star clusters, galaxies, clusters of galaxies and intergalactic matter as determined by local astronomical observations holds also at large distances, in such a way as to determine a well-defined average 4-velocity u^a, proper mass density ρ and a conserved stress-energy tensor T_{ab}; i.e. we shall adopt the conventional "fluid approximation" (see, e.g., [15]). In this context, "suitably large" (cf. p.5) means a scale on which galaxies may be regarded as particles of a cosmological fluid, with the fluid 4-velocity defined by averaging over clusters of galaxies. The existence of such an averaging scale, leading to a well-defined 4-velocity, is a fundamental aspect of the fluid approximation.

If de Vaucouleurs' further hierarchical ordering of clusters of galaxies into superclusters, and these, in turn, into super-superclusters is correct, it is possible that no suitable averaging scale exists (see, e.g., [28],[29]). Any attempt to construct hierarchical models using the fluid approximation could therefore be fundamentally misconceived. A kinetic theory type approach (see, e.g., [30]) might be more appropriate. In fact, whether any hierarchical model could be constructed within the framework of General Relativity would depend to a large extent on whether the apparently severe problems concerning the actual definition of a stress-energy tensor could be overcome.

Common to both these interpretations of the observational data (and to all but the most eccentric interpretations) is the implicit assumption that the universe is large - so large that on a cosmological scale we are unable to move off our local galactic world line. The spatial distances and time scales involved are so great that we effectively view the universe from a single space-time point - "here and now". This restriction (which will be discussed in more detail in §4.1) imposes fundamental limitations on what is empirically decidable in cosmology.

The fundamental principles sketched above facilitate the formulation of the cosmological objectives as reasonably well-defined mathematical problems. In order to render these problems tractable, it is customary to make a series of *simplifying idealisations:*

(i) For dynamical purposes, it is convenient to treat the cosmological fluid as a "perfect" fluid ([31]). Thus we shall take as our stress-energy tensor

$$T_{ab} = -(\mu+p)u_a u_b + pg_{ab}, \quad u^a u_a = 1 \tag{1.3.1}$$

where μ is the total energy density, and p is the isotropic pressure. In addition, we shall require equations of state relating the variables ρ (see p.6), μ and p. These will be introduced later.

The perfect fluid idealisation neglects entirely the effects of, e.g., heat conduction and bulk viscosity, both of which could be important at various stages during the evolution of the universe ([32],[33]). Nevertheless, it may be argued ([31]) that the kind of matter-energy content that can be described by (1.3.1) is sufficiently diverse to cover most cosmological epochs.

(ii) Probably the most important idealisation made in most cosmological theories is the assumption that, again on a sufficiently large scale, the universe possesses a high degree of symmetry. The FRW models, e.g., are based on space-times that are invariant under a 6-dimensional isometry group (see [34] for a good discussion of isometry groups in cosmology). These models have been the subject of much fruitful research: Einstein's field equations can be solved in closed form for various types of matter-energy content, and definite predictions about the exact form of the observational relations mentioned in §1.2 are possible (see Chapter 6). Typically, these predicted relations contain three free parameters: the Hubble constant, the deceleration parameter and the total energy density. The aim of "observational cosmology" is then to determine the FRW model whose parameters best fit the observed universe.

In spite of their considerable success in providing detailed and plausible explanations of several important phenomena, these models should be treated with a certain amount of caution, because the symmetry assumptions on which they are based may be challenged on a number of grounds.

Closely related to the problem of the averaging scale involved in making the fluid approximation, there is the question concerning the scale on which these symmetries are supposed to hold. Ironically, this problem has been highlighted by the recently discovered anisotropy in the micro-wave background radiation (formerly, the isotropy of this radiation was one of the observational cornerstones of the FRW models), and its interpretation as implying large scale motions of matter in our vicinity ([35],[36]). There is also some evidence of anisotropy in the observed magnitude-redshift relation ([37]). This raises the question: what degree of observed departure from isotropy (e.g. how large a motion of our local group of galaxies relative to the background radiation) will be accepted as being consistent with the exact isotropy demanded by the FRW models? As has been pointed out in [1], what is required is some operational criterion: "We shall reject the FRW hypothesis if the anisotropy in test X at redshift z is above p% in a sample of size n". Only on the basis of such an agreed statistical verification criterion is it possible to undertake a meaningful evaluation of the observational evidence relating to isotropy and spatial homogeneity. (This evidence is discussed in detail in [16]; see also [34] for a brief, recent review.)

But even the concept of "observational evidence" is in some need of clarification. One of the problems here is that the FRW models have become a firmly established part of the observational cosmologist's theoretical framework. This has had the unfortunate effect that the observational data is often adjusted before publication by means of numerous correction terms (see, e.g., [16]), and these are calculated on the basis of FRW-type assumptions. The first step in assessing the evidence is to establish what is actually observed; this requires a careful analysis of observational techniques such as that undertaken in [38], [39]. Given the results of such an analysis, the observational status of the isotropy assumption may be assessed.

The near-isotropy of the micro-wave radiation has often been used as evidence for the isotropy of the universe about our space-time point of observation, but there are certain difficulties associated with this interpretation, quite apart from any problems caused by the recently reported anisotropy. Without additional evidence, current observations do not, for example, exclude the possibility that the redshift of emission of this radiation varies with direction (in fact, even within the FRW class of models, there is a wide range of possible emission redshifts). Furthermore, even if this redshift turns out to be the same in all directions, one cannot deduce from this that the universe is isotropic about the point of observation: a counter-example is provided by the spherically symmetric, static model described in [21].

What we are raising here is the problem of *uniqueness:* isotropy of space-time implies that all observations should be isotropic. Not only the background radiation, but also the observed magnitude-redshift and number counts-redshift relations should be isotropic. No distortion effect or apparent proper motions of distant galaxies must be observed, since these would single out preferred directions in the sky. The problem of uniqueness may then be formulated (in this context) as the converse question: do isotropic observations imply that the space-time is isotropic?

If we accept that the aim of cosmology is to determine the structure of the actual universe, it is not sufficient merely to exhibit a model that is consistent with observations; it is also necessary to show that this is the only such model. Strictly speaking, of course, we should be concerned with classes of model, since all that can actually be obtained from, e.g., observations of the micro-wave radiation is an upper limit on the anisotropy, rather than a statement of zero isotropy ([1]). More generally, there is a fundamental uncertainty inherent in all cosmological observations ([1],[6],[40]). Thus any attempt to verify that the universe is isotropic must ultimately concern itself with the problem of *stability:* given that isotropic observations do, in fact, imply space-time isotropy, can one show further that small departures from observational isotropy imply small deviations from space-time isotropy?

This problem is made particularly acute by the fact that observational uncertainty increases rapidly with redshift, and by the existence, in practice, of limiting redshifts, i.e. redshifts beyond which certain observations cannot be made

([1],[39]). If one accepts the interpretation of the background radiation as relic radiation from the plasma of a primeval fireball (see, e.g., [9]), there is also an absolute limiting redshift, corresponding to the time when the radiation and matter decouple, and the universe becomes opaque. As Ellis has pointed out ([1]), a very small departure from isotropy at that time could be consistent with very large departures from isotropy at earlier times - in tilted Bianchi Type V universes, such departures could be arbitrarily small at the time of decoupling, even though the early behaviour of these models is completely different from that of the FRW universes: there could, e.g., be two separate initial singularities in such models, at both of which the density remains finite ([27], [41-43]).

The existence of limiting redshifts, together with the basic limitation that we effectively view the universe from one space-time point (on a cosmological scale), means that our knowledge of cosmological null observations is restricted to a bounded subset of one past light cone. This severely restricts the region of the universe (even within our past light cone) in which our observational inferences can be assumed to be valid (unless the universe has finite space sections, and we have already seen around it; see [6],[44]).

One way of attempting to overcome this problem of the *domain of observational inference*, is to bound the isotropy of the early universe by using indirect arguments based on, e.g., light element abundance measurements ([45],[46]). Two points need to be borne in mind when evaluating the strength of these arguments ([1]). Firstly, a number of authors have recently used the theory of element formation in FRW models, together with current abundance estimates, to impose limits on the time variation of the gravitational constant ([23]) and Planck's constant ([47]), and to bound the lifetimes of neutrinos associated with heavy leptons ([48]). The latter are also sensitive to variations in the weak and strong coupling constants, as well as the neutron-proton mass difference. If there is sufficient uncertainty in these quantities to warrant an investigation of this kind, one is presumably justified in treating with some caution the results of the element abundance calculations bounding anisotropy - these are usually performed using specific values of these fundamental constants.

Secondly, these arguments bound anisotropy only in the vicinity of our galactic world line; extrapolation to larger distances is based on the assumption of spatial homogeneity. Unless this assumption can itself be verified, independently of isotropy, such extrapolation is observationally unjustified.

But the observational status of spatial homogeneity is far more dubious than that of isotropy. If we grant isotropy about one point, spatial homogeneity immediately ensures that the universe is isotropic about each point, and thus leads unambiguously to the FRW space-times. A kind of observational converse of this result can also be proved (see Chapter 6): a space-time which (a) is spherically symmetric about a regular, geodesic world line ("centro-symmetric" in the terminology of [34]), (b) contains a pressure-free, perfect fluid and (c) satisfies Einstein's field equations, will be FRW (and hence spatially homogeneous) within some region of our past light cone iff the observer area distance (see §4.2) and number counts vary with redshift in exactly the manner predicted by the standard FRW models. Thus one could, in principle, verify spatial homogeneity within our past light cone by carrying out detailed measurements of area distances and number counts as functions of redshift. An interesting point here is that the predicted number counts appear to be contradicted by observations ([16],[49]). This discrepancy is usually explained in terms of galaxy evolution; in this case, it is obviously not possible to use number counts to prove homogeneity as well. The problems associated with checking the area distance-redshift relation are numerous, since the actual determination of area distances depends crucially on a detailed knowledge of source and detector characteristics ([39]).

Outside our past light cone, observational verification of spatial homogeneity would be very difficult indeed. Even in the standard models there exist particle horizons, and thus regions of the universe from which we have, as yet, received no information.

Thus, while there is certainly some evidence for the isotropy of the universe about our point of observation, the spatial homogeneity hypothesis has, to an uncomfortably large extent, the status of an unverified (and, in practice, probably unverifiable), a priori assumption. In fact, it is usually introduced at the outset by invoking the so-called "Copernican Principle", which may be stated in various ways, e.g.: "We do not occupy a privileged position in space-time" ([15]). The philosophical basis of this principle has recently been criticised very forcefully in [17]. Only one of the arguments presented there needs to be mentioned, very briefly. The Copernican principle has its roots in "the laudable desire to express the idea that the creation of the universe was not centered on our presence, [but it] must not be confused with the incorrect statement that we could equally well live in all places and at all times in the universe". Even within the solar system, our particular form of life could only have evolved within a very narrow range of distances from the sun, i.e. in a very special position. Equally then, the universe might well be such that life (as we know it) could only have evolved in "privileged positions".

§1.4. Alternatives: outline of thesis

Partly because of the problems mentioned above, a number of alternatives to the FRW models have been investigated in recent years. The literature now abounds with examples of anisotropic and inhomogeneous cosmological models (for an extensive review, see [34]).

Spatially homogeneous (but anisotropic) space-times have been studied in great detail, but the observational predictions of these models are difficult to derive. Except for a few special cases ([50],[51]), approximation techniques need to be used ([52],[53]). One of the major obstacles remains the problem of integrating the geodesic equations in these models. It is at this point that exact solutions of the field equations become important; knowledge of closed form solutions is often useful in deriving observational predictions. In Chapter 2 we shall present a method that may be used to obtain, in a unified manner, many of the known exact solutions for spatially homogeneous models. By way of illustration, we derive some new solutions.

As an example of the possible consequences of abandoning the Copernican Principle, we present, as Appendix III, a joint paper with G.F.R. Ellis and R. Maartens in which inhomogeneous, spherically symmetric models are discussed.

Both of the alternatives mentioned thus far suffer from the drawback that they are still based on a priori symmetry assumptions, and thus all space-times that do not satisfy these requirements are ruled out by hypothesis rather than by observation.

In Chapters 3 - 6 we shall consider a completely different approach to cosmology, many of the ideas of which are implicit in the remarkable paper by Kristian and Sachs ([54]). In this *observational approach*, the observations are taken as the point of departure and we attempt to discover what these imply about the large scale structure of the universe - without making prior assumptions about its geometry.

One may do this cosmographically, i.e. without using a set of field equations ([2],[5]), but we shall be concerned mainly with the cosmological characteristic initial value problem: cosmological null observations are regarded as initial data on our past light cone for the Einstein field equations.

In Chapter 3, we introduce a system of null coordinates that is reasonably well adapted to this problem. The assumption that the galactic world line of the observer is a regular geodesic imposes certain restrictions on the behaviour of the metric, Riemann and stress-energy tensors in the vicinity of this world line. These "central conditions", which are of extreme importance to the rest of the analysis, are also derived in Chapter 3, using the so-called NP-formalism ([55]).

Chapter 4 investigates what cosmologically significant information can be derived from astronomical observations. It is found that the most detailed information one can reasonably hope to attain by direct observations of distant galaxies can be represented by a "maximal data set", denoted $D(w_o,z^*)$ (see §4.3).

The aim of the rest of the thesis is then to determine to what extent $D(w_o,z^*)$ may be used to determine the structure of the universe. In particular, it is shown in §5.2 that if $D(w_o,z^*)$ is used as initial data for the field equations, then (under certain conditions which will be specified) the model on the light cone is completely determined. This may be regarded as the central result of the thesis.

Two problems that are common to most modern cosmological investigations are highlighted if one adopts this observational approach. Firstly, the paucity of cosmological observations means that the required initial data is not currently available, nor is it likely to become available in the forseeable future. Various ways of dealing with this deficiency will be discussed. One possibility is to make assumptions about the observations, and this gives rise to various *observational space-times* associated with a particular class of observations (see §4.5). An observationally isotropic space-time, e.g., is the equivalence class of space-times in which all cosmological observations about one point are isotropic. Thus it is suggested that symmetry assumptions might well have to be made, but these would relate to the actual observations rather than to the space-time itself. Such assumptions could, in principle, be checked at one space-time point.

Furthermore, the characteristic initial value problem, as formulated in §5.2, requires that the initial data on the initial past light cone $C^-(p)$ be prescribed as C^∞ functions of three variables, whereas, in practice, the actual measurements yield discrete sets of data points to which smooth curves are then fitted by a variety of statistical techniques. In §§5.4 - 6 we shall assume that the initial data is specified in polynomial form (with bounded remainder terms), and then actually carry out the integration of the field equations explicitly up to the highest order possible.

The second problem arising from the observational approach relates to the hyperbolic nature of the field equations. In general, the characteristic initial value problem is not well-posed if data is specified only on one past light cone $C^-(p)$. Clearly, in order to predict to the future of $C^-(p)$, one has to adopt some kind of "predictability criterion", either by assuming analyticity ([56]), or by specifying (arbitrarily) initial data on some intersecting hypersurface ([57]), or by assuming that the space-time satisfies certain global conditions that render it "deterministic" ([58]). Even within the past light cone, the existence of limiting redshifts restricts the maximal development. We have not investigated this problem in any detail in this thesis: it appears that the NP-formalism is not ideally suited to a rigorous discussion of the propagation of the solution off $C^-(p)$, except possibly in the analytic case. Nevertheless, in §5.3 (cf. also §5.6), it will be argued that the solution on $C^-(p)$

some ways the cosmological analogue of the "linguistic turn" in philosophy ([63]), we wish to emphasise, from a new perspective, the vital role of observations in what is often claimed to be - and should be - an empirical science.

2. EXPLICIT SOLUTIONS FOR SPATIALLY HOMOGENEOUS MODELS

§2.1. *Introduction*

Spatially homogeneous models (see, e.g., [31],[34]) are based on space-times that admit a group G_r ($r = 3,4,6$) of isometries acting transitively on 3-dimensional spacelike orbits. The so-called Bianchi models arise when the G_r acts simply-transitively on the group orbits. In this case, it is always possible to find coordinates (t,x^ν), $(\nu = 1,2,3)$, such that the metric is given by

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

and the group orbits are the spacelike hypersurfaces $\{t = \text{constant}\}$. (In this chapter only, we shall work with a metric of signature $(-+++)$.)

If $r = 6$ or $r = 4$, then, in addition to the simply-transitive G_3 subgroup, there is (respectively) a 3- or 1-dimensional isotropy group acting in these hypersurfaces. A space-time is called Locally Rotationally Symmetric (LRS) if it admits a spatial isotropy group at every point ([64]).

Throughout this chapter we shall take the stress-energy tensor to have the perfect fluid form

$$T_{ab} = (\mu+p)u_a u_b + pg_{ab} \quad , \quad u^a u_a = -1, \tag{2.1.1}$$

with the total energy density μ and the isotropic pressure p related by an isentropic equation of state (see, e.g., [31])

$$p = (\gamma-1)\mu \quad , \quad 1 \le \gamma \le 2. \tag{2.1.2}$$

The fluid 4-velocity may be written as ([43])

$$u^a = n^a \cosh\beta + c^a \sinh\beta \tag{2.1.3}$$

where n^a is the unit normal to, and c^a is a unit spacelike vector lying in, the group orbits; β is called the "hyperbolic determined from $D(w_o, z^*)$ may be propagated off $C^-(p)$ without additional specification of "transverse" derivatives. The argument used is essentially that employed by Bondi et. al. ([59]), Sachs ([60]), Newman and Unti ([61]) and Tamburino and Winicour ([62]). This argument would appear to guarantee existence and uniqueness in the analytic case, and it seems likely that the result may be extended to the non-analytic case by confining attention to the interior of $C^-(p)$ and using techniques similar to those employed by Müller zum Hagen and Seifert ([57]). But the problem of finding a suitable predictability criterion remains unsolved.

It should be emphasised that the approach we are advocating can be used in conjunction with more conventional methods. Thus one might, as in Chapter 6, assume a priori that the space-time is spherically symmetric about one world line, and then undertake an observational analysis. This allows one to answer such questions as: what observations in a universe that is isotropic about one world line will imply that it is isotropic about all world lines?

Finally, it should not be supposed that we are attempting to construct a presuppositionless theory of cosmology. We shall be only too conventional in adopting the fundamental principles and the perfect fluid idealisation discussed in §1.3. Furthermore, as we have tried to emphasise in §1.2, a complete cosmological theory consists of more than just a geometrical description of the universe (which is all that a solution of the initial value problem would yield, even if indefinitely accurate measurements of the required initial data were to become available). But, by adopting an approach that is in tilt angle": if $\beta = 0$, the models are called "orthogonal", while "tilted" models arise when $\beta \neq 0$.

In all spatially homogeneous models, a suitable choice of coordinates reduces the field equations

$$R_{ab} - \frac{1}{2}Rg_{ab} = T_{ab} \tag{2.1.4}$$

to ordinary differential equations with independent variable t. The standard techniques for analysing the qualitative behaviour of the solutions may then be applied (see, e.g., [42],[61]). However, there appear to be very few systematic procedures available to obtain exact (closed form) solutions of these equations. Wainwright et.al. ([65]) have recently shown how solutions for stiff matter models ($\gamma = 2$) may be obtained from the vacuum solution. In this chapter we present a systematic method for obtaining first integrals of linear combinations of the field equations. This often allows one to reduce the system of equations to one ordinary differential equation plus a set of quadratures, or to obtain closed form solutions.

We shall first present the method, which is based on the concept of a decomposable differential operator ([66], included as Appendix IV), and then give two examples to show how it is applied. Both solutions given in §2.3 are, as far as we are aware, new (although they first appeared in a joint paper with Maartens, [66]).

§2.2. Decomposability

A cursory examination of the field equations listed in §2.3 reveals that they are all of the generic form

$$A_1\frac{\ddot{X}}{X} + A_2\frac{\ddot{Y}}{Y} + A_3\frac{\dot{X}^2}{X^2} + A_4\frac{\dot{Y}^2}{Y^2} + A_5\frac{\dot{X}}{X} + A_6\frac{\dot{Y}}{Y} + A_7\frac{\dot{X}\dot{Y}}{XY} +$$

$$+ H(X,Y,t) = 0 \qquad (2.2.1)$$

where $A_i \epsilon R$ $(i=1,2,\ldots,7)$, a dot denotes differentiation with respect to the independent variable t, and X(t), Y(t) are metric components. The terms with coefficients A_5 and A_6 seldom appear explicitly in the usual coordinate formulation; they often arise when a new independent variable is introduced. Also, we are assuming that μ and p have been solved for from the conservation equations in terms of X and Y. Contributions from the stress-energy tensor are thus represented by the function H(X,Y,t).

Very simply, the method described below involves finding an integrating factor for (2.2.1) which depends explicitly only on the variables X, Y and t (and not, as would be the case in general, on $\dot{X}$ and $\dot{Y}$). By transforming to new variables

$$u^1 = \log|X| \ , \quad u^2 = \log|Y| \ , \qquad (2.2.2)$$

we may rewrite (2.2.1) in the form

$$L_{\underline{a}}\{u^A\}(t) + H(u^A,t) = 0 \ , \quad (A=1,2) \ , \qquad (2.2.3)$$

where

$$L_{\underline{a}}\colon C^k(R) \times C^k(R) \to C^{k-2}(R) \ , \quad (k\geq 4) \ ,$$

is a differential operator given by

$$L_{\underline{a}}\{u^A\}(t) \equiv \alpha_A \ddot{u}^A + \gamma_{AB}\dot{u}^A\dot{u}^B + \beta_A \dot{u}^A \ , \qquad (2.2.4)$$

with

$$\underline{a} \equiv (a_1,\ldots,a_7) \equiv (A_1,A_2,A_1+A_3,A_2+A_4,A_5,A_6,A_7) \qquad (2.2.5a)$$

and

$$\alpha_A \equiv \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} , \quad \beta_A \equiv \begin{pmatrix} a_5 \\ a_6 \end{pmatrix} , \quad \gamma_{AB} \equiv \begin{pmatrix} a_3 & \tfrac{1}{2}a_7 \\ \tfrac{1}{2}a_7 & a_4 \end{pmatrix} . \qquad (2.2.5b)$$

(Note that these are all constant matrices.)

The operator $L_{\underline{a}}$ will be said to be *decomposable* if there exist functions $F = F(u^A,t)$ and $G = G(\dot{u}^A,u^A,t)$ of differentiability class k-2 and k-1, respectively, in their arguments, such that

$$L_{\underline{a}}\{u^A\}(t) = F(u^A,t)\dot{G}(\dot{u}^A,u^A,t) \ . \tag{2.2.6}$$

Clearly F^{-1} is then an integrating factor for the equation $L_{\underline{a}}\{u^A\}(t) = 0$, so that (2.2.3) may be integrated to

$$G[u^A, \frac{du^A}{d\tau}, t(\tau)] = \tau_0-\tau \quad , \quad (\tau_0\epsilon R) \ , \tag{2.2.7}$$

where we have introduced a new independent variable, τ , by

$$\tau(t) \equiv \int H[u^A(t'),t']F^{-1}[u^A(t'),t']dt' \tag{2.2.8}$$

The first complete characterisation of decomposability for operators of the form (2.2.4) was given in [66], where it was shown that $L_{\underline{a}}$ is decomposable iff

$$(a_1)^2a_4 + (a_2)^2a_3 - a_1a_2a_7 = 0 \tag{2.2.9a}$$

and

$$(a_5)^2a_4 + (a_6)^2a_3 - a_5a_6a_7 = 0 \ . \tag{2.2.9b}$$

However, the proof presented there required consideration of a number of cases and subcases. This may be avoided by the simple device of treating α_A, β_A and γ_{AB} as spinors. Thus indices are raised and lowered by

$$\alpha^A = \varepsilon^{AB}\alpha_B \ , \quad \alpha_B = \alpha^A\varepsilon_{AB} \ , \tag{2.2.10}$$

where

$$\varepsilon^{AB} = \varepsilon^{[AB]}, \quad \varepsilon^{01} = 1 \ , \tag{2.2.11a}$$

$$\varepsilon_{AB} = \varepsilon_{[AB]}, \quad \varepsilon_{01} = 1 \ . \tag{2.2.11b}$$

It is convenient to define

$$v^A \equiv \dot{u}^A \ . \tag{2.2.12}$$

THEOREM 2.1.

$$L_{\underline{a}}\{u^A\} = \alpha_A \dot{v}^A + \gamma_{AB} v^A v^B + \beta_A v^A \quad , \text{ with } \alpha_A \neq 0$$

is decomposable iff

$$\alpha^A \alpha^B \gamma_{AB} = 0 \tag{2.2.13a}$$

and $$\beta^A \beta^B \gamma_{AB} = 0. \tag{2.2.13b}$$

(It follows immediately from (2.2.5b) and (2.2.10) that (2.2.13) is equivalent to (2.2.9).)

Proof. Necessity:

Suppose

$$\alpha_A \dot{v}^A + \gamma_{AB} v^A v^B + \beta_A v^A = F(u^A,t)\dot{G}(v^A,u^A,t). \tag{2.2.14a}$$

Since

$$\dot{G} = \frac{\partial G}{\partial v^A} \dot{v}^A + \frac{\partial G}{\partial u^A} v^A + \frac{\partial G}{\partial t} \tag{2.2.14b}$$

it follows that

$$\alpha_A = F \frac{\partial G}{\partial v^A} \tag{2.2.15a}$$

and

$$\gamma_{AB} v^A v^B + \beta_A v^A = F \left[\frac{\partial G}{\partial u^A} v^A + \frac{\partial G}{\partial t} \right] \tag{2.2.15b}$$

Differentiation of (2.2.15b) with respect to v^C gives

$$2\gamma_{AC} v^A + \beta_C = F \left[\frac{\partial^2 G}{\partial u^A \partial v^C} v^A + \frac{\partial G}{\partial u^C} + \frac{\partial^2 G}{\partial t \partial v^C} \right] \tag{2.2.16}$$

Further differentiation (using the fact that $\dfrac{\partial^2 G}{\partial v^A \partial v^B} = 0$, which follows from (2.2.15a)) gives

$$\gamma_{AB} = F \frac{\partial^2 G}{\partial u^{(A} \partial v^{B)}} \tag{2.2.17}$$

where round brackets denote symmetrisation.

Hence, by (2.2.15a)

$$\gamma_{AB} = -F^{-1} \alpha_{(A} \frac{\partial F}{\partial u^{B)}} \tag{2.2.18}$$

But, since $\alpha^A \alpha_A = 0$, by (2.2.10), it follows immediately from (2.2.18) that

$$\alpha^A \alpha^B \gamma_{AB} = 0$$

thus establishing the first decomposability condition, (2.2.13a).

Now this is just the condition that α_A be a principal null direction of the symmetric spinor γ_{AB}. Thus (see, e.g., [73])

$$\gamma_{AB} = \alpha_{(A}\rho_{B)} \text{ , some constant spinor } \rho_A \text{ .} \tag{2.2.19}$$

Now (2.2.18) gives

$$\frac{\partial F}{\partial u^A} = - F\rho_A \tag{2.2.20}$$

while substitution of (2.2.17) into (2.2.16) yields

$$\beta_C = F \left[\frac{\partial G}{\partial u^C} + \frac{\partial^2 G}{\partial t \partial v^C} - \frac{\partial^2 G}{\partial u^C \partial v^A} v^A \right] \tag{2.2.21}$$

Thus

$$F^{-2} \beta_{[C} \frac{\partial F}{\partial u^{B]}} = - \frac{\partial^3 G}{\partial t \partial v^{[C} \partial u^{B]}} = \frac{\partial}{\partial t} \left\{ F^{-2} \alpha_{[C} \frac{\partial F}{\partial u^{B]}} \right\}$$

so that, by (2.2.20)

$$\beta_{[C}\rho_{B]} = - \alpha_{[C}\rho_{B]} F^{-1} \frac{\partial F}{\partial t} \tag{2.2.22}$$

On the other hand, it follows from (2.2.15b) and (2.2.16) that

$$\gamma_{AC} v^A v^C = F \left[\frac{\partial^2 G}{\partial u^A \partial v^C} v^A v^C + \frac{\partial^2 G}{\partial t \partial v^C} v^C - \frac{\partial G}{\partial t} \right] \tag{2.2.23}$$

and hence, by (2.2.17)

$$\frac{\partial^2 G}{\partial t \partial v^C} v^C - \frac{\partial G}{\partial t} = 0 \tag{2.2.24}$$

Furthermore, (2.2.15a) and (2.2.21) imply

$$\alpha^C \beta_C = F \alpha^C \left[\frac{\partial G}{\partial u^C} - \frac{\partial^2 G}{\partial u^C \partial v^A} v^A \right]$$

which, after differentiation with respect to t, and use of (2.2.24), gives

$$\alpha^C \beta_C \frac{\partial F}{\partial t} = 0. \tag{2.2.25}$$

On comparing (2.2.22) and (2.2.25), we therefore have

$$\beta^A \beta^B \alpha_A \rho_B = 0$$

which, by virtue of (2.2.20), establishes the second condition, (2.2.13b)

Sufficiency.

Suppose (2.2.13) holds. Then (2.2.19), which now follows from the fact that α_A does not vanish identically, defines a constant spinor ρ_A. Let

$$F = \exp[-\rho_A u^A + \sigma t] \tag{2.2.26a}$$

and

$$G = F^{-1}[\alpha_A v^A + \kappa] \tag{2.2.26b}$$

where σ and κ are constants defined as follows:

$$\text{If} \quad \alpha^A \beta_A = 0, \text{ then } \beta_A = \sigma\alpha_A \quad \text{and } \kappa = 0. \tag{2.2.27a}$$

$$\text{If} \quad \alpha^A \beta_A \neq 0, \text{ then } \beta_A = \kappa\rho_A \quad \text{and } \sigma = 0. \tag{2.2.27b}$$

Then

$$\begin{aligned} F(u^A,t)\dot{G}(v^A,u^A,t) &= \alpha_A \dot{v}^A + \alpha_A \rho_B v^A v^B + [\kappa\rho_A + \sigma\alpha_A] v^A \\ &= \alpha_A \dot{v}^A + \gamma_{AB} v^A v^B + \beta_A v^A \end{aligned}$$

where the last equality follows from (2.2.12) and (2.2.27). $\square$

Maartens ([5],[66]) has shown that the decomposability conditions (2.2.9) are invariant under all transformations of variable that leave the form of (2.2.1) invariant. Thus, if $L_{\underline{a}}$ is not decomposable, no form-preserving transformation of variable can render it decomposable.

§2.3. Examples.

By means of illustrative examples, we show how the results of the foregoing discussion may be used to obtain exact solutions of the field equations for a variety of spatially homogeneous models. We shall assume throughout that the cosmological constant vanishes.

Example 1: Orthogonal, L.R.S. models

For these models, there exist co-moving coordinates $(x^a) \equiv (t,x,y,z)$ such that the metric is given by

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)[dy^2 + f^2(y)]dz^2 - X^2(t)h(y)[2dx - h(y)dz]dz \qquad (2.3.1)$$

where

$$f(y) = \begin{pmatrix} \sin y \\ y \\ \sinh y \end{pmatrix}, \quad h(y) = \begin{pmatrix} 2c \cos y \\ -c^2 y^2 \\ -2c \cosh y \end{pmatrix}, \text{ if } k = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad (2.3.2)$$

and c,k are parameters related to the symmetry group of the space-time.

The fluid 4-velocity is

$$u^a = \delta^a_o \qquad (2.3.3)$$

and the field equations are

$$\frac{\ddot{X}}{X} + \frac{\dot{X}\dot{Y}}{XY} + \frac{\ddot{Y}}{Y} + c^2 \frac{X^2}{Y^4} = (1-\gamma)\mu \qquad (2.3.4a)$$

$$2\frac{\ddot{Y}}{Y} + \frac{\dot{Y}^2 + k}{Y^2} - 3c^2\frac{X^2}{Y^4} = (1-\gamma)\mu \qquad (2.3.4b)$$

$$2\frac{\dot{X}\dot{Y}}{XY} + \frac{\dot{Y}^2 + k}{Y^2} - c^2\frac{X^2}{Y^4} = \mu \qquad (2.3.4c)$$

while the conservation equations, $T^{ab}{}_{;b} = 0$, reduce to

$$\frac{1}{\gamma}\frac{\dot{\mu}}{\mu} + \frac{\dot{X}}{X} + 2\frac{\dot{Y}}{Y} = 0 \qquad (2.3.4d)$$

We shall present here only the solution for the case $ck \neq 0$, corresponding to Bianchi Types VIII ($k = -1$) and IX ($k = 1$), with a "stiff matter" equation of state, $\gamma = 2$.

Equations (2.3.4a,c) give

$$\frac{\ddot{X}}{X} + \frac{\ddot{Y}}{Y} + \frac{\dot{Y}^2}{Y^2} + 3\frac{\dot{X}\dot{Y}}{XY} + \frac{k}{Y^2} = 0 \qquad (2.3.5)$$

By performing the transformation (2.2.2), we may write this in the form (2.2.4), with

$$\alpha_A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \gamma_{AB} = \begin{pmatrix} 1 & 3/2 \\ 3/2 & 2 \end{pmatrix}, \quad \beta_A = 0 \qquad (2.3.6a)$$

It is easily verified that (2.2.13) is satisfied, and from (2.2.19) we find

$$\rho_A = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad (2.3.6b)$$

Thus (2.3.5) decomposes to

$$(XY^2)^{-1}\{(XY^2)(\frac{\dot{X}}{X} + \frac{\dot{Y}}{Y})\}^{\cdot} = -kY^{-2}$$

or

$$\{Y(XY)^{\cdot}\}^{\cdot} = -kX \qquad (2.3.7)$$

This immediately suggests defining a new time parameter, τ, by

$$\tau(t) = \int X(t')dt' , \qquad (2.3.8)$$

and now (2.3.7) may be integrated twice to yield

$$(XY)^2 = -k\tau^2 + a\tau + b \ , \ (a,b \in R) \tag{2.3.9}$$

Integration of (2.3.4d) yields

$$\mu = e^2 X^{-2} Y^{-4} \ , \ (e \in R) \tag{2.3.10}$$

Now substitution of Y and μ from (2.3.9,10) into (2.3.4c) provides a differential equation for $X(\tau)$:

$$\left\{ \frac{d}{d\tau}(X^{-2}) \right\}^2 = \{(a^2 + 4bk - e^2)X^{-2} - 4c^2\} \times$$
$$\times (-k\tau^2 + a\tau + b)^{-2} \tag{2.3.11}$$

This is separable, and thus $X(\tau)$ may be determined by a simple quadrature, after which $Y(\tau)$ and $\mu(\tau)$ follow from (2.3.9,10). The full solution is therefore obtained in parametric form, where the proper time t is related to the parameter τ by

$$t(\tau) = \int X^{-1}(\tau')d\tau' \tag{2.3.12}$$

The relation between the vacuum and $\gamma = 2$ solutions (cf. [65]) is particularly simple in this case: the vacuum solution is found by setting the constant e to zero throughout.

Example 2: Tilted, L.R.S. models

The simplest models in this class are the Bianchi Type V models, for which there exist co-moving coordinates $(x^a) \equiv (t,x,y,z)$ such that the metric is given by

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)\exp(-2A_o x)(dy^2 + dz^2) \tag{2.3.13}$$

where $A_o \in R$, $A_o \neq 0$. The fluid 4-velocity is

$$u^a = (\cosh\beta)\delta^a_o + (X^{-1}\sinh\beta)\delta^a_1 \tag{2.3.14}$$

where β is the hyperbolic tilt angle (see §2.1).

The field equations are

$$\frac{\ddot{X}}{X} + 2\,\frac{\dot{X}\dot{Y}}{XY} - 2\,\frac{A_o^2}{X^2} = \tfrac{1}{2}(2-\gamma)\mu + \gamma\mu\sinh^2\beta \tag{2.3.15a}$$

$$\frac{\ddot{Y}}{Y} + \frac{\dot{Y}^2}{Y^2} + \frac{\dot{X}\dot{Y}}{XY} - 2\,\frac{A_o^2}{X^2} = \tfrac{1}{2}(2-\gamma)\mu \tag{2.3.15b}$$

$$2\,\frac{A_o}{X}\left(\frac{\dot{X}}{X} - \frac{\dot{Y}}{Y}\right) = \gamma\mu\sinh\beta\cosh\beta \tag{2.3.15c}$$

$$\frac{\dot{Y}^2}{Y^2} + 2\,\frac{\dot{X}\dot{Y}}{XY} - 3\,\frac{A_o^2}{X^2} = \mu\cosh^2\beta + (\gamma-1)\sinh^2\beta \tag{2.3.15d}$$

while the conservation equations can be written as

$$\{\log(\mu^\gamma XY^2\cosh\beta)\}^{\cdot} = 2\,\frac{A_o}{X}\tanh\beta \tag{2.3.16a}$$

$$\{\log[\mu^{(\gamma-1)/\gamma}X\sinh\beta]\}^{\cdot} = 0 \tag{2.3.16b}$$

Again the case $\gamma = 2$ will be solved. A special case of this solution was first found jointly with Maartens ([66]).

By (2.3.16b)

$$\mu = A_o^2(\alpha X\sinh\beta)^{-2} \quad , \quad (\alpha \in R) \tag{2.3.17}$$

But (2.3.15a,b) imply

$$\frac{\ddot{X}}{X} - \frac{\ddot{Y}}{Y} - \frac{\dot{Y}^2}{Y^2} + \frac{\dot{X}\dot{Y}}{XY} = 2\mu\sinh^2\beta = 2A_o^2(\alpha X)^{-2} \tag{2.3.18}$$

where the second equality follows from (2.3.17).

Now (2.3.18) decomposes to give

$$\left\{ (XY^2)\left(\frac{\dot{X}}{X} - \frac{\dot{Y}}{Y}\right) \right\}^{\cdot} = 2A_o^2\alpha^{-2}X^{-1}Y^2 \ , \tag{2.3.19}$$

while (2.3.15b) also decomposes and yields

$$\{ X(Y^2)^{\cdot} \}^{\cdot} = 4A_o^2X^{-1}Y^2 \tag{2.3.20}$$

Let

$$\xi = \int X^{-1}(t')dt' \tag{2.3.21}$$

Then (2.3.19) and (2.3.20) become, respectively

$$\frac{d}{d\xi}\{\, Y^2 \left(\frac{1}{X}\frac{dX}{d\xi} - \frac{1}{Y}\frac{dY}{d\xi} \right) \} = 2A_o^2\alpha^{-2}Y^2 \qquad (2.3.22)$$

and

$$\frac{d^2}{d\xi^2}(Y^2) - 4A_o^2Y^2 = 0 \qquad (2.3.23)$$

Now (2.3.23) integrates immediately to give

$$Y^2(\xi) = a\cosh(2A_o\xi) + b\sinh(2A_o\xi) \quad , \quad (a,b \in R) \qquad (2.3.24)$$

In order to integrate (2.3.22) we use (2.3.24):

$$\begin{aligned} Y^2 \frac{d}{d\xi}\{\log\left(\frac{X}{Y}\right)\} &= 2A_o^2\alpha^{-2}\int Y^2(\xi)\,d\xi \\ &= \tfrac{1}{2}\alpha^{-2}\frac{d}{d\xi}(Y^2) + c \quad , \quad (c \in R) \end{aligned} \qquad (2.3.25)$$

and hence

$$X(\xi) = e\, Y^{(1+\alpha^{-2})}\exp\{\, c\int Y^{-2}(\xi)\,d\xi\} \quad , \quad (e \in R) \qquad (2.3.26)$$

In order to find $\beta(\xi)$ and $\mu(\xi)$, note that, under the transformation (2.3.21), equation (2.3.15c) becomes

$$\begin{aligned} A_o\frac{d}{d\xi}\{\log\left(\frac{X}{Y}\right)\} &= \mu\sinh\beta\cosh\beta \\ &= A_o^2\alpha^{-2}X^{-2}\coth\beta \end{aligned}$$

where the last equality follows from (2.3.17). Hence, by (2.3.25)

$$\coth\beta = A_o^{-1}X^2Y^{-2}\left(\alpha^2 c + Y\frac{dY}{d\xi}\right) \qquad (2.3.27)$$

and now μ may be obtained directly from (2.3.17).

Once again, the solution has therefore been found in parametric form, with the proper time t related to the parameter ξ by

$$t(\xi) = \int X(\xi)\,d\xi \qquad (2.3.28)$$

In both the examples considered in this section, the solutions have been left in terms of elementary quadratures, since the actual properties of the solutions will not concern us here. Further examples may be found in Appendix IV.

3. CENTRAL CONDITIONS FOR OBSERVATIONAL COSMOLOGY

§3.1. Introduction

The aim of this chapter is two-fold: firstly, we introduce the Newman-Penrose (NP) formalism and an associated null coordinate system that will be used extensively in the sequel. Secondly, from the assumption that the central world line of the observer is a regular geodesic, we derive the limiting behaviour of the spin coefficients and the metric and stress-energy components in the vicinity of this world line. These "central conditions" (see also Appendix I of [2]) are crucial to the rest of the analysis (see, in particular, Chapter 5). The method used to derive these conditions is an adaptation of techniques employed by Newman, Unti and Posadas in their study of vacuum and Einstein-Maxwell fields (see, e.g., ([61],[67],[68]).

§3.2. The Newman-Penrose formalism

The Newman-Penrose formalism ([55]) is based on a complex null tetrad field

$$\{e_{\underline{a}}\} = \{\underline{n},\underline{k},\underline{m},\underline{\bar{m}}\}$$

where $\underline{n},\underline{k}$ are real null vector fields and $\underline{m}$ is a complex null vector field whose real and imaginary parts are both spacelike. The only non-zero scalar products between these vectors are

$$\underline{k}.\underline{n} = 1 = -\underline{m}.\underline{\bar{m}} \tag{3.2.1}$$

and thus the metric tensor is given by the completeness relation

$$g_{ab} = 2\{ k_{(a}n_{b)} - m_{(a}\bar{m}_{b)} \} \tag{3.2.2}$$

The affine connection is described by the 12 so-called spin coefficients:

$$\kappa = k_{a;b}m^a k^b \qquad \nu = -n_{a;b}\bar{m}^a n^b$$

$$\sigma = k_{a;b}m^a m^b \qquad \lambda = -n_{a;b}\bar{m}^a \bar{m}^b$$

$$\rho = k_{a;b}m^a \bar{m}^b \qquad \mu = -n_{a;b}\bar{m}^a m^b$$

$$\tau = k_{a;b}m^a n^b \qquad \pi = -n_{a;b}\bar{m}^a k^b \tag{3.2.3}$$

$$\beta = \tfrac{1}{2}(k_{a;b}n^a m^b - m_{a;b}\bar{m}^a m^b)$$

$$\alpha = \tfrac{1}{2}(k_{a;b}n^a \bar{m}^b - m_{a;b}\bar{m}^a \bar{m}^b)$$

$$\varepsilon = \tfrac{1}{2}(k_{a;b}n^a k^b - m_{a;b}\bar{m}^a k^b)$$

$$\gamma = \tfrac{1}{2}(k_{a;b}n^a n^b - m_{a;b}\bar{m}^a n^b)$$

The irreducible pasrt of the curvature tensor R_{abcd} are represented as follows: the Weyl tensor C_{abcd} is described by

$$\Psi_0 = - C_{abcd}k^a m^b k^c m^d$$

$$\Psi_1 = - C_{abcd}k^a n^b k^c m^d$$

$$\Psi_2 = - C_{abcd}\bar{m}^a n^b k^c m^d \tag{3.2.4a}$$

$$\Psi_3 = C_{abcd}k^a n^b n^c \bar{m}^d$$

$$\Psi_4 = - C_{abcd}n^a \bar{m}^b n^c \bar{m}^d ,$$

the trace free part of the Ricci tensor by

$$\Phi_{00} = - \tfrac{1}{2} R_{ab}k^a k^b \qquad \Phi_{11} = - \tfrac{1}{4} R_{ab}(k^a n^b + m^a \bar{m}^b)$$

$$\Phi_{01} = - \tfrac{1}{2} R_{ab}k^a m^b \qquad \Phi_{12} = - \tfrac{1}{2} R_{ab}n^a m^b \tag{3.2.4b}$$

$$\Phi_{02} = - \tfrac{1}{2} R_{ab}m^a m^b \qquad \Phi_{22} = - \tfrac{1}{2} R_{ab}n^a n^b$$

with

$$\Phi_{IJ} = \bar{\Phi}_{IJ} , \quad (I,J = 0,1,2) \tag{3.2.4c}$$

while the Ricci scalar is given by

$$\Lambda = \frac{1}{24} R \tag{3.2.4d}$$

The tetrad derivatives are assigned the symbols

$$\Delta \equiv n^a \nabla_{\underline{a}} \ , \quad D \equiv k^a \nabla_{\underline{a}} \ , \quad \delta \equiv m^a \nabla_{\underline{a}} \tag{3.2.5}$$

The Ricci, Bianchi, and contracted Bianchi identities may now be expressed in terms of the spin coefficients and the irreducible parts of the curvature tensor (see, e.g., [69], but note that there are two misprints in the equations given there). These identities will be given in §3.4, after we have introduced a special coordinate system and null tetrad.

§3.3. Null coordinates and null tetrad

§3.3.1

The coordinates we shall use are based on the world line of our galaxy (hereafter referred to as the "central world line" and denoted C) and on the family of past light cones $C^-(p)$ of points p on C. The coordinates $(x^a) \equiv (w, y, x^A)$ are defined as follows:

(i) $x^0 = w$: Let the surfaces {w = constant} be the past light cones $C^-(p)$ of events on C. Normalise w by requiring that it should measure proper time t along C, and let $w = w_0$ correspond to the event "here and now". This defines w uniquely once the constant w_0 has been chosen. It is a function which is C^1 almost everywhere within a simply convex normal neighbourhood of the point $w = w_0$ on C, but is not C^1 on C. Thus the one-form dw is not well-defined on C, but the vector field $\frac{\partial}{\partial w}$ is well-defined on C; in fact, the tangent vector to C is given by

$$\left.\frac{\partial}{\partial w}\right|_C = \left.\frac{\partial}{\partial t}\right|_C \tag{3.3.1}$$

The past-directed null geodesic tangent vector field whose integral curves are the ruling geodesics of the past light cones $C^-(p)$ is then

$$\underline{k} = \frac{\partial}{\partial v}\ , \text{ i.e. } k^a = \frac{dx^a}{dv} \tag{3.3.2a}$$

where

$$k_a = -w,_a \quad , \quad k^a k_a = 0, \tag{3.3.2b}$$

and v is an affine parameter down these geodesics.

Now C is itself a member of the time-like congruence with tangent vector

$$u^a = \frac{dx^a}{dt}. \tag{3.3.3}$$

Thus

$$u^a k_a = -u^a w,_a = -u^0 = -\frac{dw}{dt}\ , \tag{3.3.4a}$$

so that

$$\lim_{v \to 0} (\underline{u}.\underline{k}) = -1\ , \tag{3.3.4b}$$

which means that

$$\left(\frac{\partial}{\partial w}.\underline{k}\right)\Big|_C = -1\ . \tag{3.3.4c}$$

(The minus sign in (3.3.4b,c) reflects the fact that $\underline{k}$ is past-directed, while $\underline{u}$ and $\frac{\partial}{\partial w}$ are both future directed.)

We now impose the condition that the central world line should correspond to $v = 0$; it then follows from (3.3.2a,4c) that v is uniquely determined.

(ii) $\underline{x^1 = y}$: Choose y to be some parameter (not necessarily affine) down the null geodesics ruling $C^-(p)$ and let $y = 0$ on C. A number of choices present themselves; e.g. we might take

(a) $y = v$, where v is the unique affine parameter above,

(b) $y = z$, where z is the redshift (see §4.2) of events as observed from C,

(c) $y = r$, where r is the observer area distance (see §4.2) of events with respect to C.

In general, if y is chosen as one of (a)-(c) in an open neighbourhood of C, it will not be co-moving with the fluid. Alternatively, one may specify y to be one of (a)-(c) on *one* light cone, and then drag y off this light cone by the fluid 4-velocity $\underline{u}$. In this case one has

$$\mathcal{L}_{\underline{u}}\, y = 0 \quad \text{iff} \quad y_{,a}u^a = 0 \quad \text{iff} \quad u^1 = 0, \tag{3.3.5}$$

where $\mathcal{L}$ denotes the Lie derivative. Given a choice of y from the possibilities (a)-(c), either of these alternatives determines y uniquely.

(iii) x^A: Label the null geodesics ruling the past light cones by the "angular coordinates" x^A. (This terminology will be justified in §3.4.) Thus

$$\mathcal{L}_{\underline{k}}\, x^A = 0, \quad \text{or} \quad x^A{}_{,a}k^a = 0, \tag{3.3.6a}$$

so that

$$k^a = \left(\frac{dy}{dv}\right)\delta^A_1 \,. \tag{3.3.6b}$$

This determines x^A up to a relabelling of the null generators:

$$x^{A'} = x^{A'}(w, x^A) \tag{3.3.7}$$

Clearly, these coordinates do not cover all of space-time. Firstly, there is the standard singularity associated with the angular coordinates on C. But even off C, the coordinates do not necessarily give a one-one covering of that part of space-time that is observable from C.

They may give a many-one covering: this will occur when the light cones develop caustics, either because the space-sections are compact, or because the gravitational lens effect ([70]) may cause neighbouring null geodesics to join the same event to the observer. Thus it may be possible to observe the same object in different parts of the sky, at different redshifts and at different times in its history. Actually to decide what identifications need to be made amongst the observed objects will therefore be difficult - nevertheless, such (possibly non-local) identifications need to be made in order to assign local coordinates to events. This problem will be discussed in more detail in §4.1.

On the other hand, these coordinates could also give a one-many covering (i.e. the same coordinates might refer to distinct events). This could occur if r or z were to be used as the y coordinate since, in general, there is no guarantee that either of these quantities is a monotonic function of affine distance down the null geodesics. In a spherically symmetric, static space-time, e.g., it seems that the area distance is not a monotonic function of the redshift ([21]), and thus one of these quantaties is not a monotonic function of v.

§*3.3.2.*

The particular choice of y adopted will depend on the problem at hand. If one is interested in obtaining exact solutions of the field equations, a co-moving coordinate offers the advantage that if, in addition, the fluid 4-velocity components u^A vanish $(A = 2,3)$, the contracted Bianchi identities may be integrated (this will be the case when the proper motions vanish - see §4.2). Choosing y to be either r or z means that y is an observable coordinate; this is useful for a cosmographic analysis of the kind undertaken in [2],[5]. A distinct disadvantage of both r and z as coordinate choices is that, as Penrose has pointed out ([71]), they are both essentially non-local quantities (the monotonicity problem mentioned above arises because of this). Throughout this thesis, we shall therefore choose $y = v$ *in a neighbourhood of C.*

A further advantage of this choice is that it implies $g_{01} = -1 = g^{01}$, as follows from (3.3.6b) with $\frac{dy}{dv} = 1$. In fact,

$$g_{1a} = g_{ab}k^b = -g_a{}^b w_{,b} = -\delta_a^o, \tag{3.3.8a}$$

while

$$g^{ao} = g^{ab}w_{,b} = -g^{ab}k_b = -k^a = -\delta_1^a . \tag{3.3.8b}$$

Thus the coordinate components of the metric tensor are given by

$$g_{ab} = \begin{pmatrix} g_{00} & -1 & g_{02} & g_{03} \\ -1 & 0 & 0 & 0 \\ g_{02} & 0 & & \\ g_{03} & 0 & & g_{AB} \end{pmatrix} \tag{3.3.8c}$$

$$g^{ab} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & g^{11} & g^{12} & g^{13} \\ 0 & g^{12} & & \\ 0 & g^{13} & & g^{AB} \end{pmatrix} \tag{3.3.8d}$$

where g^{ab} and g_{ab} are related by

$$g_{AB}g^{BC} = \delta_A^C \tag{3.3.8e}$$

$$g^{1A} = g_{OB}g^{AB}, \text{ i.e } g^{12} = (g_{02}g_{33} - g_{03}g_{23})h^{-1},$$

$$g^{13} = (g_{03}g_{22} - g_{02}g_{23})h^{-1} \tag{3.3.8f}$$

$$g^{11} = - g_{00} + g_{0A}g^{1A} = - g_{00} + g_{0A}g_{0B}g^{AB} \qquad (3.3.8g)$$

where

$$h \equiv \det g_{AB} \qquad (3.3.8h)$$

A less direct, but nonetheless significant advantage of the choice y = v is that this is also the coordinate used in most studies of asymptotically flat space-times employing the NP-formalism; one of the most important of these (for our purposes) being the Newman-Unti analysis of the characteristic initial value problem in asymptotically flat space-times ([61]; see also [72]). It turns out that there is a close analogy between this problem and the cosmological initial value problem we are studying. In asymptotically flat space-times, one imposes certain conditions on the metric tensor and the matter distribution on $\mathscr{I}^+$, i.e. at future null infinity (see, e.g., [73]); whereas in our case, the assumption that C is geodesic, timelike and regular imposes certain restrictions on the geometry in the vicinity of C. Using the same y coordinate allows one to exploit this analogy directly.

Of course, v is not an "observational" coordinate, since there is no observational (or cosmographic) method of relating v to either of the observational parameters r or z, and it is precisely for this reason that the results of a cosmographic analysis ([2]) are rather limited. But if a set of field equations is assumed (see chapter 5), the null Raychaudhuri equation may be used to determine v as a function of observable quantaties (cf. equation (5.2.7)). Thus, in a cosmological analysis, the fact that v is not observational poses no great problems.

As yet, the coordinates (x^A) are also not necessarily observational, in the following sense: suppose x_1^A and x_2^A label particular null generators with tangent vectors $\underline{k}_1$ and $\underline{k}_2$ respectively, in the past light cones $C^-(p_1)$ and $C^-(p_2)$ of two distinct points p_1 and p_2 on C. Suppose further that $\underline{k}_1$ and $\underline{k}_2$ are related by Fermi-Walker transport along C, i.e. they point in the "same" null direction with respect to a non-rotating frame on C. Then (x^A) will be observational coordinates iff this implies $x_1^A = x_2^A$, i.e. if generators starting off from C in the same null direction are assigned the same coordinate labels x^A. In [2], the coordinate freedom (3.3.7) is used to make (x^A) observational, by demanding that they be based on the direction cosines of $\underline{k}$ with respect to a parallelly propagated orthonormal tetrad on C (since C is assumed geodesic, parallel transport is equivalent to Fermi-Walker transport).

Here we shall proceed differently. Instead of using (3.3.7) to make (x^A) observational, we shall use this freedom to choose some of the arbitrary functions of integration that arise during the derivation of the central conditions in §3.4. As far as possible, we shall choose these functions to have the same values (up to sign) that they are given in the Newman-Unti (NU) integration. A remarkable consequence of this strategy is that, as will be shown in §3.4, the coordinates (x^A) thus obtained are in fact observational in the sense defined above.

§3.3.3

The most general null tetrad satisfying the normalisation conditions (3.2.1), with $\underline{k}$ tangent to the generators of the past light cones $C^-(p)$, is given by

$$k^a = \delta^a_1 \tag{3.3.9a}$$

$$n^a = -\,\delta^a_o + U\,\delta^a_1 + X^A\,\delta^a_A \tag{3.3.9b}$$

$$m^a = \omega\delta^a_1 + \xi^A\,\delta^a_A \tag{3.3.9c}$$

$(A=2,3)$, where U and X^A are real, while ω and ξ^A are complex arbitrary functions of the coordinates, which are related to the metric components by the completeness relation (3.2.2):

$$g^{11} = 2(U - \omega\bar{\omega}) \tag{3.3.10a}$$

$$g^{1A} = X^A - (\xi^A\bar{\omega} + \bar{\xi}^A\omega) \tag{3.3.10b}$$

$$g^{AB} = -(\xi^A\bar{\xi}^B + \bar{\xi}^A\xi^B) \tag{3.3.10c}$$

This tetrad is well-defined only off C; on C, it is singular because there $\underline{k}$ points in a 2-sphere's worth of null directions (this notion will be made more precise later, cf. (3.4.35) below).

By (3.3.2), $\underline{k}$ is the hypersurface orthogonal ($k_{[a;b]}=0$) null geodesic ($k^a k_a = 0$, $k^a{}_{;b}k^b = 0$) vector field orthogonal to the null hypersurfaces $\{w = \text{constant}\}$ ($w_{,a}k^a = 0$), with v an affine parameter along these geodesics. Hence, by (3.2.3),

$$\kappa = 0, \quad \varepsilon + \bar{\varepsilon} = 0, \quad \rho - \bar{\rho} = 0, \quad \tau = \bar{\alpha} + \beta \tag{3.3.11}$$

With the simplifications induced by (3.2.11), the commutators between the tetrad derivatives D, Δ, and δ may be written as ([55]):

$$\Delta D - D\Delta = (\gamma+\bar{\gamma})D - (\tau+\bar{\pi})\bar{\delta} - (\bar{\tau}+\pi)\delta \tag{3.3.12a}$$

$$\delta D - D\delta = (\bar{\alpha}+\beta-\bar{\pi})D - \sigma\bar{\delta} - (\bar{\rho}+\varepsilon-\bar{\varepsilon})\delta \tag{3.3.12b}$$

$$\delta\Delta - \Delta\delta = -\bar{\nu}D + \bar{\lambda}\bar{\delta} + (\mu-\gamma+\bar{\gamma})\delta \tag{3.3.12c}$$

$$\bar{\delta}\delta - \delta\bar{\delta} = (\bar{\mu}-\mu)D - (\bar{\alpha}-\beta)\bar{\delta} - (\bar{\beta}-\alpha)\delta \tag{3.3.12d}$$

By applying these commutators to each of the coordinates w, v, x^A in turn, the "metric variables" U, X^A, ω, ξ^A and the spin coefficients may be related by the "metric equations ([72]):

$$D\xi^A = (\rho+\varepsilon-\bar{\varepsilon})\xi^A + \sigma\bar{\xi}^A \tag{3.3.13a}$$

$$D\omega = (\rho+\varepsilon-\bar{\varepsilon})\omega + \sigma\bar{\omega} - \tau + \bar{\pi} \tag{3.3.13b}$$

$$DX^A = (\bar{\tau}+\pi)\xi^A + (\tau+\bar{\pi})\bar{\xi}^A \tag{3.3.13c}$$

$$DU = (\bar{\tau}+\pi)\omega + (\tau+\bar{\pi})\bar{\omega} - (\gamma+\bar{\gamma}) \tag{3.3.13d}$$

$$\delta X^A - \Delta\xi^A = (\mu+\bar{\gamma}-\gamma)\xi^A + \bar{\lambda}\bar{\xi}^A \tag{3.3.14a}$$

$$\delta U - \Delta\omega = (\mu+\bar{\gamma}-\gamma)\omega + \bar{\lambda}\bar{\omega} - \bar{\nu} \tag{3.3.14b}$$

$$\delta\bar{\xi}^A - \bar{\delta}\xi^A = (\bar{\beta}-\alpha)\xi^A + (\bar{\alpha}-\beta)\bar{\xi}^A \tag{3.3.14c}$$

$$\delta\bar{\omega} - \bar{\delta}\omega = (\bar{\beta}-\alpha)\omega + (\bar{\alpha}-\beta)\bar{\omega} + (\mu-\bar{\mu}) \tag{3.3.14d}$$

By applying the Ricci identities to the tetrad vectors $\underline{n},\underline{k},\underline{m},\bar{\underline{m}}$, the 18 NP equations are obtained ([72]):

$$D\rho = \rho^2 + \sigma\bar{\sigma} + \Phi_{00} \tag{3.3.15a}$$

$$D\sigma = (2\rho+4\varepsilon)\sigma + \Psi_0 \tag{3.3.15b}$$

$$D\tau = (\tau+\bar{\pi})\rho + (\bar{\tau}+\pi)\sigma + 2\varepsilon\tau + \Psi_1 + \Phi_{01} \tag{3.3.15c}$$

$$D\alpha - \bar{\delta}\varepsilon = (\rho-3\varepsilon)\alpha + \beta\bar{\sigma} - \bar{\beta}\varepsilon + (\rho+\varepsilon)\pi + \Phi_{10} \tag{3.3.15d}$$

$$D\beta - \delta\varepsilon = (\alpha+\pi)\sigma + (\rho+\varepsilon)\beta + (\bar{\pi}-\bar{\alpha})\varepsilon + \Psi_1 \tag{3.3.15e}$$

$$D\gamma - \Delta\varepsilon = (\tau+\bar{\pi})\alpha + (\bar{\tau}+\pi)\beta - (\gamma+\bar{\gamma})\varepsilon + \tau\pi + \\ + \Psi_2 - \Lambda + \Phi_{11} \tag{3.3.15f}$$

$$D\lambda - \bar{\delta}\pi = (\rho-4\varepsilon)\lambda + \bar{\sigma}\mu + (\pi+\alpha-\bar{\beta})\bar{\pi} + \Phi_{20} \tag{3.3.15g}$$

$$D\mu - \delta\pi = \rho\mu + \sigma\lambda + \pi\bar{\pi} - (\bar{\alpha}-\beta)\pi + \Psi_2 + 2\Lambda \tag{3.3.15h}$$

$$D\nu - \Delta\pi = (\pi+\bar{\tau})\mu + (\bar{\pi}+\tau)\lambda + (\gamma-\bar{\gamma})\pi - 2\varepsilon\nu + \\ + \Psi_3 + \Phi_{21} \tag{3.3.15i}$$

$$\Delta\lambda - \bar{\delta}\nu' = (\bar{\gamma}-3\gamma-\mu-\bar{\mu})\lambda + (2\alpha+\pi)\nu - \Psi_4 \tag{3.3.16a}$$

$$\delta\rho - \bar{\delta}\sigma = \rho\tau - \sigma(3\alpha-\bar{\beta}) - \Psi_1 + \Phi_{01} \tag{3.3.16b}$$

$$\delta\alpha - \bar{\delta}\beta = \mu\rho - \lambda\sigma + \alpha\bar{\alpha} + \beta\bar{\beta} - 2\alpha\beta + \varepsilon(\mu-\bar{\mu}) \\ - \Psi_2 + \Lambda + \Phi_{11} \tag{3.3.16c}$$

$$\delta\lambda - \bar{\delta}\mu = (\mu-\bar{\mu})\pi + \mu\bar{\tau} + \lambda(\bar{\alpha}-3\beta) - \Psi_3 + \Phi_{21} \tag{3.3.16d}$$

$$\delta\nu - \Delta\mu = \lambda\bar{\lambda} + (\mu+\gamma+\bar{\gamma})\mu - \bar{\nu}\pi - 2\beta\nu + \Phi_{22} \tag{3.3.16e}$$

$$\delta\gamma - \Delta\beta = \mu\tau - \sigma\nu - \epsilon\bar{\nu} - \beta(\gamma-\bar{\gamma}-\mu) + \alpha\bar{\lambda} + \Phi_{12} \qquad (3.3.16f)$$

$$\delta\tau - \Delta\sigma = (\mu+\bar{\gamma}-3\gamma)\sigma + \bar{\lambda}\rho + 2\beta\tau + \Phi_{02} \qquad (3.3.16g)$$

$$\Delta\rho - \bar{\delta}\tau = (\gamma+\bar{\gamma}-\bar{\mu})\rho - \sigma\lambda - 2\alpha\tau - \Psi_2 - 2\Lambda \qquad (3.3.16h)$$

$$\Delta\alpha - \bar{\delta}\gamma = (\rho+\epsilon)\nu - (\tau+\beta)\lambda + (\bar{\gamma}-\gamma-\bar{\mu})\alpha - \Psi_3 \qquad (3.3.16i)$$

The NP equations imply the Bianchi identities and contracted Bianchi identities ([72]):

$$\begin{aligned} D(\Phi_{01} - \Psi_1) - \delta\Phi_{00} + \bar{\delta}\Psi_0 \\ &= (4\alpha-\pi)\Psi_0 - (4\rho+2\epsilon)\Psi_1 + (\bar{\pi}-2\tau)\Phi_{00} + \\ &+ 2(\rho+\epsilon)\Phi_{01} + 2\sigma\Phi_{10} \end{aligned} \qquad (3.3.17a)$$

$$\begin{aligned} D(\Phi_{11}+\Lambda-\Psi_2) + \bar{\delta}\Psi_1 - \delta\Phi_{10} \\ &= \lambda\Psi_0 + 2(\alpha-\pi)\Psi_1 - 3\rho\Psi_2 - \mu\Phi_{00} + \pi\Phi_{01} + \\ &+ (\bar{\pi}-2\bar{\alpha})\Phi_{10} + 2\rho\Phi_{11} + \sigma\Phi_{20} \end{aligned} \qquad (3.3.17b)$$

$$\begin{aligned} D(\Phi_{21}-\Psi_3) + \bar{\delta}(\Psi_2+2\Lambda) - \delta\Phi_{20} \\ &= 2\lambda\Psi_1 - 3\pi\Psi_2 + 2(\epsilon-\rho)\Psi_3 - 2\mu\Phi_{10} + 2\pi\Phi_{11} + \\ &- (2\bar{\alpha}-\bar{\pi}-2\beta)\Phi_{20} + 2(\rho-\epsilon)\Phi_{21} \end{aligned} \qquad (3.3.17c)$$

$$\begin{aligned} D\Psi_4 - \bar{\delta}(\Phi_{21}+\Psi_3) + \Delta\Phi_{20} \\ &= -3\lambda\Psi_2 + 2(\alpha+2\pi)\Psi_3 + (\rho-4\epsilon)\Psi_4 + 2\nu\Phi_{10} + \\ &- 2\lambda\Phi_{11} + (2\bar{\gamma}-2\gamma-\mu)\Phi_{20} - 2\bar{\beta}\Phi_{21} + \bar{\sigma}\Phi_{22} \end{aligned} \qquad (3.3.17d)$$

$$\begin{aligned} \Delta\Psi_0 - \delta(\Phi_{01}+\Psi_1) + D\Phi_{02} \\ &= (4\gamma-\mu)\Psi_0 - 2(2\tau+\beta)\Psi_1 + 3\sigma\Psi_2 - \bar{\lambda}\Phi_{00} + \\ &+ 2(\bar{\pi}-\beta)\Phi_{01} + 2\sigma\Phi_{11} + (\rho+4\epsilon)\Phi_{02} \end{aligned} \qquad (3.3.18a)$$

$$\begin{aligned} \Delta(3\Psi_1-\Phi_{01}) - \delta(2\Phi_{11}+3\Psi_2) + \bar{\delta}\Phi_{02} + 2D\Phi_{12} \\ &= 3\nu\Psi_0 + 6(\gamma-\mu)\Psi_1 - 9\tau\Psi_2 + 6\sigma\Psi_3 - \bar{\nu}\Phi_{00} + \\ &+ 2(\bar{\mu}-\mu-\gamma)\Phi_{01} - 2\bar{\lambda}\Phi_{10} + 2(\tau+2\bar{\pi})\Phi_{11} + \\ &+ (3\alpha-\bar{\beta}+2\pi)\Phi_{02} + 4\epsilon\Phi_{12} + 2\sigma\Phi_{21} \end{aligned} \qquad (3.3.18b)$$

$$\begin{aligned} \Delta(3\Psi_2-2\Phi_{11}) - \delta(\Phi_{21}+3\Psi_3) + 2\bar{\delta}\Phi_{12} + D\Phi_{22} \\ &= 6\nu\Psi_1 - 9\mu\Psi_2 - 6\bar{\alpha}\Psi_3 + 3\sigma\Psi_4 - 2\nu\Phi_{01} \\ &- 2\bar{\nu}\Phi_{10} + 2(2\bar{\mu}-\mu)\Phi_{11} + 2\lambda\Phi_{02} - \bar{\lambda}\Phi_{20} \\ &+ 2(\pi+\bar{\tau}-2\bar{\beta})\Phi_{12} + 2(\beta+\tau+\bar{\pi})\Phi_{21} - \rho\Phi_{22} \end{aligned} \qquad (3.3.18c)$$

$$\Delta(\Psi_3-\Phi_{21}) - \delta\Psi_4 + \bar{\delta}\Phi_{22}$$
$$= 3\nu\Psi_2 - 2(\gamma+2\mu)\Psi_3 + (4\beta-\tau)\Psi_4 - 2\nu\Phi_{11} +$$
$$- \bar{\nu}\Phi_{20} + 2\lambda\Phi_{12} + 2(\gamma+\bar{\mu})\Phi_{21} - \bar{\tau}\Phi_{22} \qquad (3.3.18d)$$

$$D(\Phi_{11}+3\Lambda) - \delta\Phi_{10} - \bar{\delta}\Phi_{01} + \Delta\Phi_{00}$$
$$= (2\gamma+2\bar{\gamma}-\mu-\bar{\mu})\Phi_{00} + (\pi-2\alpha-2\bar{\tau})\Phi_{01} +$$
$$+ (\bar{\pi}-2\bar{\alpha}-2\tau)\Phi_{10} + 4\rho\Phi_{11} + \bar{\sigma}\Phi_{02} +$$
$$+ \sigma\Phi_{20} \qquad (3.3.19a)$$

$$D\Phi_{12} - \delta(\Phi_{11}-3\Lambda) - \bar{\delta}\Phi_{02} + \Delta\Phi_{01}$$
$$= \bar{\nu}\Phi_{00} + (2\gamma-\mu-2\bar{\mu})\Phi_{01} - \bar{\lambda}\Phi_{10} + 2(\bar{\pi}-\tau)\Phi_{11} +$$
$$+ (\pi+\bar{\beta}-3\alpha)\Phi_{02} + (3\rho+2\varepsilon)\Phi_{12} + \sigma\Phi_{21} \qquad (3.3.19b)$$

$$D\Phi_{22} - \delta\Phi_{21} - \bar{\delta}\Phi_{12} + \Delta(\Phi_{11}+3\Lambda)$$
$$= \nu\Phi_{01} + \bar{\nu}\Phi_{10} - 2(\mu+\bar{\mu})\Phi_{11} - \lambda\Phi_{02} - \bar{\lambda}\Phi_{20} +$$
$$+ (2\pi-\bar{\tau}+2\bar{\beta})\Phi_{12} + (2\bar{\pi}-\tau+2\beta)\bar{\Phi}_{12} + 2\rho\Phi_{22} \qquad (3.3.19c)$$

The remaining tetrad freedom that preserves all the relations introduced so far is ([61]),

(i) *a null rotation about* $\underline{k}$*:*

$$k^{a'} = k^a, \quad n^{a'} = n^a + B\bar{B}\,k^a + \bar{B}\,m^a + B\,\bar{m}^a,$$
$$m^{a'} = m^a + B\,k^a, \quad (B \in C) \qquad (3.3.20a)$$

(ii) *a spatial rotation in the* $(\underline{m}\text{-}\underline{\bar{m}})$*-plane:*

$$k^{a'} = k^a,\ n^{a'} = n^a, \quad m^{a'} = e^{iC}m^a, \quad (C \in R) \qquad (3.3.20b)$$

Equations (3.3.12-19) may be simplified in a number of ways by further specialisation of the null tetrad (3.3.9). In §5.7 (and only in this section) we shall use the tetrad freedom to set $X^A = 0$. The resulting tetrad will then, in general, not be parallelly propagated along the null generators. *In all other parts of this thesis we shall demand that the null tetrad be parallelly propagated along the* *integral curves of $\underline{k}$, so that*

$$\nabla_{\underline{k}}\underline{e}_a = 0. \tag{3.3.21}$$

(For a = 1, this is already satisfied, since $\underline{k}$ is geodesic and v affine.) The spin coefficients then satisfy the stronger conditions

$$\kappa = 0,\ \varepsilon = 0,\ \pi = 0,\ \rho - \bar{\rho} = 0,\ \tau = \bar{\alpha} + \beta \tag{3.3.22}$$

while the remaining tetrad freedom is then given by (3.3.20), but with B and C now independent of v, i.e.

$$DB = 0,\ \ DC = 0. \tag{3.3.23}$$

The limiting behaviour of the spin coefficients and the metric tensor can now be derived.

§3.4. Central conditions

§3.4.1.

The fact that C is a regular world line implies that all tetrad components of the curvature tensor must be bounded as functions of v in the limit $v \to 0$, i.e.

$$\Phi_{IJ} = o(1),\ \ \Lambda = o(1),\ \ \Psi_Q = o(1), \tag{3.4.1}$$

$$(I,J = 0,1,2;\ Q = 0,1,\dots,4)$$

where the order symbol is defined by

$$f(w,v,x^a) = o[g(v)] \quad \text{iff} \quad |f(w,v,x^a)| < g(v)F(w,x^a) \tag{3.4.2}$$

for some $F(w,x^a)$ and for all small v.

For suppose that (3.4.1) does not hold. The remaining tetrad freedom is given by (3.3.20) subject to (3.3.23). From the transformation laws for the Weyl and Ricci tensor components listed in Appendix I, it follows that (3.4.1) will then be violated in any parallelly propagated null tetrad, and hence it will be possible to construct an orthonormal frame from $\{\underline{e}_a\}$ in which at least one of the curvature tensor components diverges, contrary to the assumption that C is regular.

Of course, this argument only works because $\underline{k}$ has been determined uniquely, and thus "boost" transformations of the form $k^{a'} = sk^a$, $n^{a'} = s^{-1}n^a$ are excluded from the group of permissible tetrad transformations.

Since C is geodesic, with tangent vector $\left.\frac{\partial}{\partial w}\right|_C$,

$$\nabla_{\left.\frac{\partial}{\partial w}\right|_C}\left\{\left.\frac{\partial}{\partial w}\right|_C\right\} = 0. \tag{3.4.3}$$

The next assumption required derives from the fact that in Minkowski space, a null cone (as opposed to any other kind of null hypersurface) is characterised by $\rho = -\frac{1}{v}$, $\sigma = 0$. Now Posadas has shown ([67]) that the only solutions of (3.3.15a,b) subject to (3.4.1):

$$D\rho = \rho^2 + \sigma\bar{\sigma} + o(1) \tag{3.4.4a}$$

$$D\sigma = 2\rho\sigma + o(1) \tag{3.4.4b}$$

which have the limiting behaviour $\rho \to -\frac{1}{v}$, $\sigma \to 0$, are given by

$$\rho = -\frac{1}{v} + o(v) \tag{3.4.5a}$$

$$\sigma = o(v). \tag{3.4.5b}$$

(The other solutions of (3.4.4) correspond to what are termed "plane", "cylindrical" and "shearing cylindrical" rays.) Since the surfaces {w = constant} are light cones of points on C, and C is regular, it follows that (3.4.5) must hold.

Finally, we shall assume that as many angular derivatives of (3.4.1,5) as may be required in the subsequent analysis do not spoil the limiting behaviour given by (3.4.1,5), i.e.

$$\Phi_{IJ,A_1A_2\ldots A_n} = o(1); \quad \rho_{,B_1B_2\ldots B_n} = o(v)$$

$$(A_j, B_j = 2,3) \qquad (3.4.6)$$

with similar expressions for the angular derivatives of Λ, Ψ_Q and σ.

§3.4.2.

Lemma 3.1. (Newman and Penrose,[55]).

Let the complex $(n\times n)$ matrix $\hat{G}$ and the complex column n-vector $\hat{h}$ be given functions of s, with

$$\hat{G} = \hat{o}(s^{-2}), \quad \hat{h} = \hat{o}(s^{-2}) \qquad (3.4.7)$$

where

$$f(w,s,x^a) = \hat{o}[g(s)] \text{ iff } |f(w,s,x^a)| < g(s)F(w,x^a) \qquad (3.4.8)$$

for some $F(w,x^a)$ and all large s.

Let the $(n\times n)$ matrix $\hat{H}$ be independent of s and have no eigenvalue with positive real part. Suppose further that any eigenvalue of $\hat{H}$ with vanishing real part is regular, i.e. its multiplicity is equal to the number of linearly independent eigenvectors corresponding to it. Then all solutions of

$$\frac{du}{ds} = (\hat{H}s^{-1} + \hat{G})u + \hat{h} \qquad (3.4.9)$$

are bounded functions of s as $s\to\infty$, i.e. $u = \hat{o}(1)$.

Lemma 3.2.

Let the complex $(n\times n)$ matrix G and the complex column n-vector h be given functions of v, with

$$G = o(1), \quad h = o(v^{-1}). \qquad (3.4.10)$$

Then all solutions of

$$Dx = (-Iv^{-1} + G)x + h \qquad (3.4.11)$$

satisfy $x = o(v^{-1})$. Here I denotes the identity (n×n) matrix.

Proof.

Let $s = v^{-1}$. Then $Dx = -s^2 \frac{dx}{ds}$, and $G = \hat{o}(1)$, $h = \hat{o}(s)$. Thus

$$\frac{dx}{ds} = (-I s^{-1} - G s^{-2})x - h s^{-2}.$$

Now let $x = us$. Then

$$\frac{du}{ds} = -G s^{-2} u - h s^{-3}.$$

Lemma 3.1, with $\hat{G} = -G s^{-2} = \hat{o}(s^{-2})$, $\hat{h} = -h s^{-3} = \hat{o}(s^{-3})$ and $\hat{H} = 0$ thus gives $u = \hat{o}(1)$, so that $x = \hat{o}(s)$, or $x = o(v^{-1})$. □

§3.4.3.

It is now easy to show that ([67])

$$g_{AB} = \overset{0}{g}_{AB} v^2 + o(v^4), \qquad \overset{0}{g}_{AB} = \overset{0}{g}_{AB}(w, x^a), \tag{3.4.12}$$

since, using (3.4.5), we may rewrite (3.3.13a) as

$$D\xi^A = [-v^{-1} + o(v)]\xi^A + [o(v)]\bar{\xi}^A. \tag{3.4.13a}$$

Lemma (3.2) now applies, with

$$x = \begin{pmatrix} \xi^A \\ \bar{\xi}^A \end{pmatrix},$$

and hence $\xi^A = o(v^{-1})$. Substituting this result back into (3.4.13) gives

$$D\xi^A + v^{-1}\xi^A = o(1), \tag{3.4.13b}$$

or $\quad D(v\xi^A) = o(v). \qquad (3.4.13c)$

Since it is permissible to integrate order symbols (but not to differentiate them), it follows that

$$\xi^A = \xi^{AO} v^{-1} + o(v), \qquad \xi^{AO} = \xi^{AO}(w, x^a) \tag{3.3.14}$$

so that, by (3.3.10c),

$$g^{AB} = -(\xi^{AO}\bar{\xi}^{BO} + \bar{\xi}^{AO}\xi^{BO})v^{-2} + o(1) \tag{3.4.15a}$$

$$\equiv \overset{0}{g}{}^{AB} v^{-2} + o(1) \tag{3.4.15b}$$

If we therefore define $\overset{0}{g}_{AB}$ by

$$\overset{0}{g}_{AB}\overset{0}{g}{}^{BC} = \delta_A^C \text{ , i.e. } \overset{0}{g}_{AB}(\xi^{BO}\bar{\xi}^{CO} + \bar{\xi}^{BO}\xi^{CO}) = -\delta_A^C \text{ ,} \tag{3.4.16}$$

then (3.4.12) is an immediate consequence of (3.3.8e) and (3.4.15). Now (3.4.16) defines $\overset{0}{g}_{AB}$ iff det $\overset{0}{g}{}^{AB} \neq 0$. In order to ensure that this condition is satisfied (as it must be, since C is regular) the coordinate freedom (3.3.7) is used to set

$$\xi^{20} = i\xi^{30} = -P \text{ ,} \tag{3.4.17a}$$

where

$$P = P(w,\zeta,\bar{\zeta}) \neq 0, \quad \text{with} \quad \zeta \equiv x^2+ix^3. \tag{3.4.17b}$$

By using the tetrad freedom (3.3.20b), it is furthermore possible to make

$$P = \bar{P}. \tag{3.4.17c}$$

The coordinates (x^A) are now, in the limit $v \to 0$, conformally flat (or isothermal) coordinates:

$$\overset{0}{g}{}^{AB} = \text{diag } (-2P^2, -2P^2) \text{ ,} \tag{3.4.18a}$$

$$\overset{0}{g}_{AB} = \text{diag } (-\tfrac{1}{2}P^{-2}, -\tfrac{1}{2}P^{-2}) \text{,} \tag{3.4.18b}$$

and thus the 2-dimensional cross-sections of the past light cones, {w = constant, v = constant} have the line element

$$dq^2 \equiv -g_{AB}dx^A dx^B = (\tfrac{1}{2}P^{-2}d\zeta d\bar{\zeta})\, v^2 + o(v^4) \tag{3.4.18c}$$

The remaining tetrad freedom is (3.3.20a), while the remaining coordinate freedom is ([67]) a holomorphic transformation

$$\zeta' = \zeta'(w,\zeta). \tag{3.4.19}$$

Now consider equations (3.3.15d,e). Using (3.4.1,5), these may be written as

$$D\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}-v^{-1} & 0\\ 0 & -v^{-1}\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \begin{pmatrix}0 & o(v)\\ o(v) & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + o(1) \qquad (3.4.20)$$

Lemma 3.2 implies $\alpha = o(v^{-1})$, $\beta = o(v^{-1})$. Substituting these results back into (3.4.20), we find (cf.(3.4.13,14)):

$$\alpha = \alpha^0 v^{-1} + o(v), \qquad \alpha^0 = \alpha^0(w,x^a) \qquad (3.4.21a)$$

$$\beta = \beta^0 v^{-1} + o(v), \qquad \beta^0 = \beta^0(w,x^a) \qquad (3.4.21b)$$

and hence, by (3.3.22),

$$\tau = \tau^0 v^{-1} + o(v), \qquad \tau^0 = \bar{\alpha}^0 + \beta^0. \qquad (3.4.21c)$$

The remaining tetrad freedom is now used to set $\tau^0 = 0$, as follows. Under a null rotation (3.3.20a), the spin coefficients ρ and σ are left invariant (see Appendix I), while τ transforms as

$$\tau' = \tau + B\sigma + \bar{B}\rho = (\tau^0 - \bar{B})v^{-1} + o(v).$$

By choosing $\bar{B} = \tau^0$, we therefore have $(\tau')^0 = 0$. Hence (dropping primes)

$$\tau = o(v) \qquad (3.4.22a)$$

and

$$\bar{\alpha}^0 + \beta^0 = 0. \qquad (3.4.22b)$$

Equations (3.3.13b) and (3.3.15g,h) now read

$$D\omega = [-v^{-1} + o(v)]\omega + [o(v)]\bar{\omega} + o(v) \qquad (3.4.23a)$$

$$D\lambda = [-v^{-1} + o(v)]\lambda + [o(v)]\mu + o(1) \qquad (3.4.23b)$$

$$D\mu = [-v^{-1} + o(v)]\mu + [o(v)]\lambda + o(1) \qquad (3.4.23c)$$

and hence

$$\omega = \omega^0 v^{-1} + o(v), \qquad \omega^0 = \omega^0(w,x^a) \qquad (3.4.24a)$$

$$\lambda = \lambda^0 v^{-1} + o(v), \qquad \lambda^0 = \lambda^0(w,x^a) \qquad (3.4.24b)$$

$$\mu = \mu^0 v^{-1} + o(v), \qquad \mu^0 = \mu^0(w,x^a) \qquad (3.4.24c)$$

From (3.3.13c) and (3.3.13f,i) it now follows that DX^A, $D\gamma$ and $D\nu$ are all $o(1)$, and thus

$$X^A = X^{A0}(w,x^a) + o(v) \tag{3.4.24d}$$

$$\gamma = \gamma^0(w,x^a) + o(v) \tag{3.4.24e}$$

$$\nu = \nu^0(w,x^a) + o(v). \tag{3.4.24f}$$

Before integrating (3.3.13d), we first substitute the results obtained thus far into (3.3.16b) which, with the aid of (3.2.5), (3.3.9) and (3.4.1), may be written as

$$\omega D\rho + \xi^A \rho,_A - \bar{\omega} D\sigma - \bar{\xi}^A \sigma,_A = o(1). \tag{3.4.25}$$

But, by (3.4.1,4,5,6),

$$D\rho = [-v^{-1} + o(v)]^2 + o(1) = -v^{-2} + o(1),$$

$$D\sigma = o(1), \qquad \rho,_A = o(v), \qquad \sigma,_A = o(v). \tag{3.4.26}$$

Substituting this, together with (3.4.14,24a), into (3.4.25) then shows that $\omega^0 = 0$, so that, by (3.4.24a), $\omega = o(v)$. But now (3.4.23a) gives

$$D\omega + v^{-1}\omega = o(v),$$

which integrates to

$$\omega = o(v^2). \tag{3.4.27}$$

Hence, by (3.3.13d) and (3.4.24e)

$$U = U^0 - (\gamma^0 + \bar{\gamma}^0)v + o(v^2), \tag{3.4.28a}$$

where U^0 is fixed, using (3.3.10a), by the requirement that w measures proper time along C:

$$U^0 = -\frac{1}{2}. \tag{3.4.28b}$$

From (3.3.14a) we then find

$$(\mu^0 + \frac{1}{2})\, \xi^{A0} + \bar{\lambda}^0 \bar{\xi}^{A0} = 0, \tag{3.4.29a}$$

$$X^{20},_2 - iX^{20},_3 = X^{30},_3 + iX^{30},_2 = 0, \tag{3.4.29b}$$

$$2\bar{\gamma}^0 + X^{A0}(\log P),_A - (\log P),_0 = 0. \tag{3.4.29c}$$

Together with (3.4.17), equation (3.4.29a) implies

$$\mu^0 = -\frac{1}{2}, \quad \lambda^0 = 0, \qquad (3.4.30a)$$

while (3.4.29b) shows that $X^{20} + iX^{30}$ is holomorphic, and thus the coordinate freedom (3.4.19) can be used to set

$$X^{A0} = 0. \qquad (3.4.30b)$$

Now (3.4.29c) gives, since P is real,

$$\gamma^0 = \bar{\gamma}^0 = \frac{1}{2}(\log P)_{,0}. \qquad (3.4.30c)$$

The remaining coordinate freedom is then

$$\zeta' = \zeta'(\zeta). \qquad (3.4.31)$$

In the NU integration, γ^0 is set to zero by a coordinate transformation involving w - in our case no such freedom exists, because (see §3.3) w is uniquely determined. Instead, we shall deduce $\gamma^0 = 0$ from the condition that C is geodesic. The same argument also establishes that $\nu^0 = 0$.

Let $\{\underline{E}_a\} = \{\underline{E}_0, \underline{E}_i\}$ ($\underline{E}_0$ timelike, $\underline{E}_i$ spacelike, i=1,2,3) be an orthonormal tetrad that is well-defined on C. Clearly such a tetrad exists, since C is regular. Let $\underline{E}_0|_C$ be tangent to C:

$$(\underline{E}_0) = \left(\frac{\partial}{\partial w}\right), \qquad (3.4.32)$$

where we have introduced the notation (valid for the rest of this section) that round brackets denote the limit of a quantity on approaching C along a particular null generator $\{x^a\}$. Because both the null tetrad (3.3.9) and the null coordinate system used are singular on C, a quantity such as $(\gamma) = \gamma^0$ will, in general, be a function of the angular coordinates (x^a). In what follows, this functional dependence will not be indicated explicitly by writing $(\gamma)(x^a)$, it being understood that all such expressions are to be evaluated by taking the limit $v \to 0$ along a particular null generator. Similarly, $(\underline{k})$ denotes the tangent vector, on C, to some specific null geodesic.

Since C is geodesic, $\nabla_{(\underline{E}_o)}\{(\underline{E}_o)\} = 0$ (cf.(3.4.3)). In addition, we demand that $(\underline{E}_i)$ also be parallelly propagated along C. Thus

$$\nabla_{(\underline{E}_o)}\{(\underline{E}_a)\} = 0. \tag{3.4.33}$$

The null tetrad $\{\underline{e}_a\}$ is related to $\{\underline{E}_a\}$ by

$$\underline{e}_a = e_a{}^b \underline{E}_b\,, \quad \text{some } e_a{}^b. \tag{3.4.34}$$

Now (3.3.4c) and (3.4.32) imply $(e_1{}^o) = -1$, i.e.

$$(\underline{k}) = -(\underline{E}_o) + (k^i \underline{E}_i) \tag{3.4.35a}$$

where, since $\underline{k}$ is null, the 3-dimensional scalar product $k^i k_i$ satisfies

$$(k^i k_i) \equiv (e_1{}^i e_{1i}) = -1, \tag{3.4.35b}$$

indices being raised and lowered by the orthonormal tetrad metric, $g_{ab} = \text{diag}\,(1,-1,-1,-1)$.

The singularity of the null tetrad $\{e_a\}$ is here reflected by the fact that $\underline{k}$ does not tend to a well-defined limit as C is approached along different generators.

From (3.3.9) and the central limits derived thus far, it is possible to show that

$$(\underline{E}_o) = \left(\frac{\partial}{\partial w}\right) = -(\underline{n} + \tfrac{1}{2}\underline{k}) \tag{3.4.36a}$$

and hence

$$(\underline{n}) = -\tfrac{1}{2}(\underline{E}_o) - \tfrac{1}{2}(k^i \underline{E}_i). \tag{3.4.36b}$$

Then $\underline{k}.\underline{m} = 0 = \underline{n}.\underline{m}$ implies that $(e_2{}^o) = 0 = (e_3{}^o)$, and thus

$$(\underline{m}) = (m^i \underline{E}_i), \tag{3.4.37a}$$

where

$$(m^i m_i) = 0 = (m^i k_i), \quad (m^i \bar{m}_i) = -1. \tag{3.4.37b}$$

Using the definitions (3.2.3) of the spin coefficients, a straightforward calculation shows that $\nabla_{(\underline{E}_o)}\{(\underline{E}_o)\} = 0$ iff

$$(\gamma+\bar{\gamma})(k^i) + (\nu - \tfrac{1}{2}\bar{\tau})(m^i) - (\bar{\nu} - \tfrac{1}{2}\tau)(\bar{m}^i) = 0.$$

By (3.4.35b,37b), it follows that $(\gamma+\bar{\gamma})$ and $(\nu - \frac{1}{2}\bar{\tau})$ vanish. But, by (3.4.22a,24d,e), $(\tau) = 0$, $(\gamma+\bar{\gamma}) = \gamma^0 + \bar{\gamma}^0$ and $(\nu) = \nu^0$. Thus, by (3.4.30c),

$$\gamma^0 = 0 = \nu^0, \tag{3.4.38a}$$

$$P,_o = 0. \tag{3.4.38b}$$

Now (3.4.28) gives

$$U = -\frac{1}{2} + o(v^2). \tag{3.4.39}$$

The explicit form of P may now be derived from (3.3.14c), (3.3.16c) and (3.4.38b). From (3.3.14c) it is found that

$$\alpha^0 = -\frac{\partial P}{\partial \bar{\zeta}}, \tag{3.4.40}$$

so that (3.3.16c) implies

$$P\frac{\partial^2 P}{\partial \zeta \partial \bar{\zeta}} - \frac{\partial P}{\partial \zeta}\frac{\partial P}{\partial \bar{\zeta}} = \frac{1}{8}. \tag{3.4.41}$$

(In deriving these equations, many of the central limits deduced thus far have been used.)

The general 1-parameter family of solutions of (3.4.41) is

$$P = a(1 + b\zeta\bar{\zeta}), \quad a^2 b = \frac{1}{8}. \tag{3.4.42a}$$

where a and b are real constants (recall that, by (3.4.17b) and (3.4.38b), $P,_o = 0 = P,_1$). Using the coordinate freedom (3.4.31), we may set $b = 1$, in which case $a = \frac{1}{2\sqrt{2}}$. Thus

$$P = \frac{1}{2\sqrt{2}}(1 + \zeta\bar{\zeta}), \tag{3.4.42b}$$

and now (3.4.18b) implies

$$dq_o^2 \equiv \lim_{v\to 0}[-v^{-2}g_{AB}dx^A dx^B] = 4(1+\zeta\bar{\zeta})^{-1}d\zeta d\bar{\zeta} \tag{3.4.42c}$$

which is simply the line element of the unit 2-sphere expressed in terms of the stereographic coordinate

$$\zeta \equiv e^{i\phi}\cot\frac{\theta}{2}, \tag{3.4.43}$$

where (θ,ϕ) are the usual 2-sphere polar coordinates.

The remaining coordinate freedom is (3.4.31) subject to $\frac{d\zeta'}{d\zeta} = \pm 1$, i.e.

$$\zeta' = \pm\zeta + c \qquad (c \in C) \tag{3.4.44}$$

These are the so-called "rigid rotations" ([74]), which form a 3-parameter subgroup of the full 6-parameter group of conformal mappings of the unit 2-sphere onto itself.

Collecting results, the limiting behaviour of the spin coefficients (other than κ, ϵ and π, which satisfy (3.3.22)) is therefore

$$\sigma,\tau,\nu,\lambda,\gamma = o(v)$$

$$\rho = -v^{-1} + o(v), \quad \mu = -\frac{1}{2}v^{-1} + o(v),$$

$$\alpha = \alpha^0 v^{-1} + o(v), \quad \beta = \beta^0 v^{-1} + o(v) \tag{3.4.45a}$$

where

$$\bar{\beta}^0 = -\alpha^0 = \frac{\partial P}{\partial\bar{\zeta}} = \frac{1}{2\sqrt{2}}\zeta . \tag{3.4.45b}$$

The metric variables satisfy

$$U = -\frac{1}{2} + o(v^2), \quad \omega = o(v^2), \quad X^A = o(v),$$

$$\xi^A = \xi^{AO}v^{-1} + o(v), \tag{3.4.46a}$$

where

$$\xi^{2O} = i\xi^{3O} = -P = -\frac{1}{2\sqrt{2}}(1+\zeta\bar{\zeta}) \tag{3.4.46b}$$

The coordinate components of the metric tensor follow from (3.3.8,10) and (3.4.46):

$$g^{oa} = -\delta^a_1, \quad g^{11} = -1 + o(v^2), \quad g^{1A} = o(v),$$

$$g^{AB} = -2P^2\delta^{AB}v^{-2} + o(1) \tag{3.4.47a}$$

$$g_{1a} = -\delta^o_a, \quad g_{oo} = 1 + o(v^2), \quad g_{oA} = o(v^3),$$

$$g_{AB} = -\frac{1}{2}P^{-2}\delta_{AB}v^2 + o(v^4). \tag{3.4.47b}$$

§3.4.4.

It remains to be shown that the coordinates (x^A) are now observational coordinates in the sense defined in §3.3.2. The orthonormal tetrad $\{\underline{E}_a\}$ introduced in §3.4.3 is completely specified on C once a spacelike triad $\{\underline{E}_i\}$ has been chosen at one point on C. There remains, however, the freedom to choose how this tetrad is propagated off C, and hence we may set

$$\nabla_{(\underline{E}_i)}\{(\underline{E}_a)\} = 0. \tag{3.4.48}$$

The projection $_\perp\underline{k}$ of the null vector $\underline{k}$ into the instantaneous rest space of an observer with world line C is, by (3.4.35a)

$$_\perp\underline{k} = (\underline{k}) + (\underline{E}_o) = (k^i\underline{E}_i). \tag{3.4.49}$$

Thus (k^i) are simply the direction cosines of $_\perp\underline{k}$ with respect to the parallelly propagated orthonormal triad $\{\underline{E}_i\}$. But from (3.4.36a) and the central limits (3.4.45), it is easy to show that

$$\nabla_{(\underline{E}_o)}\{(\underline{k})\} = 0, \tag{3.4.50a}$$

so that, by (3.4.33),

$$0 = \nabla_{(\underline{E}_o)}\{(k^i)\} = (k^i)_{,o} \tag{3.4.50b}$$

Thus the same direction cosines are assigned to the same null directions (in the non-rotating frame $\{\underset{-}{E}_a\}$). It does not yet follow that the same null directions are assigned the same coordinate labels x^A. In order to show that this is indeed the case, we shall prove that

$$(k^i) = [\sin\theta\sin\phi, \sin\theta\cos\phi, \cos\theta] \qquad (3.4.51)$$

where (θ,ϕ) are defined by (3.4.43). The required result then follows immediately.

In deriving (3.4.51), it is convenient to use the differential operator $\eth$ and the associated class of functions, the so-called spin-s spherical harmonics ([74],[75]). For a 2-sphere with line element

$$dq_o^2 = \frac{1}{2} P^{-2} d\zeta d\bar{\zeta} \qquad (3.4.52)$$

we define

$$\eth\eta = 2\sqrt{2}P^{1-s} \frac{\partial}{\partial\zeta} (P^s\eta) \qquad (3.4.53a)$$

$$\bar{\eth}\eta = 2\sqrt{2}P^{1+s} \frac{\partial}{\partial\bar{\zeta}} (P^{-s}\eta) \qquad (3.4.53b)$$

for any quantity η of spin weight s, i.e. if under a transformation of the form

$$\tilde{m}^{a'} = \tilde{m}^a e^{iC}, \qquad (C \in R) \qquad (3.4.54a)$$

η transforms as

$$\eta' = \quad e^{isC}, \qquad (3.4.54b)$$

where $Re(\underset{-}{\tilde{m}})$ and $Im(\underset{-}{\tilde{m}})$ are two real vector fields tangent to the 2-sphere, and $\underset{-}{\tilde{m}}.\underset{-}{\tilde{m}} = 0$, $\underset{-}{\tilde{m}}.\underset{-}{\bar{\tilde{m}}} = -1$. Now $Re(\underset{-}{m})$ and $Im(\underset{-}{m})$ are tangent to the 2-surfaces {w = constant, v = constant}; (3.4.42) shows that, in the limit $v \to 0$, these 2-surfaces are conformal to the unit 2-sphere, with conformal factor $- v^{-2}$. Thus, for any quantity η which is independent of v we may replace $\underset{-}{\tilde{m}}$ in the definition (3.4.54) of s by $\underset{-}{m}$. It then follows from (3.4.45,46) that

$$\eth\eta = -\sqrt{2}(\xi^{AO}\eta,_A + 2s\bar{\alpha}^0\eta) \tag{3.4.55a}$$

$$\bar{\eth}\eta = -\sqrt{2}(\bar{\xi}^{AO}\eta,_A - 2s\alpha^0\eta). \tag{3.4.55b}$$

The spin-s spherical harmonics ${}_sY_{\ell m}$ ([74]) are defined by

$${}_sY_{\ell m} = a_{\ell m}[(\ell-s)!(\ell+s)!]^{-1/2}(1+\zeta\bar{\zeta})^{-\ell}\times$$

$$\times\sum_p \binom{\ell-s}{p}\binom{\ell+s}{p+s-m}\zeta^p(-\bar{\zeta})^{p+s-m} \tag{3.4.56a}$$

with

$$a_{\ell m} = (-1)^{\ell-m}[(\ell+m)!(\ell-m)!(2\ell+1)/4\pi]^{1/2}$$

where

$$|s| \le \ell, \quad |m| \le \ell,$$

and satisfy

$${}_s\bar{Y}_{\ell m} = (-1)^{m+s}\,{}_{-s}Y_{\ell m} \qquad \text{(no summation implied)} \tag{3.4.56b}$$

$$\eth\,{}_sY_{\ell m} = [(\ell-s)(\ell+s+1)]^{1/2}\,{}_{s+1}Y_{\ell m} \tag{3.4.56c}$$

$$\bar{\eth}\,{}_sY_{\ell m} = -[(\ell+s)(\ell-s+1)]^{1/2}\,{}_{s-1}Y_{\ell m} \tag{3.4.56d}$$

$$\bar{\eth}\eth\,{}_sY_{\ell m} = -(\ell-s)(\ell+s+1)\,{}_sY_{\ell m}\,. \tag{3.4.56e}$$

They form a complete orthonormal set of functions for each value of s; i.e. any function which is regular on the 2-sphere and which has spin weight s can be expanded in a series in ${}_sY_{\ell m}$.

Since, by (3.2.3),

$$\nabla_{\underline{m}}\underline{m} = -\bar{\sigma}\underline{n} + \bar{\lambda}\underline{k} + (\beta-\bar{\alpha})\underline{m} \tag{3.4.57a}$$

$$\nabla_{\underline{m}}\underline{k} = \tau\underline{k} - \rho\underline{m} - \sigma\bar{\underline{m}}\,, \tag{3.4.57b}$$

it follows from (3.3.9,21), (3.4.35-37,45,46,48,53-55) that

$$\eth(m^j) = 0, \tag{3.4.58a}$$

$$\eth(k^j) = -\sqrt{2}(m^j). \tag{3.4.58b}$$

Moreover, similar to (3.4.50), we have $(m^j)_{,0} = 0$. Thus, by (3.4.56c,58a)

$$(m^j) = a^{jp}\,{}_1Y_{1p} \qquad (p = -1.0,1)$$

(the summation convention holds), where a^{jp} are complex constants determined by the conditions (3.4.37b). These imply

$$(m^j) = \frac{1}{\sqrt{2}}(1+\zeta\bar{\zeta})^{-1}[-i(1+\bar{\zeta}^2),\ 1-\bar{\zeta}^2,\ 2\bar{\zeta}] \tag{3.4.59a}$$

and now (3.4.58b) may be solved, with the constants of integration being determined by (3.4.35b,37b). We find

$$(k^j) = (1+\zeta\bar{\zeta})^{-1}[i(\bar{\zeta}-\zeta),\ \zeta+\bar{\zeta},\ \zeta\bar{\zeta}-1] \tag{3.4.59b}$$

and now, finally, (3.4.51) is obtained by substituting for ζ from (3.4.43) into (3.4.59b).

Since (3.4.58) is crucial to this analysis, it is worth giving an alternative derivation of this equation. We use the so-called GHP-formalism developed in [76].

Let $\{\underline{n},\underline{k},\underline{m},\underline{\bar{m}}\}$ be a null tetrad satisfying (3.2.1). Consider the transformation

$$k^{a'} = \Omega\bar{\Omega}k^a\ ,\quad n^{a'} = \Omega^{-1}(\bar{\Omega})^{-1}n^a,\quad m^{a'} = \Omega(\bar{\Omega})^{-1}m^a \tag{3.4.60a}$$

where

$$\Omega^2 \equiv fe^{iC} \quad (\text{some } f,C \in R). \tag{3.4.60b}$$

If, under a transformation of the form (3.4.60a), a quantity η transforms as

$$\eta' = \eta^p(\bar{\eta})^q \tag{3.4.61}$$

then η is said to be a spin-and-boost-weighted scalar of type $\{p,q\}$, with spin-weight $\frac{1}{2}(p-q)$ and boost-weight $\frac{1}{2}(p+q)$. (Note that this is consistent with the previous definition of spin-weight, since (3.4.54) is a special case of (3.4.60, 61), obtained by setting $f=1$.)

Define an operator $\eth_*$ by

$$\eth_*\eta = (\delta - p\beta - q\bar{\alpha})\eta, \tag{3.4.62}$$

where δ is given by (3.2.5) and (3.3.9), and α, β are the spin coefficients defined by (3.2.3).

Now suppose $\eta = o(1)$. Then

$$\eth_*\eta = v^{-1}[\xi^{AO}\eta,_A + (p-q)\bar{\alpha}^0\eta] + o(v) \tag{3.4.63}$$

$$\eth_*\eta = v^{-1}[\xi^{AO}\eta,_A + 2s\bar{\alpha}^0\eta] + o(v),$$

since $s = \frac{1}{2}(p-q)$.

Hence, by (3.4.55),

$$\eth_*\eta = -\frac{1}{\sqrt{2}}v^{-1}\eth\eta + o(v). \tag{3.4.64}$$

Let $\{o^A, \iota^A\}$ be a spin frame associated with the null tetrad $\{\underline{n},\underline{k},\underline{m},\bar{\underline{m}}\}$, so that

$$k^a = o^A\bar{o}^{A'}, \quad n^a = \iota^A\bar{\iota}^{A'}, \quad m^a = o^A\bar{\iota}^{A'},$$

$$\bar{m}^a = \iota^A\bar{o}^{A'}. \tag{3.4.65}$$

Then ([76])

$$\eth_* o^A = -\sigma\iota^A, \quad \eth_*\bar{o}^{A'} = -\rho\bar{\iota}^{A'}$$

$$\eth_*\iota^A = \mu o^A, \quad \eth_*\bar{\iota}^{A'} = \bar{\lambda}\bar{o}^{A'} \tag{3.4.66}$$

and hence

$$\eth_* m^a = -\sigma n^a + \bar{\lambda}k^a \tag{3.4.67a}$$

$$\eth_* k^a = -\sigma\bar{m}^a - \rho m^a \tag{3.4.67b}$$

(cf. (3.4.57)). Hence by (3.4.64) and the central limits

(3.4.45),

$$v^{-1}\eth m^a = \sqrt{2}[o(v)n^a + o(v)k^a]$$

$$v^{-1}\eth k^a = \sqrt{2}[o(v)\bar{m}^a + o(v)m^a] - \sqrt{2}v^{-1}m^a \;,$$

from which (3.4.58) follows immediately.

§3.4.5.

The central behaviour of the fluid 4-velocity components can now be calculated. Using the central limits (3.4.46), it is readily shown that (3.3.9) implies

$$\frac{\partial}{\partial w} = -\underline{n} + [-\tfrac{1}{2} + o(v^2)]\underline{k} + o(v^2)\underline{m} + o(v^2)\underline{\bar{m}} \tag{3.4.68a}$$

$$\frac{\partial}{\partial v} = \underline{k} \tag{3.4.68b}$$

$$\frac{\partial}{\partial x^2} = o(v^3)\underline{k} - \frac{1}{2P}v\,[\,1+o(v)][\,\underline{m}+\underline{\bar{m}}] \tag{3.4.68c}$$

$$\frac{\partial}{\partial x^3} = o(v^3)\underline{k} - \frac{i}{2P}v\,[\,1+o(v)][\,\underline{m}-\underline{\bar{m}}] \;, \tag{3.4.68d}$$

from which it follows that

$$\begin{aligned}\underline{u} = -u^o\underline{n} &+ \{[\,-\tfrac{1}{2}+o(v^2)\,]u^o + u^1 + o(v^3)u^2 + o(v^3)u^3\}\underline{k} + \\ &+ \{o(v^2)u^o - \frac{1}{2P}v[1+o(v)\,][u^2+iu^3]\}\underline{m} \\ &+ \{o(v^2)u^o - \frac{1}{2P}v[1+o(v)\,][u^2-iu^3]\}\underline{\bar{m}}\;.\end{aligned} \tag{3.4.69}$$

But, by (3.3.1) and (3.4.36a)

$$\underline{u}|_C = \frac{\partial}{\partial w}\Big|_C = -(\underline{n}+\tfrac{1}{2}\underline{k})_C = E_o|_C\;. \tag{3.4.70}$$

Comparison of (3.4.69) and (3.4.70) gives

$$\lim_{v\to 0} u^o = 1 \tag{3.4.71a}$$

$$\lim_{v\to 0} u^1 = 0 \tag{3.4.71b}$$

$$\lim_{v\to 0} u^A = u^{AO}(w,x^B)\;, \tag{3.4.71c}$$

where $u^{AO}(w,x^B)$ are arbitrary functions. (The physical significance of these functions is discussed in §5.5.)

Finally, the total energy density, denoted M from now on in order to distinguish it from the spin coefficient μ, as well as the pressure p must both tend to well-defined, finite limits (since they are scalars, and C is regular), and thus

$$\lim_{v\to 0} M = M^0(w) \tag{3.4.72a}$$

$$\lim_{v\to 0} p = p^0(w). \tag{3.4.72b}$$

4. COSMOLOGICALLY SIGNIFICANT OBSERVATIONS

§4.1. Introduction.

In order to study the cosmological characteristic initial value problem, it is necessary to investigate what cosmologically significant information can be derived from observations. Since the most important observations relevant to cosmology are those of distant galaxies and of the background radiation, we shall be concerned almost exclusively with these. Three other kinds of observation (discussed in more detail in [2]), however, also deserve to be mentioned.

(i) Local physical experiments are important in so far as they provide the theoretical framework in terms of which all observations are interpreted (see chapter 1). Whether detailed cosmological information could ever be extracted from the results of purely local experiments, is uncertain (but see, e.g., [77]).

(ii) Local astronomical observations provide knowledge of the local ordering of matter, and one of the basic assumptions made in chapter 1 is that this ordering holds also on cosmological scales. Together with geophysical data, local astronomical observations provide important information about, e.g., light element abundances. Such information needs to be incorporated in any detailed cosmological theory.

(iii) Cosmic rays present a slightly different problem: if one could prove that they were extragalactic in origin and find ways of "separating out" the interactions they have undergone, observations of these massive high energy particles could well yield significant information about distant regions of the universe.

For our present purposes, therefore, these observations will be regarded as imposing constraints that any viable cosmological model must satisfy; they do not provide information that can be used as initial data for the initial value problem under consideration. Only the observations mentioned in the first paragraph will therefore be discussed in detail.

A common feature of all observations that can be made of distant galaxies and of the background radiation is that the information reaches the observer on the central world line C by means of electromagnetic signals, so that these observations give information about the past light cones of points on C.

Each single observation is made over a finite time interval, and thus involves a 1-parameter family of past light cones, rather than a single light cone. But in a characteristic initial value problem, initial data needs to be specified on just one past light cone $C^-(p)$ (we are assuming, for the moment, that the solution is required only in the interior of $C^-(p)$). This problem can be overcome by appealing to the fluid approximation already assumed and to the scale of time variation of cosmological quantities, in order to construct an *idealised initial past light cone* $C^-(p)$: all the available cosmological information obtained from actual observations made at different times and in different directions is simply assumed to have been obtained by observing from a single point, p, on C, and is used as initial data on the single past light cone $C^-(p)$. This is clearly a "smoothing approximation" of sorts, similar to to that used in conventional cosmology when one refers to spatially homogeneous hypersurfaces in the universe. The justification for this approximation is the fact that the scale of time variation of cosmological quantities is expected to be extremely large compared to the time intervals over which observations are made on C. In fact, it can be argued that such an approximation is necessary for reasons of consistency once the fluid approximation has been adopted: the averaging scale involved in the fluid approximation is that of clusters of galaxies, while the entire recorded history of mankind corresponds to a spatial distance of a few kiloparsecs.

Now, the structure of the *actual* past light cones of points on C will, in general, be very complicated; in particular, caustics are bound to occur in the vicinity of massive stars and black holes. The existence of such caustics might cause a number of problems that need to be considered carefully when the actual observations are interpreted. (One interesting point, raised by Penrose ([71]), is whether there are any observational effects, such as anomalous brightness of certain sources, that can be used to detect such caustics.)

But as far as the *idealised* past light cone $C^-(p)$ is concerned, we may assume that it does not develop such caustics, precisely because the fluid approximation (as formulated in §1.3) already includes the assumption that local irregularities in the matter distribution have been smooothed out - and we have interpreted this as implying that, for the purposes of the characteristic initial value problem, black holes may be ignored. (Strictly speaking, by the very nature of singularities, any such "hole-free" assumption needs to be formulated in global terms; see, e.g., [27]).

Caustics caused by the global topological properties of the universe cannot be smoothed out in this way. A good example is provided by the spherically symmetric, static universe with 2 regular centres (see Appendix III). The conformal structure of such a universe may be represented in the usual way (see, e.g., [25]) by a cylinder with topology $R \times S^3$ (see Appendix III). The past light cone of an event p on C wraps around the cylinder and refocusses at a second point p' to the past of, and conjugate to, p. Thus, if the universe has compact space sections and its age is large enough so that we have already seen around it, the observer on C will see different images, in different directions, of the same object.

It was noted in chapter 3 that the null coordinate system (w,v,x^A) breaks down at this point and one needs to determine what identifications must be made amongst the observed objects. In principle, this could be done, given a sufficiently detailed catalogue of source characteristics. One could then define a group G of identifications that states, for each object observed on $C^-(p)$ in the direction x^A and at redshift z, which other objects seen in different directions x^A_α and at different redshifts z_α, are in fact the same object.

In practice, such a procedure would be exceedingly difficult to carry out, mainly because a sufficiently detailed catalogue of source characterisitics is not likely to become available in the forseeable future. In order to circumvent this difficulty we shall, for the remainder of this thesis, restrict attention to that part of the past light cone $C^-(p)$ on which the null generators have not yet started to refocus; this means that the spin coefficient ρ satisfies

$$\rho < 0 \tag{4.1.1}$$

throughout.

§4.2. Observations of distant galaxies

§4.2.1. Redshift

If light emitted at a wavelenth λ_e by some distant galaxy is observed on C to have wavelength λ_o, the redshift of the source is defined to be

$$z \equiv \frac{\lambda_o}{\lambda_e} - 1. \tag{4.2.1}$$

Since w measures proper time t along C, the time dilation observed from C is determined by $\frac{dw}{dt}$, so that

$$1+z = \frac{dw}{dt} = u^o. \tag{4.2.2}$$

This expression can also be derived using the relation ([15])

$$1+z = (u^a k_a)/(u^a k_a)_o \tag{4.2.3a}$$

which implies, by (3.3.3,4),

$$1+z = -\,(\,u^a k_a\,) = u^a \delta^o_a = u^o. \tag{4.2.3b}$$

Thus the 4-velocity component u^o is a directly observable quantity. In fact, the components u^A are also observable (in principle), and may be determined by measurements of the proper motions.

§4.2.2. Proper motions

The celestial sphere of an observer at p on C may be defined as the set of all unit spacelike vectors in the instantaneous rest space of the observer, i.e.

$$S_p = \{\underline{X} \in T_pM \mid \underline{X}.\underline{u} = 0,\ \underline{X}.\underline{X} = -1\}$$

But from the discussion in §3.4.4, it follows that S_p is simply the set of projected null directions at p, i.e.

$$S_p = \{{}_\perp k(x^A) \mid x^A \in R\},$$

where ${}_\perp k$ is defined by (3.4.49).

Now consider a source on $C^-(p)$ which is assigned angular coordinates (x^A_p), i.e. a source joined to p by a null geodesic whose tangent vector has spatial projection ${}_\perp k(x^A_p)$. In general, these coordinates will vary with time on C; at some later time , corresponding to the event p' on C, the same source will be assigned different angular coordinates $(x^A_{p'})$. Suppose p and p' correspond to proper times w_o and w, respectively, on C. The proper motion of the source at time w_o is defined to be

$$M^A \equiv \lim_{w \to w_o} \left(\frac{x^A_{p'} - x^A_p}{w - w_o} \right) \tag{4.2.4}$$

or

$$M^A \equiv \left. \frac{dx^A}{dw} \right|_{w=w_o} . \tag{4.2.5}$$

Thus M^A describes the (rate of) change in position of a galaxy on the celestial sphere as observed from C, relative to a local non-rotating reference frame. Of course, this definition only makes sense because x^A are *observational* coordinates, as shown in §3.4.4. An alternative definition of proper motions, used in [54], is obtained by considering $\frac{dk^i}{dw}$, where k^i are defined by (3.4.35). These two definitions are clearly equivalent and may be related by (3.4.59b); for our purposes, (4.2.5) is more convenient.

In order to relate the proper motions M^A to the 4-velocity components u^A, we use (3.3.3) and (4.2.2) to find

$$u^A = (1+z)^{-1}M^A. \tag{4.2.6}$$

Thus u^A are directly observable quantities.

The remaining 4-velocity component, u^1, is not observable, but has to be found from the normalisation condition $u^a u_a = 1$:

$$(1+z)^2 g_{00} - 2(1+z)u^1 + 2(1+z)g_{0A}u^A + u^A u^B g_{AB} = 1 \tag{4.2.7}$$

(If the x^1 coordinate had been chosen to be co-moving with the fluid, it would have followed from (3.3.3) that $u^1 = 0$, in which case (4.2.7) could have been used to determine one of the metric components once the others were known.)

From the central limits (3.4.71c) it follows that the proper motions do not necessarily vanish on C. This reflects the fact that the fluid shear and vorticity may be non-zero on C (see (5.4.17) and (5.5.11) below).

§4.2.3. *Observer area distance and distortion*

Both the shape and size of the image of an observed object depend on the path taken through space-time by the null geodesics which convey information from the source to the observer, because the space-time curvature introduces distortion (Weyl curvature) and causes focussing (Ricci curvature). The 2-dimensional line element at the source is given by

$$dl^2 = g_{AB} dx^A dx^B, \tag{4.2.8}$$

where dl^2 represents distances at the object perpendicular to the line of sight and dx^A are the corresponding angular displacements. If the orientation and intrinsic size and shape of distant objects are known from astrophysical considerations, dl^2 and dx^A are directly measurable. By carrying out measurements in all directions down to some limiting redshift $z^* = z^*(x^A)$, one can therefore determine (in principle) the metric components g_{AB} as functions of (z,x^A).

It is well-known (see, e.g., [69],[78]) that, with the choice of null tetrad (3.3.9), the focussing effect may be described in terms of the spin coefficient ρ, while the distortion of the null geodesics is represented, up to phase, by the null shear σ. It follows from (3.3.13a) that

$$\begin{aligned}\rho &= \tfrac{1}{2}(\xi^2\bar{\xi}^3 - \bar{\xi}^2\xi^3)^{-1} D(\xi^2\bar{\xi}^3 - \bar{\xi}^2\xi^3) \\ &= \tfrac{1}{4} D[\log(\det g^{AB})]\end{aligned} \tag{4.2.9a}$$

and

$$\sigma = (\xi^2\bar{\xi}^3 - \bar{\xi}^2\xi^3)^{-1}[\xi^2 D\xi^3 - \xi^3 D\xi^2]. \tag{4.2.9b}$$

Now since g_{AB} is measurable, (3.3.8e) implies that g^{AB} is measurable. Thus, by (3.3.10c), the following quantities are measurable (i.e. determined as functions of (z,x^A)):

$\xi^2\bar{\xi}^2$, $\xi^3\bar{\xi}^3$, $\xi^2\bar{\xi}^3 + \bar{\xi}^2\xi^3$.

Hence, if we set

$$\xi^A = r_A e^{i\chi_A}, \quad \text{(no sum)}, \tag{4.2.10}$$

then r_A and $\cos(\chi_2 - \chi_3)$ are measurable.

But, by (4.2.9),

$$\rho = \frac{1}{2} D[\log(r_2 r_3)] + \frac{1}{2}\cot(\chi_2-\chi_3)D(\chi_2-\chi_3) \tag{4.2.11a}$$

$$\sigma = -\frac{i}{2} e^{i(\chi_2+\chi_3)} [D\log(\frac{r_3}{r_2}) + iD(\chi_3-\chi_2)]\operatorname{cosec}(\chi_2-\chi_3). \tag{4.2.11b}$$

Hence $\rho \frac{dv}{dz}$ and $\sigma \frac{dv}{dz} e^{-i(\chi_2+\chi_3)}$ are measurable. The factor $\frac{dv}{dz}$ arises because there is no observational way of determining the relationship between v and z, while the phase factor $\exp[-i(\chi_2+\chi_3)]$ represents the arbitrariness in the choice of the tetrad vectors $\underline{m}$ and $\bar{\underline{m}}$, whose real and imaginary parts span the 2-dimensional screen space orthogonal to the null generators. If a parallelly propagated tetrad is used, so that, in addition to (3.3.11), conditions (3.3.22) also hold, this phase factor may be determined as follows: from (3.3.13a) it follows that

$$D(\xi^2\xi^3) = 2\rho\xi^2\xi^3 + \sigma(\xi^2\bar{\xi}^3 + \bar{\xi}^2\xi^3) \tag{4.2.12}$$

and hence

$$D(\chi_2 + \chi_3) = D\log(\frac{r_2}{r_3}) \tag{4.2.13}$$

which integrates to

$$\chi_2 + \chi_3 = \log(\frac{r_2}{r_3}) + H(w,x^A), \quad \text{some } H. \tag{4.2.14}$$

But, by (3.4.17),

$$\xi^2 = P\, e^{i\pi} v^{-1} + o(v) \tag{4.2.15a}$$

$$\xi^3 = P\, e^{i\pi/2} v^{-1} + o(v) \tag{4.2.15b}$$

and hence

$$\chi_2 + \chi_3 = \log(\frac{r_2}{r_3}) + (2n+1)\frac{\pi}{3} \quad (n \in \mathbb{N}) \tag{4.2.16}$$

Since r_A is measurable, it follows that $\exp[-i(\chi_2+\chi_3)]$ is also measurable. Thus measurement of g_{AB} implies that $\rho\frac{dv}{dz}$ and $\sigma\frac{dv}{dz}$ are measurable.

In practice, these quantities will be determined first, so that g_{AB} will then be deduced by reversing the argument above. But even $\rho\frac{dv}{dz}$ and $\sigma\frac{dv}{dz}$ are not the primary observable quantities; what one actually measures are the observer area distance and the integrated distortion (see below), from which $\rho\frac{dv}{dz}$ and $\sigma\frac{dv}{dz}$ can be derived by differentiation.

The observer area distance r may be defined by

$$\frac{d}{dv}(\log r^2) = k^a{}_{;a} = -2\rho, \qquad (4.2.17a)$$

$$r|_C = 0. \qquad (4.2.17b)$$

By (4.2.9a), the unique solution of (4.2.17a) subject to (4.2.17b) is given by

$$r^4 = F(w,x^A)(\det g_{AB}), \qquad (4.2.18a)$$

where the "constant" of integration $F(w,x^A)$ must be found by evaluating (4.2.18a) in the limit $v \to 0$.

The limiting behaviour of r is obtained by integrating (4.2.17a) using (3.4.5a). We find

$$r = v + o(v^3). \qquad (4.2.18b)$$

If the 2-metric is expressed in terms of stereographic coordinates $(\zeta,\bar{\zeta})$, this gives $F = 4P^4$, and thus

$$r^2 = 2P^2(\det g_{AB})^{1/2}. \qquad (4.2.19a)$$

A more familiar expression is obtained if standard angular coordinates $(x^{A'}) \equiv (\theta,\phi)$, related to x^A by (3.4.43), are used. In this case

$$r^2 \sin\theta = (\det g_{A'B'})^{1/2} \qquad (4.2.19b)$$

The observer area distance is usually defined by considering a bundle of null geodesics diverging from the observer. If this bundle subtends a solid angle $d\Omega$ at the observer and spans a cross-sectional area dS at some point along the rays, the observer area distance of that point is defined by the relation

$$dS = r^2 d\Omega. \tag{4.2.20}$$

The equivalence of this definition and (4.2.17) is a consequence of the the central conditions (3.4.47b), (4.2.19a) and the area law of ray optics (see, e.g.,[78]), which may be written as

$$\frac{d}{dv}(dS) = (k^a{}_{;a})(dS). \tag{4.2.21}$$

Now, not only is r measurable in principle (as follows from (4.2.17) and the fact that $\rho\frac{dv}{dz}$ is measurable), but there are in fact practical ways of measuring it. These are discussed in some detail in [38],[39]. We shall not enter here into a description of these quite complex observational techniques, except to mention that they depend crucially on a detailed knowledge of source characteristics such as surface brightness and radial evolution. It is customary to express the results of such measurements as the so-called (r,z) or "observer area distance relation". (The qualifier "observer" in this nomenclature is added to distinguish r from other area distances such as the "galaxy area distance", which is defined by considering a bundle of null geodesics diverging from the source (see, e.g., [15]). Since we shall not use such distance measures in this thesis, we shall henceforth refer to r simply as the "area distance".) Once $r = r(z,x^A)$ has been determined, $\rho\frac{dv}{dz}$ follows from (4.2.17).

Similarly, $\sigma\frac{dv}{dz}$ may be determined (although practical observational techniques do not yet seem to exist) by measuring the distortion present in the images of objects whose intrinsic shape and size are known. An intrinsically spherical object, for example, will appear elliptical in general, with the ratio of major to minor axes determined by $|\sigma|$ and the orientation of the ellipse (with respect to the preassigned 2-plane element basis vectors $Re(\underline{m})$ and $Im(\underline{m})$) determined by $\arg\sigma$ (see, e.g., [78]). What one is really measuring in this way is the integrated shear

$$\int_0^v \sigma\, dv = \int_0^z (\sigma \frac{dv}{dz})\, dz$$

rather than the shear itself; but by carrying out measurements for all values of z down a fixed null direction, one may deduce $\sigma\frac{dv}{dz}$ by differentiation, thereby obtaining what we shall refer to as the $(\sigma\frac{dv}{dz}, z)$ or "distortion-redshift relation".

§4.2.4. Number counts

Suppose the observer on C counts galaxies seen in a solid angle $d\Omega$ in the null direction x^A down to an affine distance v. An affine increment dv will include dN new galaxies in the count, where dN is the number of new sources *detected* in the proper volume (see [15])

$$dV = r^2 d\Omega (1+z) dv.$$

If the number density of sources at a distance v is n, then ndV new sources will be included in this volume, and thus

$$dN = fnr^2 d\Omega (1+z) (\frac{dv}{dz}) dz, \tag{4.2.22}$$

where f is a selection function (see [38]) representing the fraction of sources in the volume element dV that are actually detected. Thus the total number of galaxies counted within a solid angle down to a redshift z is given by

$$N(z) = \int_0^z (\frac{dN}{dz'})\, dz'$$

Provided that the selection function f is known (in practice one would have to estimate f from a knowledge of galactic brightness and spectra and the detection limits of the apparatus used), number counts of distant galaxies therefore determines $r^2 n \frac{dv}{dz}$ as a function of redshift and direction. Thus, since r may be determined by other means, $n \frac{dv}{dz}$ is measurable.

§4.3. Maximal Data Set

Given the results of ideal observations, i.e. observations carried out to indefinite accuracy in all directions, the discussion in the previous section shows that the following quantities are, in principle, measurable: $u^o, u^A, \rho\frac{dv}{dz}, \sigma\frac{dv}{dz}, n\frac{dv}{dz}$.

Because of limitations in the detection apparatus (and, possibly, because the universe becomes opaque at some era in the past) there will be limiting redshifts beyond which further observations cannot be made. These will not necessarily be the same for all measurable quantities, or even the same in all directions for any particular quantity, but in order to simplify the discussion somewhat, we shall assume that these limiting redshifts have the same value $z^*(x^A)$ for all observables (alternatively, z^* is just a suitably defined minimum limiting redshift).

In most calculations, it is convenient to work with the (r,z) rather than the $(\rho\frac{dv}{dz}, z)$ relation; we shall therefore take as our initial data the set

$$D(w_o,z^*) \equiv \{(u^o,u^A,r,\sigma\frac{dv}{dz},n\frac{dv}{dz}) \mid w=w_o,\ 0 \leq z \leq z^*,\ x^A \in R\}$$

where w_o is the proper time on C at which the observations are assumed to have been carried out.

It is conceivable that more information could be obtained from direct observation: in principle, there is no reason why one should not be able to measure the time derivatives of some of the quantities in $D(w_o,z^*)$. Thus, e.g., one might be able to measure time variations in the redshifts of certain sources. However, any attempt to include such derivatives in the initial data set faces the severe problem of having to decide when such time variations are of a cosmological rather than astrophysical nature. The variety of explanations put forward to explain the recently discovered variation in redshift of the radio source SS433, e.g., seems to indicate that it would be very difficult (or perhaps impossible) to establish conclusively that such time variations are indeed cosmological in origin. (The proper motions of galaxies are included in $D(w_o,z^*)$ because they do *not* seem to suffer from the same inherent ambiguity).

The set $D(w_o,z^*)$ will therefore be taken to represent the most detailed cosmological information one could hope to obtain by direct observation of distant galaxies, and will henceforth be referred to as the *maximal data set*. In the unlikely event that information not included in $D(w_o,z^*)$ does come to light, we shall use such information as a check of both the primary data and the field equations used.

§4.4. *Background radiation*

The term "background radiation" refers to radiation received from all unresolved sources down the line of sight, i.e. both to unresolved discrete sources and to continuous sources such as intergalactic gas or the primeval plasma.

For our purposes, the most important information to be gained by measurements of this radiation are its spectrum and the energy density $M_r|_p$ at the point p of observation. Observations appear to indicate that the background radiation (or at least the microwave component) has a well-defined (black body) temperature and is reasonably isotropic about p. If this is indeed the case, then the radiation density on $C^-(p)$ is given by ([15])

$$M_r = (M_r|_p)(1+z)^4 \tag{4.4.1}$$

In chapters 5 and 6 we shall make the assumption that, for dynamical purposes, the radiation contribution to the total stress-energy tensor is neglible - the galactic fluid will be regarded as a pressure-free perfect fluid. This assumption is almost certainly violated at early times. A slightly more realistic matter-energy description could be achieved by assuming that the stress-energy tensor T_{ab} consists of separately conserved matter and radiation components, represented by the same average 4-velocity u^a, with equations of state

$$p_{matter} = 0, \quad p_{radiation} = \frac{1}{3}M_{radiation}. \tag{4.4.2}$$

(For the remainder of the thesis we shall use the symbol "M" to denote the energy density, since the more conventional "μ" is being used to denote one of the spin coefficients.) If this description is used, the radiation density on $C^-(p)$ is an essential part of the required initial data for the field equations. In chapter 7 it will be shown how the discussion in chapters 5 and 6 can easily be modified to cope with this more complicated matter-energy description. For the present, however, the information provided by the background radiation will be relegated (like that obtained from, e.g., light element abundance measurements) to the role of indirect constraint data.

§4.5. *Observational space-times*

We shall not examine here the practical difficulties that arise when one actually attempts to measure any of the quantities defined in the previous sections (these are discussed at length in [1],[39]). For the purposes of observational cosmology - as conceived in this thesis - the major problem is that neither the distortion nor the proper motions are likely to be measurable in practice in the near future. Also, there is not much data currently available concerning the angular variation of the remaining elements of $D(w_o,z^*)$ (but see [80] and references cited there).

In order to implement the observational approach, one could proceed in a number of ways. Two possibilities are the following:

(i) All (or some) of those elements of $D(w_o,z^*)$ that cannot be determined by observation could simply be "carried" as unknown functions in the subsequent analysis. The extent to which the geometry and matter distribution are then undetermined would constitute a measure of "cosmological uncertainty" attributable to the paucity of currently available observations.

(ii) All (or some) of the unknown quantities could be *assumed* to have a particular functional dependence. This gives rise to various classes of *observational space-times: the observational space-time associated with a particular maximal data set* $D(w_o,z^*)$ *is the class of perfect fluid general relativistic cosmological models in which a null hypersurface* Σ *may be imbedded such that* $\Sigma \cup \{p\}$ *is the past light cone* $C^-(p)$ *of some point* p *on a regular geodesic world line C, and the observational relations predicted by this model are precisely those described by* $D(w_o,z^*)$.

Of course, the concept of an observational space-time is applicable irrespective of whether the elements of $D(w_o,z^*)$ are assumed or observed to have a particular functional form, and the most interesting cases are likely to arise when one of the following sets of *observational assumptions* is adopted:

	Area distance, Number counts:	Proper Motions, Distortion:
Ideal Observations (IO):	Measured	Measured
Restricted Ideal Observations (RIO):	Measured	Unknown
Simplified Ideal Observations (SIO):	Measured	Vanish
Isotropic Ideal Observations (IIO):	Isotropic	Vanish

As an example, consider the case of isotropic ideal observations. These are characterised by ([2])

$u^A = 0, \quad r = r(w_o,z), \quad \sigma = 0;$

in addition, the group G of identifications of observed objects (see §4.1) is invariant under rotations. Maartens has shown ([5]) that the only pressure-free perfect fluid solutions of Einstein's field equations consistent with these assumptions are spherically symmetric about the line C. Thus the observational space-time associated with IIO is just the Bondi-Tolman solution (see, e.g., [81]). This result will be used in chapter 6.

Suppose one is given a pressure-free perfect fluid model, with C and $p \in C$ as above. Suppose further that observational coordinates have been introduced as in §3.3. The coordinate system is then uniquely determined up to the freedom (3.3.44), corresponding to a rigid rotation about one point on C of the orthonormal triad used to define the angular coordinates. For any given v^*, the metric and stress-energy tensor of the model determine uniquely the quantities u^a, M and g_{ab} on $C^-(p,v^*)$, where $C^-(p,v^*)$ denotes the subset of $C^-(p)$ consisting of points lying within an affine distance v^* of p, i.e. they determine uniquely the set

$$S(w_o,v^*) \equiv \{(u^a,M,g_{ab}) \mid w=w_o,\ 0 \le v \le v^*,\ x^A \in R\}.$$

On the other hand, a necessary condition for a null hypersurface Σ to be imbeddable in the model so that $\Sigma \cup \{p\}$ is the past light cone of p, is that there be a one-one mapping of $C^-(p,v^*)$ into $\Sigma \cup \{p\}$ and that the elements of $S(w_o,v^*)$ take the same values at corresponding points of $\Sigma \cup \{p\}$ and $C^-(p,v^*)$. Any variation in the data in $S(w_o,v^*)$ (modulo the remaining coordinate freedom) will correspond to a real difference in the space-time geometry and matter distribution. We shall therefore consider $S(w_o,z^*)$ as characterising the local structure of the model on the light cone.

One way of constructing the observational space-times associated with a given set of observations $D(w_o,z^*)$ would therefore be to try and deduce the local structure $S(w_o,z^*)$ from the data $D(w_o,z^*)$, where now v^* is the affine distance corresponding to the limiting redshift z^*, and then to propagate the solution off $C^-(p)$.

It was shown in [2] that the deduction of $S(w_o,v^*)$ from $D(w_o,z^*)$ cannot be made cosmographically; in chapter 5 we shall prove that, given $D(w_o,z^*)$, the local structure on the past light cone is uniquely determined by Einstein's field equations. In addition, we shall indicate briefly how the solution may be propagated into the interior of $C^-(p)$ without assuming any further initial data.

5. THE COSMOLOGICAL CHARACTERISTIC INITIAL VALUE PROBLEM

§5.1. Introduction

The main aim of this chapter is to show that when Einstein's field equations for a pressure-free perfect fluid

$$R_{ab} = -M\,u_a u_b + \tfrac{1}{2} M g_{ab}$$

are assumed, the maximal data set $D(w_o, z^*)$ determines uniquely the local structure of the cosmological model on at least some part of the initial past light cone $C^-(p)$, i.e. it determines the set $S(w_o, v_L^{**})$ (see §4.5), where v_L^{**} satisfies $0 \le v_L^{**} \le v^*$, with v^* being the affine distance corresponding to the limiting redshift z^*. This result is established in §5.2.

It has been remarked before that the NP formalism does not appear to be particularly well-suited to a rigorous discussion of the problem of propagating the solution off $C^-(p)$, except in the analytic case. In §§5.3 and 5.7 we shall show that, without specifying any additional data, it is possible to deduce the w-derivatives of a suitable initial data set from the field equations, Bianchi identities and contracted Bianchi identities. This has generally been taken to imply that one can then determine the initial data "on a neighbouring hypersurface separated by an infinitesimal w-displacement" ([62]), or "at the next instant of w" ([59]), so that the model may be determined on this "next" light cone by again carrying out the hypersurface integration discussed in §5.2. Repeated application of this process then formally generates a solution in some region in the interior of $C^-(p)$.

Clearly, these rather vague and imprecise statements gloss over a number of severe technical difficulties. In order to obtain a proper existence and uniqueness proof for the solution off $C^-(p)$, it would seem that techniques similar to those employed by Müller zum Hagen and Seifert need to be used ([57]; in this connection, see also [82]).

In practice, the initial data will never be available in the explicit functional form required; rather, from a discrete set of data points, it is necessary to construct smooth (i.e. C^∞) functions by a variety of interpolating methods, such as fitting least squares polynomials to the data. In §5.4 we treat the case where the data is given in this polynomial form

$$a = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + o(z^{n+1}), \qquad (5.1.1)$$

where the "expansion coefficients" a_i are given functions of x^A, e.g.

$$\left(\sigma\frac{dv}{dz}\right)(w_o, z, x^A) = \left(\sigma\frac{dv}{dz}\right)_1(w_o, x^A)\, z + o(z^2) \qquad (5.1.2)$$

(cf. (5.4.1)). It is shown that the integration procedure of §5.2 may be simplified slightly when the data is given in this form, and the integration is then carried out explicitly, the solution being expressed in the form

$$a = a^0 + a^1 v + a^2 v^2 + \ldots + a^m v^m + o(v^{m+1}), \qquad (5.1.3)$$

where the expansion coefficients a^i of the solution are determined by the a_i through the field equations. The exact order to which the initial data will be assumed known as a function of z is given in (5.4.1); this is different for different elements of $D(w_o, z^*)$. Thus r is assumed known up to order z^5, while u^A is given only up to first order in z. The orders have been chosen partly with realistic observational problems in mind (thus, e.g., there are practical techniques for determining r(z), but $u^A(z)$ is unlikely to be determined to any degree of accuracy in the near future) and partly for consistency reasons (even if r is known up to, say, $o(z^{16})$, no further terms in the solution can be calculated unless u^A is known to at least second order in z). Given the data as specified in (5.4.1), we have calculated the solution to the highest order possible in v.

This explicit integration serves three purposes. Firstly, since it assumes a data set of the form that is likely to arise in practice, it may be used as a first step in actually implementing the observational approach; the next stage would involve a numerical calculation using the results of §5.4 in conjunction with the best currently available data.

Secondly, it provides a crude estimate of the "incompleteness" of the solution, given incomplete specification of the initial data, and may thus be seen as a first attempt to come to grips with one aspect of the "cosmological uncertainty" discussed in chapter 1.

Finally, the case of analytic initial data can be obtained from (5.4.1) by formally letting $n \to \infty$ in all expressions of the form (5.1.1).

In §5.5 it will be shown that when the initial data has the form (5.4.1), it is possible to derive the general angular behaviour of all the expansion coefficients in terms of the spin-s spherical harmonics introduced in §3.4. This appears to contradict a claim made by Chlellone and Williams ([83]) in their analysis of the perfect fluid characteristic initial value problem. For example, it can be shown that, if $\sigma\frac{dv}{dz}$ is given by (5.1.2), then

$$(\sigma\frac{dv}{dz}) = \frac{1}{3}[\frac{1}{3}\theta(w_o) - \sigma_{ij}(w_o)k^i k^j]^{-2} a^m(w_o)\,{}_2Y_{2m}, \quad (5.1.4)$$

where the k^i are defined by (3.4.59b) (or, equivalently, (3.4.51)) and θ, σ_{ij} and a^m are constants on $C^-(p)$ (the physical significance of these constants is discussed in §5.5). Here the summation convention again operates (m=-2,...,2).

The discussion in §§5.2 and 5.4 shows that the data set $D(w_o,z^*)$ is sufficient to determine the local structure of the model on $C^-(p)$ and (at least in the analytic case) to propagate the solution towards the interior of $C^-(p)$. However, not all 37 of the equations (3.3.13-19) are used during this integration and it is thus possible that the remaining equations will impose constraints on the initial data; it is even conceivable that the elements of $D(w_o,z^*)$ are not all independent.

In §5.6 it will be shown that no such constraints arise when the data is specified in the form (5.4.1). This case is relatively easy to handle, since one need only prove consistency up to the order in v to which the solution is known.

The general case is more difficult, one of the problems being that equations (3.3.13-19) are not all independent. As an example of this redundancy, note that, with the simplification induced by (3.3.22), equations (3.3.15d,e) imply (3.3.15c); more generally, it is well-known that the Ricci identities imply the Bianchi identities. The consistency problem is therefore greatly simplified if one works only with the 10 Einstein field equations and the contracted Bianchi identities (rather than with the full set of NP equations). This approach, originally used by Bondi et. al. ([59]) and Sachs ([60]) in their pioneering work on gravitational radiation, will be adopted in §5.7.

§5.2. The model on the light cone

Einstein's field equations (with vanishing cosmological constant) for a pressure-free perfect fluid with total energy density M and 4-velocity u^a may be written as

$$R_{ab} = -M\,(u_a u_b - \tfrac{1}{2} g_{ab}). \tag{5.2.1}$$

From the definition (3.2.4b) it follows that

$$\Phi_{00} = \tfrac{1}{2} M(1+z)^2 \tag{5.2.2a}$$

$$\Phi_{01} = -\tfrac{1}{2} M(1+z) Y^A \xi^B g_{AB} \tag{5.2.2b}$$

$$\Phi_{02} = \tfrac{1}{2} M(\, Y^A \xi^B g_{AB}\,)^2 \tag{5.2.2c}$$

$$\Phi_{11} = \tfrac{1}{8} M(1 - 2Y^A Y^B g_{AB}) \tag{5.2.2d}$$

$$\Phi_{12} = \tfrac{1}{4} M(1+z)^{-1} (Y^A \xi^B g_{AB})(Y^C Y^D g_{CD} - 1) \tag{5.2.2e}$$

$$\Phi_{22} = \tfrac{1}{8} M(1+z)^{-2} (Y^A Y^B g_{AB} - 1)^2 \tag{5.2.2f}$$

where

$$Y^A \equiv u^A + (1+z) X^A. \tag{5.2.3}$$

The central limits (3.4.47a,72a) imply

$$\Phi_{00} = o(1), \quad \Phi_{01} = o(v), \quad \Phi_{02} = o(v^2),$$
$$\Phi_{11} = o(1), \quad \Phi_{12} = o(v), \quad \Phi_{22} = o(1). \tag{5.2.4}$$

The Ricci scalar is given by (see (3.2.4d))

$$\Lambda = \frac{1}{24}R = \frac{1}{24}M \tag{5.2.5a}$$

and has the limiting behaviour

$$\Lambda = o(1). \tag{5.2.5b}$$

Suppose $D(w_o, z^*)$ is specified, so that $u^o = (1+z)$, u^A, r, $\sigma\frac{dv}{dz}$ and $n\frac{dv}{dz}$ are given as C^∞ functions of (z, x^A) on $C^-(p, z^*)$. We shall assume throughout that the following conditions are satisfied:

$$\frac{dr}{dz} \neq 0, \quad \rho \neq 0. \tag{5.2.6}$$

From the condition that the fluid be pressure-free, it can be shown (see [15]) that the total energy density M is proportional to the number density of sources n. We shall assume that the constant of proportionality is known from local astronomical observations, so that $M\frac{dv}{dz}$ may be regarded as given on $C^-(p, z^*)$.

By substituting for Φ_{00} from (5.2.2a) into (3.3.15a), we obtain the null Raychaudhuri equation

$$D\rho = \rho^2 + \sigma\bar{\sigma} + \frac{1}{2}M(1+z)^2 .$$

(Except for §5.7, all references to (3.3.13-19) assume that these equations have been simplified using (3.3.22).)

Using (4.2.17a), we may rewrite this in the form

$$f_1(z, x^A)\frac{d}{dz}\left(\frac{dz}{dv}\right) + f_2(z, x^A)\left(\frac{dz}{dv}\right) + f_3(z, x^A) = 0, \tag{5.2.7a}$$

where

$$f_1 \equiv r^{-1}\frac{dr}{dz} \tag{5.2.7b}$$

$$f_2 \equiv r^{-1}\frac{d^2r}{dz^2} + \sigma\bar{\sigma}\left(\frac{dv}{dz}\right)^2 \tag{5.2.7c}$$

$$f_3 \equiv \frac{1}{2}\left(M\frac{dv}{dz}\right)(1+z)^2. \tag{5.2.7d}$$

Now (5.2.7a) constitutes a linear first order differential equation for the unknown quantity $\frac{dz}{dv}$, provided $f_1 \neq 0$, a condition satisfied by virtue of (5.2.6). The coefficients f_i of this differential equation are all observable quantities, i.e. they are determined by the data set $D(w_o, z^*)$. From the central conditions (3.4.71a) and (4.2.18b), it can be shown that

$$\lim_{v \to 0} \left(\frac{dz}{dv} \right) = \lim_{v \to 0} \left(\frac{dz}{dr} \right) \equiv H_o(x^A), \tag{5.2.8}$$

where H_o is the Hubble parameter. In general, H_o will vary with direction (see, e.g., [80]).

The unique solution of (5.2.7a) subject to (5.2.8) is given by

$$\frac{dv}{dz} = F(z,x^A) \left(H_o - \int_o^z \frac{f_3(z',x^A) F(z',x^A)\, dz'}{f_1(z',x^A)} \right)^{-1} \tag{5.2.9a}$$

where

$$F(z,x^A) \equiv \left(\frac{dr}{dz} \right) \exp \left(\int_o^z \frac{[\sigma\bar{\sigma}(\frac{dv}{dz})^2](z',x^A)\, r(z',x^A)\, dz'}{(\frac{dr}{dz})(z',x^A)} \right) \tag{5.2.9b}$$

and now a further quadrature determines $v = v(z,x^A)$ uniquely, the constant of integration being fixed by the central condition $\lim_{v \to 0} z = 0$. Since $\frac{dv}{dz} \neq 0$ (a consequence of (5.2.6)), this relation may be inverted to give $z = z(v,x^A)$.

Thus the null Raychaudhuri equation determines the relationship between the affine distance v and observable measures of distance such as z or r. It is precisely the inability to determine this relationship without the use of field equations that prevents the deduction of $S(w_o, v^*)$ from $D(w_o, z^*)$ in the cosmographic analysis carried out in [2].

Once z has been determined as a function of (v,x^A), the 4-velocity components u^{0} and u^{A}, the spin coefficients ρ and σ and the metric variables ξ^{A} are determined on $C^{-}(p,v^{*})$; and now the Weyl tensor component Ψ_0 is given uniquely by

$$\Psi_0 = D\sigma - 2\rho\sigma. \tag{5.2.10}$$

At this stage of the integration, the only Ricci curvature components that are known are Λ and Φ_{00}. The others will be known once Y^{A} has been found and thus, by (5.2.3), once X^{A} has been determined. In fact, by (5.2.2a,b), we have

$$\Phi_{01} = h_1 - \Phi_{00}h_2, \tag{5.2.11a}$$

where

$$h_1 = -\frac{1}{2}M(1+z)u^{A}\xi^{B}g_{AB} , \tag{5.2.11b}$$

$$h_2 = X^{A}\xi^{B}\ g_{AB} . \tag{5.2.11c}$$

The identities

$$\xi^{A}\xi^{B}g_{AB} = 0 , \quad \xi^{A}\bar{\xi}^{B}g_{AB} = -1 , \tag{5.2.12}$$

which follow from (3.3.8e,10c), imply that, provided $\det g_{AB} \neq 0$, the relation (5.2.11c) inverts to give

$$X^{A} = -(h_2\bar{\xi}^{A} + \bar{h}_2\xi^{A}). \tag{5.2.11d}$$

Thus X^{A}, and hence Y^{A}, are determined once we know h_2. A differential equation for h_2 may be derived using (3.3.8e, 10c,13a,13c) and (5.2.12). We find

$$Dg_{AB} = -g_{AC}g_{BC}\,Dg^{CD} \tag{5.2.13}$$

and hence

$$Dh_2 = -\rho h_2 - \sigma\bar{h}_2 - (\bar{\alpha} + \beta). \tag{5.2.14}$$

Since α and β are also as yet unknown, an examination of (3.3.13,15,17) shows that we need to solve the following system of coupled differential equations:

$$D\alpha = \rho\alpha + \beta\bar{\sigma} - \Phi_{00}\bar{h}_2 + \bar{h}_1 \tag{5.2.15a}$$

$$D\beta = \rho\beta + \alpha\bar{\sigma} + \Psi_1 \tag{5.2.15b}$$

$$D\omega = \rho\omega + \sigma\bar{\omega} - \bar{\alpha} - \beta \tag{5.2.15c}$$

$$Dh_2 = -\rho h_2 - \sigma\bar{h}_2 - \bar{\alpha} - \beta \tag{5.2.15d}$$

$$\begin{aligned} D\Psi_1 = {} & 4\rho\Psi_1 + 3\bar{\alpha}\Phi_{00} - 4\alpha\Psi_0 + 3\beta\Phi_{00} - \omega D\Phi_{00} + \bar{\omega}D\Psi_0 + \\ & + h_2(3\rho\Phi_{00} - D\Phi_{00}) + 3\bar{h}_2\sigma\Phi_{00} + (Dh_1 - 2\rho h_1 - 2\sigma\bar{h}_1) + \\ & + \bar{\xi}^A\Psi_{0,A} + \xi^A\Phi_{00,A} \end{aligned} \tag{5.2.15e}$$

where we have used the fact, which follows from (3.2.5) and (3.3.9), that

$$\delta = \omega D + {}_{,A}\xi^A .$$

In order to show that (5.2.15), together with the appropriate central limits (see below), determines uniquely the unknown quantities α, β, ω, h_2 and Ψ_1, we employ a standard technique from the theory of differential equations, viz. the principle of contraction maps (see, e.g., [84]). The proof below was suggested by a study of the slightly different method of establishing existence and uniqueness used in [82].

The central limits (3.4.45a) imply

$$\rho = -v^{-1} + \tilde{\rho}, \quad \alpha = \alpha^0 v^{-1} + \tilde{\alpha} , \quad \beta = -\bar{\alpha}^0 v^{-1} + \tilde{\beta} , \tag{5.2.16a}$$

where

$$\lim_{v\to 0} \tilde{\rho} = 0, \quad \lim_{v\to 0} \tilde{\alpha} = 0, \quad \lim_{v\to 0} \tilde{\beta} = 0 . \tag{5.2.16b}$$

and α^0 is defined by (3.4.45b).

Since $\Psi_1 = o(1)$, $\omega = o(v^2)$ and $h_2 = o(v^2)$ (a consequence of (3.4.46,47) and (5.2.11c)),

$$\lim_{v\to 0} h_2 = 0, \quad \lim_{v\to 0} v\Psi_1 = 0, \quad \lim_{v\to 0} \omega = 0. \tag{5.2.16c}$$

Substituting (5.2.16a) into (5.2.15) now gives

$$D(v\tilde{\alpha}) = \tilde{\rho}(v\tilde{\alpha}) + \bar{\sigma}(v\tilde{\beta}) - v\Phi_{00}(h_2) + J_1 \tag{5.2.17a}$$

$$D(v\tilde{\beta}) = \tilde{\rho}(v\tilde{\beta}) + \sigma(v\tilde{\alpha}) + (v\Psi_1) + J_2 \tag{5.2.17b}$$

$$D(v\omega) = \tilde{\rho}(v\omega) + \sigma(v\bar{\omega}) - (v\bar{\tilde{\alpha}}) - (v\tilde{\beta}) \tag{5.2.17c}$$

$$D(h_2) = (v^{-1} - \tilde{\rho})(h_2) - \sigma(\bar{h}_2) - v^{-1}(v\bar{\tilde{\alpha}}) - v^{-1}(v\tilde{\beta}) \tag{5.2.17d}$$

$$\begin{aligned} D(v\Psi_1) = {} & (-3v^{-1} + 4\tilde{\rho})(v\Psi_1) + 3(v\bar{\tilde{\alpha}})\Phi_{00} - 4(v\tilde{\alpha})\Psi_0 + \\ & + 3(v\tilde{\beta})\Phi_{00} - (v\omega)D\Phi_{00} - (v\bar{\omega})D\Psi_0 + \\ & + (h_2)(3\Phi_{00} + \tilde{\rho}v\Phi_{00} - vD\Phi_{00}) + 3\sigma v\Phi_{00}(\bar{h}_2) + J_3 \end{aligned} \tag{5.2.17e}$$

with initial conditions

$$\lim_{v\to 0} (v\tilde{\alpha}, v\tilde{\beta}, v\omega, h_2, v\Psi_1) = 0, \tag{5.2.18}$$

where

$$J_1 = \tilde{\rho}\alpha^0 - \bar{\sigma}\bar{\alpha}^0 + \bar{h}_1 v \tag{5.2.19a}$$

$$J_2 = \bar{h}_1 v - J_1 \tag{5.2.19b}$$

$$\begin{aligned} J_3 = {} & v(Dh_1 - 2\rho h_1 - 2\sigma\bar{h}_1) + v(\bar{\xi}^A\Psi_{0,A} - \xi^A\Phi_{00,A}) + \\ & - 4\alpha^0\Psi_0 \end{aligned} \tag{5.2.19c}$$

are known functions at this stage of the integration, and satisfy

$$J_1 = o(v), \quad J_2 = o(v), \quad J_3 = o(1). \tag{5.2.19d}$$

Let $C[0,v^*]$ denote the set of continuous functions on $[0,v^*]$ that vanish at 0, and denote by $(\times C[0,v^*])^5$ the 5-fold product. Let $\phi,\psi \in C[0,v^*]$ and $\Phi \equiv (\phi^1,\phi^2,\ldots,\phi^5)$, $\Psi \equiv (\psi^1,\psi^2,\ldots,\psi^5) \in (\times C[0,v^*])^5$. It is well-known that if we define (positive definite) metrics e and E on $C[0,v^*]$ and $(\times C[0,v^*])^5$, respectively, by

$$e(\phi,\psi) \equiv \sup_{v\in[0,v^*]} |\phi(v) - \psi(v)| \tag{5.2.20a}$$

$$E(\Phi,\Psi) \equiv \sup_{v\in[0,v^*]} \{|\phi^1(v) - \psi^1(v)| + |\phi^2(v) - \psi^2(v)| + \ldots + |\phi^5(v) - \psi^5(v)|\} \qquad (5.2.20b)$$

then $\{C[0,v^*],e\}$ and $\{(\times C[0,v^*])^5,E\}$ are Banach spaces†.

A *contraction map* F on a metric space (X,d) is a map that satisfies

$$d(Fx,Fy) \le k\,d(x,y) \quad \text{for all } x,y \in X, \qquad (5.2.21a)$$

where k is a real number and

$$0 < k < 1. \qquad (5.2.21b)$$

The principle of contraction maps (see, e.g., [84]) states that a contraction map on a complete metric space has a unique fixed point, i.e. a point $x_0 \in X$ such that

$$Fx_0 = x_0.$$

Define

$$H: (\times C[0,v^*])^5 \to (\times C[0,v^*])^5$$

by

$$(H\phi^1)(v) = v^{-1}\int_0^v v'\{\tilde{\rho}\phi^1 + \bar{\sigma}\phi^2 - 4\Phi_{00}v'\phi^4 + 4J_1\}dv' \qquad (5.2.22a)$$

$$(H\phi^2)(v) = v^{-1}\int_0^v v'\{\tilde{\rho}\phi^2 + \sigma\phi^1 + \frac{1}{2}(v')^{-1}\phi^5 + 4J_2\}dv' \qquad (5.2.22b)$$

$$(H\phi^3)(v) = v^{-1}\int_0^v v'\{\tilde{\rho}\phi^3 + \sigma\phi^3 - \frac{1}{4}\phi^1 - \frac{1}{4}\phi^2\}dv' \qquad (5.2.22c)$$

$$(H\phi^4)(v) = v^{-1/4}\int_0^v \{[\frac{1}{4}(v')^{-3/4} - \tilde{\rho}(v')^{1/4}]\phi^4 + -\sigma(v')^{1/4}\bar{\phi}^4 - \frac{1}{4}(v')^{-3/4}\times \times(\bar{\phi}^1 + \phi^2)\}dv' \qquad (5.2.22d)$$

$$(H\phi^5)(v) = v^{-3}\int_0^v (v')^3\{4\tilde{\rho}\phi^5 + 6\Phi_{00}v'\bar{\phi}^1 - 8\Psi_0 v'\phi^1 +$$

---

† Strictly, $\{C[0,v^*],\|\ \|_e\}$ is a Banach space, with $\|\phi\|_e \equiv e(\phi,\phi)$ being a norm on the linear space $C[0,v^*]$, but we shall allow ourselves this slight abuse of notation.

$$+ 6\Phi_{00}v'\phi^2 - 8v'(D\Phi_{00})\phi^3 + 8v'(D\Psi_0)\bar{\phi}^3 +$$

$$+ 8\{3\tilde{\rho}(v')^2\Phi_{00} - 3v'\Phi_{00} - (v')^2D\Phi_{00}\}\phi^4 +$$

$$+ 24\sigma(v')^2\Phi_{00}\bar{\phi}^4 + 8J_3\}dv' \qquad (5.2.22e)$$

for $v \in (0,v^*]$ and

$$(H\Phi)(0) = 0. \qquad (5.2.22f)$$

Suppose for the moment that H is well-defined (this will be proved shortly), and consider the system of integral equations

$$\Phi = H\Phi . \qquad (5.2.23)$$

If we let

$$\Phi = (4\tilde{\alpha}, 4\tilde{\beta}, \omega, v^{-1}h_2, 8v\Psi_1), \qquad (5.2.24)$$

it is readily shown that (5.2.23) is equivalent to the system of differential equations (5.2.17) subject to the initial conditions (5.2.18). Hence, in order to establish an existence and uniqueness result for (5.2.17), we need merely prove existence and uniqueness for (5.2.23), i.e. we need to show that H is a contraction map on the complete metric space $\{(\times C[0,v^*])^5, E\}$.

Lemma 5.1.

The map H defined by (5.2.22) maps $(\times C[0,v^*])^5$ into itself, i.e.

$$H\Phi \in (\times C[0,v^*])^5 \quad \text{for all} \quad \Phi \in (\times C[0,v^*])^5.$$

Proof.

Since all the functions involved in the definition (5.2.22) are continuous on $[0,v^*]$, the only point at which $H\Phi$ can possibly be discontinuous is at 0. But, from the limiting behaviour of these functions, it follows that there exists some positive real number B such that $|\tilde{\rho}|$, $|\tilde{\rho}v|$, $|\sigma|$, $|\Phi_{00}v|$, $|\Psi_0 v|$, $|vD\Phi_{00}|$, $|vD\Psi_0|$ and $|J_i|$ are all bounded by B on $[0,v^*]$. Thus

$$|H\phi^1| \le \tfrac{1}{2}Bv\,(\sup|\phi^1| + \sup|\phi^2| + 4\sup|\phi^4|) \tag{5.2.25a}$$

$$|H\phi^2| \le \tfrac{1}{2}Bv\,(\sup|\phi^1| + \sup|\phi^2| + 4) + \tfrac{1}{2}\sup|\phi^5| \tag{5.2.25b}$$

$$|H\phi^3| \le \tfrac{1}{2}Bv\,(\tfrac{1}{4}\sup|\phi^1| + \tfrac{1}{4}\sup|\phi^2| + 2\sup|\phi^3|) \tag{5.2.25c}$$

$$|H\phi^4| \le (\tfrac{1}{3} + \tfrac{8}{5}Bv)\sup|\phi^4| + \tfrac{1}{3}(\sup|\phi^1| + \sup|\phi^2|) \tag{5.2.25d}$$

$$\begin{aligned}|H\phi^5| \le{}& \tfrac{1}{2}Bv\,(7\sup|\phi^1| + 3\sup|\phi^2| + 8\sup|\phi^3| + \\ &+12\sup|\phi^4| + 2\sup|\phi^5| + 4) + \\ &+ \tfrac{8}{5}Bv^2(3B+4)\sup|\phi^4|\ ,\end{aligned} \tag{5.2.25e}$$

where all supremums are taken over $[0,v^*]$.

Since $\phi^i(0) = 0$ and ϕ^i are continuous, it follows that $\lim_{v\to 0}\{\sup|\phi^i(v)|\} = 0$; and now (5.2.25) shows that

$$\lim_{v\to 0}(H\Phi)(v) = 0 = (H\Phi)(0)$$

as required. □

Theorem 5.1.

The map H defined by (5.2.22) is a contraction map on $(\times C[0,v^{**}])^5$ for some v^{**} satisfying $0 < v^{**} \le v^*$.

Proof.

Let B be defined as in Lemma 5.1. As in the proof above, we obtain bounds on $|H\phi^i - H\psi^i|$ which yield

$$\begin{aligned}\sum_{i=1}^{5}|H\phi^i(v) - H\psi^i(v)| \le{}& (\tfrac{37}{8}Bv + \tfrac{1}{3})(\sup|\phi^1 - \psi^1|) + \\ &+ (\tfrac{21}{8}Bv + \tfrac{1}{3})(\sup|\phi^2 - \psi^2|) + \\ &+ 5Bv(\sup|\phi^3 - \psi^3|) +\end{aligned}$$

$$+ \{\frac{48}{5}Bv + \frac{8}{5}Bv^2(3B+4) + \frac{1}{3}\}(\sup|\phi^4 - \psi^4|) +$$

$$+ \frac{5}{4}Bv(\sup|\phi^5 - \psi^5|) \tag{5.2.26a}$$

Now let $\frac{1}{3} < k < 1$ and choose v** to satisfy

$$\frac{37}{8}Bv^{**} + \frac{1}{3} \le k, \quad 5Bv^{**} \le k,$$

$$\frac{48}{5}Bv^{**} + \frac{8}{5}B(v^{**})^2(3B+4) + \frac{1}{3} \le k, \quad 0 < v^{**} \le v^*. \tag{5.2.26b}$$

Then

$$\sum_{i=1}^{5} |H\phi^i(v) - H\psi^i(v)| \le k \sum_{i=1}^{5} \sup_{v\in[0,v^{**}]} |\phi^i(v) - \psi^i(v)| \tag{5.2.26c}$$

for all v in $[0,v^{**}]$. Now it follows immediately from (5.2.20b) and (5.2.26c) that

$$E(H\Phi,H\Psi) < kE(\Phi,\Psi) \quad \text{on } [0,v^{**}]$$

as required.□

At this stage of the integration, the following quantities are therefore determined on $C^-(p,v^{**})$: ρ, σ, α, β, $\tau\ (=\bar{\alpha}+\beta)$, ξ^A, ω, X^A, Φ_{IJ} (I,J=0,1,2), Ψ_0, Ψ_1; and the following equations have been satisfied: (3.3.13a-c,15a-e,17a). (Note that(3.3.15c) is implied by (3.3.15d,e).)

Next consider the group of equations (3.3.15g,h) and (3.3.17b).

$$D\lambda = \rho\lambda + \bar{\sigma}\mu + \Phi_{20} \tag{5.2.27a}$$

$$D\mu = \rho\mu + \sigma\lambda + \Psi_2 + 2\Lambda \tag{5.2.27b}$$

$$D\Psi_2 = 3\rho\Psi_2 + D(\Phi_{11}+\Lambda) - \delta\Phi_{10} + \bar{\delta}\Psi_1 - \lambda\Psi_0 +$$

$$- 2\alpha\Psi_1 + \mu\Phi_{00} + 2\bar{\alpha}\Phi_{10} - 2\rho\Phi_{11} - \sigma\Phi_{20}. \tag{5.2.27c}$$

Let

$$\tilde{\mu} \equiv \mu - \mu^0 v^{-1} \tag{5.2.28a}$$

and set

$$\Phi \equiv (\phi^1, \phi^2, \phi^3) \equiv (v\lambda, v\tilde{\mu}, 2v\Psi_2).$$

Then substitution of (5.2.28) into (5.2.27) yields a system of differential equations which is equivalent to the system of integral equations

$$\Phi = F\Phi \ , \tag{5.2.29a}$$

where

$$F:(\times C[0,v^{**}])^3 \to (\times C[0,v^{**}])^3$$

is defined by

$$(F\phi^1)(v) = \int_0^v \{\tilde{\rho}\phi^1 + \bar{\sigma}\phi^2 + G_1\}dv' \tag{5.2.29b}$$

$$(F\phi^2)(v) = \int_0^v \{\tilde{\rho}\phi^2 + \sigma\phi^1 + \frac{1}{2}\phi^3 + G_2\}dv' \tag{5.2.29c}$$

$$(F\phi^3)(v) = v^{-2}\int_0^v (v')^2\{3\tilde{\rho}\phi^3 - 2\Psi_0\phi^1 + 2\Phi_{00}\phi^2 + $$

$$+ G_3\}dv' \tag{5.2.29d}$$

for $v \in (0,v^{**}]$ and

$$(F\Phi)(0) = 0, \tag{5.2.29e}$$

with

$$G_1 = \bar{\sigma}\mu^0 + v\Phi_{20} \quad = o(v) \tag{5.2.29f}$$

$$G_2 = \tilde{\rho}\mu^0 + 2v\Lambda \quad = o(v) \tag{5.2.29g}$$

$$G_3 = 2vD(\Phi_{11} + \Lambda) - 2v\delta\Phi_{10} + 2v\bar{\delta}\Psi_1 - 4v\Psi_1\alpha +$$

$$+2\mu^0\Phi_{00} + 4\bar{\alpha}v\Phi_{10} - 4\rho v\Phi_{11} - 2\sigma v\Phi_{20}$$

$$= o(1). \tag{5.2.29h}$$

It is readily verified F is a contraction map on $(\times C[0,v])^3$ for v sufficiently small, and thus λ, μ and Ψ_2 are uniquely determined on $C^-(p,v^{***})$, where $0 < v^{***} \le v^{**}$.

All remaining quantities may now be found by integrating sequentially the following system of *uncoupled* differential equations:

$$D\gamma = \tau\alpha + \bar{\tau}\beta + \Psi_2 - \Lambda + \Phi_{11} \tag{5.2.30a}$$

$$DU = \bar{\tau}\omega + \tau\bar{\omega} - (\gamma+\bar{\gamma}) \tag{5.2.30b}$$

$$\begin{aligned} D\Psi_3 = {} & D\Phi_{21} - \delta\Phi_{20} + \bar{\delta}(\Psi_2+2\Lambda) - 2\lambda\Psi_1 + 2\rho\Psi_3 + \\ & + 2\mu\Phi_{10} + 2(\bar{\alpha}-\beta)\Phi_{20} - 2\rho\Phi_{21} \end{aligned} \tag{5.2.30c}$$

$$D\nu = \bar{\tau}\mu + \tau\lambda + \Psi_3 + \Phi_{21} \tag{5.2.30d}$$

$$\begin{aligned} D\Psi_4 = {} & \bar{\delta}(\Phi_{21}+\Psi_3) - 3\lambda\Psi_2 + 2\alpha\Psi_3 + \rho\Psi_4 + 2\nu\Phi_{10} + \\ & -2\lambda\Phi_{11} + (2\bar{\gamma}-2\gamma-\bar{\mu})\Phi_{20} - 2\bar{\beta}\Phi_{21} + \bar{\sigma}\Phi_{22} + \Delta\Phi_{20} \end{aligned} \tag{5.2.30e}$$

Equations (5.2.30a,b,d) determine γ,U and ν immediately by simple quadratures, the constants of integration being determined by the central conditions (3.4.45a,46a); while (5.2.30c) is a linear first order differential equation for Ψ_3:

$$\begin{aligned} D(v^2\Psi_3) = {} & 2\tilde{\rho}(v^2\Psi_3) + 2\mu v^2\Phi_{10} + 2v^2(\bar{\alpha}-\beta)\Phi_{20} + \\ & -2\rho v^2\Phi_{21} - 2\lambda v^2\Psi_1 + v^2 D\Phi_{21} - v^2\delta\Phi_{20} + \\ & + v^2\bar{\delta}(\Psi_2+2\Lambda) \end{aligned} \tag{5.2.31}$$

Since $\lim\limits_{v\to 0} v^2\Psi_3 = 0$, (5.2.31) determines $v^2\Psi_3$ and hence Ψ_3 uniquely. (If the central condition $\Psi_3 = o(1)$ did not hold, Ψ_3 would be determined only up to an unknown term in v^{-2}.)

Similarly (5.2.30e) determines Ψ_4 uniquely even though the right hand side of this equation contains the term $\Delta\Phi_{20}$, which involves derivatives in directions off the light cone. This is because, by (5.2.2),

$$\Phi_{02} = (\Phi_{01})^2(\Phi_{00})^{-1},$$

and $\Delta\Phi_{00}$, $\Delta\Phi_{01}$ are known at this stage of the integration; they are given explicitly by the contracted Bianchi identities (3.3.19a,b).

At each step s in the integration, the solution is determined on some bounded subset $C^-(p,v_s)$ of the past light cone $C^-(p)$, where $v^* \geq v_1 \geq v_2 \geq \ldots \geq v_n$. Thus, even if the initial data is completely specified down to a redshift z^* corresponding to an affine distance v^*, the proof above guarantees existence and uniqueness of the solution on $C^-(p)$ only down to an affine distance which we shall denote by v_L^{**}, with $0 < v_L^{**} \leq v^*$.

Collecting results, we have therefore shown that, under the assumptions (5.2.1) and (5.2.6), the maximal data set $D(w_o,z^*)$ determines uniquely all the spin coefficients, the metric variables and the tetrad components of the Riemann tensor on $C^-(p,v_L^{**})$. By substituting the metric variables (ξ^A,ω,X^A,U) into (3.3.10), the metric tensor on $C^-(p,v_L^{**})$ may be found. Furthermore, by (5.2.2a), M is known (once z has been determined as a function of (v,x^A)); u^A are measurable; $u^o = (1+z)$ and u^1 is determined by the normalisation condition $u^a u_a = 1$ (see equation (4.2.7)). Thus $S(w_o,v_L^{**})$ is uniquely determined.

In the course of deducing $S(w_o,v_L^{**})$, the following equations have been satisfied: (3.3.13,15,17). While the argument above establishes that $D(w_o,z^*)$ is sufficient to determine $S(w_o,v_L^{**})$, it does not demonstrate that $D(w_o,z^*)$ is also the minimal data required to determine $S(w_o,v_L^{**})$. For it is conceivable that the elements of $D(w_o,z^*)$ are not independent; this would be the case if any of the remaining non-propagation equations (i.e. equations which contain terms involving only derivatives *within* the past light cone) imply interrelations between the elements of $D(w_o,z^*)$. In order to prove that this is not the case we shall show that, once the central conditions (3.4.45,46,71a) and the radial equations (3.3.13,15,17) have been satisfied, all the remaining non-propagation equations - (3.3.14c,d) and (3.3.16b, c,d) - are automatically satisfied. The proof is similar to that used in the vacuum case ([82]) and will therefore only be sketched. Let

$$a_1 \equiv \delta\rho - \bar{\delta}\sigma - \rho\tau + \sigma(3\alpha-\bar{\beta}) + \Psi_1 - \Phi_{01} \tag{5.2.32a}$$

$$a_2 \equiv \delta\alpha - \bar{\delta}\beta - \mu\rho + \lambda\sigma - \alpha\bar{\alpha} - \beta\bar{\beta} + 2\alpha\beta + \Psi_2 - \Lambda - \Phi_{11} \tag{5.2.32b}$$

$$a_3 \equiv \delta\lambda - \bar{\delta}\mu - \mu\bar{\tau} + \lambda(3\beta-\bar{\alpha}) + \Psi_3 - \Phi_{21} \tag{5.2.32c}$$

$$b_1^A \equiv \delta\bar{\xi}^A - \bar{\delta}\xi^A + (\alpha-\bar{\beta})\xi^A + (\beta-\bar{\alpha})\bar{\xi}^A \tag{5.2.32d}$$

$$b_2 \equiv \delta\bar{\omega} - \bar{\delta}\omega + (\alpha-\bar{\beta})\omega + (\beta-\bar{\alpha})\bar{\omega} + \bar{\mu} - \mu \tag{5.2.32e}$$

so that (3.3.14c,d) and (3.3.16b,c,d) are given, respectively, by $b_1^A = 0$, $b_2 = 0$, $a_1 = 0$, $a_2 = 0$, $a_3 = 0$. Label the Bianchi identities (3.3.17) similarly:

$$B_1 \equiv D(\Phi_{01}-\Psi_1) - \delta\Phi_{00} + \bar{\delta}\Psi_0 - 4\alpha\Psi_0 + 4\rho\Psi_1 + \\ + 2\tau\Phi_{00} - 2\rho\Phi_{01} - 2\sigma\Phi_{10} \tag{5.2.33a}$$

$$B_2 \equiv D(\Phi_{11}+\Lambda-\Psi_2) + \bar{\delta}\Psi_1 - \delta\Phi_{10} - \lambda\Psi_0 - 2\alpha\Psi_1 + \\ + 3\rho\Psi_2 + \mu\Phi_{00} + 2\bar{\alpha}\Phi_{10} - 2\bar{\rho}\Phi_{11} - \sigma\Phi_{20} \tag{5.2.33b}$$

$$B_3 \equiv D(\Phi_{21}-\Psi_3) + \bar{\delta}(\Psi_2+2\Lambda) - \delta\Phi_{20} - 2\lambda\Psi_1 + \\ + 2\rho\Psi_3 + 2\mu\Phi_{10} + 2(\bar{\alpha}-\beta)\Phi_{20} - 2\bar{\rho}\Phi_{21} \,. \tag{5.2.33c}$$

Using the commutators (3.3.12), an extremely tedious calculation gives

$$Da_1 = 3\rho a_1 - \sigma\bar{a}_1 + B_1 \tag{5.2.34a}$$

$$Da_2 = 2\rho a_2 + \alpha a_1 - \beta\bar{a}_1 + B_2 \tag{5.2.34b}$$

$$Da_3 = 2\rho a_3 + \lambda a_1 - \mu\bar{a}_1 + B_3 \tag{5.2.34c}$$

$$Db_1^A = 2\rho b_1^A + a_1\bar{\xi}^A - \bar{a}_1\xi^A \tag{5.2.34d}$$

$$Db_2 = 2\rho b_2 + \bar{\omega}a_1 - \omega\bar{a}_1 + \bar{a}_2 - a_2 \ . \tag{5.2.34e}$$

Now, by (3.4.45,46), it follows that $\lim\limits_{v\to 0} a_1 v^3 = 0$. But (5.2.16,34a) imply

$$D(a_1v^3) = 3\tilde{\rho}(a_1v^3) - \sigma(\bar{a}_1v^3) + B_1v^3. \tag{5.2.35}$$

Provided $B_1 = 0$ (which follows from the assumption that (3.3.17a) is satisfied), the unique solution of (5.2.35) subject to $\lim\limits_{v\to 0} a_1v^3 = 0$ is

$$a_1v^3 = 0 \quad \text{on } [0,v_L^{**}]$$

and hence

$$a_1 = 0 \quad \text{on } (0,v_L^{**}].$$

Thus $a_1 = 0$ is satisfied on $[0,v_L^{**}]$ provided $a_1|_0 = 0$. But this is precisely the condition (see (3.4.24b) and the discussion following) that $\lim\limits_{v\to 0} v\omega = 0$, which is satisfied by virtue of the central condition $\omega = o(v^2)$.

In an entirely analogous fashion, it may be shown that a_2, a_3, b_1^A and b_2 all vanish identically on $[0,v_L^{**}]$, i.e. (3.3.14c,d) and (3.3.16b,c,d) are automatically satisfied on $[0,v_L^{**}]$.

§5.3. Propagation of the solution off $C^-(p)$

It has been remarked before (see §5.1; cf. also §2.3 of [82]) that the NP-formalism does not appear to be ideally suited to discussing, rigorously, the propagation of the solution on $C^-(p,v_L^{**})$ into the interior of the past light cone, except perhaps in the analytic case. It seems that the methods employed in [57] are more suitable for this purpose.

In this section we shall indicate briefly how, without additional specification of initial data, the solution obtained in §5.2 may be propagated into the interior of $C^-(p,v^{**}_L)$. The method we shall use is similar to that of Bondi et. al. ([59]), Sachs ([60]) and Newman and Unti ([61]). Even though this method guarantees the existence of a unique solution off $C^-(p,v^{**}_L)$ only in the analytic case, it may nevertheless be regarded as a necessary first step in obtaining a rigorous existence and uniqueness proof for the non-analytic case; almost all such proofs start by assuming analyticity and then extend the results to function spaces in which the analytic functions are dense with respect to the topology induced by a suitably chosen norm.

Suppose then that the integration procedure described in §5.2 has been carried out, so that equations (3.3.13), (3.3.14c,d), (3.3.15), (3.3.16b,c,d) and (3.3.17) are all satisfied. The spin coefficients, metric variables and the tetrad components of the metric and curvature tensors are known on $C^-(p,v^{**}_L)$.

Now consider equations (3.3.14a) and (3.3.19). The first of these gives

$$\Delta\xi^A = \delta X^A + (\gamma-\bar{\gamma}-\mu)\xi^A - \bar{\lambda}\bar{\xi}^A \qquad (5.3.1)$$

as a known quantity on $C^-(p,v^{**}_L)$. Since $D\xi^A$ and $\delta\xi^A$ are also known, (3.2.5) and (3.3.9) show that $\xi^A{}_{,0}$ is determined. The contracted Bianchi identities (3.3.19) may now, using (5.2.2a,b,d) and (5.2.5a), be written as a system of linear algebraic equations for the unknown quantities ΔM, Δz and ΔY^A:

$$Fx = G, \tag{5.3.2a}$$

where

$$F \equiv -\frac{1}{2}\begin{pmatrix} -(1+z)^2 & -2M(1+z) & 0 & 0 \\ (1+z)Y^A\xi^B_R g_{AB} & MY^A\xi^B_R g_{AB} & M(1+z)\xi^B_R g_{2B} & M(1+z)\xi^B_R g_{3B} \\ (1+z)Y^A\xi^B_I g_{AB} & MY^A\xi^B_I g_{AB} & M(1+z)\xi^B_I g_{2B} & M(1+z)\xi^B_I g_{3B} \\ \frac{1}{2}(Y^AY^Bg_{AB}-1) & 0 & MY^Bg_{2B} & MY^Bg_{3B} \end{pmatrix} \tag{5.3.2b}$$

(here subscripts R and I denote, respectively, real and imaginary parts),

$$x \equiv \begin{pmatrix} \Delta M \\ \Delta z \\ \Delta Y^2 \\ \Delta Y^3 \end{pmatrix} \tag{5.3.2c}$$

and G is a (4×4) real matrix which is known once $\Delta\xi^A$ has been determined using (5.3.1). This system of equations has a unique solution provided $\det F \neq 0$; a short calculation gives

$$\det F = \frac{1}{32}M^3(1+z)^3(\det g_{AB})^{3/2} \neq 0. \tag{5.3.3}$$

(More generally, if one does not assume the fluid to be pressure-free, the contracted Bianchi identities can be shown to determine ΔM, Δz and ΔY^A provided the equation of state relating the pressure p and the energy density M is such that $p \neq M$.)

By taking the Δ-derivative of (3.3.13c) and using (3.3.11a,13c,14a) and (3.3.16f,g), it can be shown that

$$D(\Delta X^A) = 2Re\{\xi^A[\bar{\delta}\gamma + \delta\bar{\gamma} + \rho\nu + \bar{\sigma}\bar{\nu} - 3(\bar{\alpha}+\beta)\lambda - \bar{\mu}(\alpha+2\bar{\beta}) + 2(\gamma-\bar{\gamma})\alpha + (\gamma+\bar{\gamma})(\alpha+\bar{\beta}) - \mu\alpha] + 2\tau\delta X^A\} \tag{5.3.4}$$

Since the right-hand side of (5.3.4) is known, one may integrate this equation (the central conditions imply $\Delta X^A|_0 = 0$) to determine ΔX^A on $C^-(p,v_L^{**})$. Once ΔX^A and Δz are known, it follows from (5.2.3) that Δu^A is determined on $C^-(p,v_L^{**})$. Furthermore, (3.3.16g) gives $\Delta\sigma$ directly on $C^-(p,v_L^{**})$. Thus ΔM, Δz, Δu^A and $\Delta\sigma$ are known, and now (3.3.9) shows that $M,_0$, $z,_0$, $u^A,_0$ and $\sigma,_0$ are known on $C^-(p,v_L^{**})$.

Now it is clear from the discussion in §5.2 that the set

$$I(w_0,v^*) = \{(u^0,u^A,\sigma,M)\,|\,w=w_0,\ 0\le v\le v^*\}$$

is a perfectly good (but not observable) initial data set for the NP equations; the integration proceeds exactly as before, except that, at the first step, the null Raychaudhuri equation (3.3.15a) is written, using (4.2.17a), as

$$r^{-1}\frac{d^2r}{dv^2} = \sigma\bar{\sigma} + \tfrac{1}{2}M(1+z)^2. \tag{5.3.5}$$

Since both terms on the right-hand side of (5.3.5) are known (as functions of v and x^A) from $I(w_0,v^*)$, this equation constitutes a differential equation for the unknown quantity r and therefore determines the relationship between the area distance and affine distance. (When the set $D(w_0,z^*)$ is used as initial data, (3.3.15a) determines the relationship between z and v.)

Thus the initial data set $D(w_0,z^*)$ is used to find the solution on $C^-(p,v_L^{**})$; in particular, it determines the initial data set $I(w_0,v_L^{**})$ as well as the w-derivatives of all the elements of $I(w_0,v_L^{**})$. Now, following Bondi et. al. (see §5.1), one argues that the initial data set $I(w_0,v_L^{**})$ can therefore be determined on some bounded subset $C^-(p',v'^*)$

of a "neighbouring" past light cone $C^-(p')$, where p' is a point on C to the past of p, corresponding to the proper time $w = w_o - dw$ on C. Once $I(w,v'^*)$ is known on $C^-(p,v'^*)$, the solution may be found on $C^-(p',v_L'^{**})$ by performing the hypersurface integration described in §5.2. The w-derivatives of the elements of $I(w,v_L'^{**})$ are again determined, and thus an iterative procedure is established which, it is claimed, determines the solution off the past light cone $C^-(p)$.

There are several problems associated with this approach. Firstly, it is clear from the hyperbolic nature of the field equations that, in general, specification of initial data on one null hypersurface cannot determine the solution throughout the space-time (except in the analytic case), unless one assumes, a priori, that the space-time satisfies certain global restrictions that render it "deterministic".

The problem of finding such global restrictions that will allow one to predict to the future of $C^-(p)$, given data only on $C^-(p)$, has been discussed by Budic and Sachs ([58]). They consider space-times $(\mathcal{M},g)$ which satisfy the condition that, for each point p in $\mathcal{M}$, every past-endless causal curve which intersects the chronological future of p also intersects its chronological past. Unfortunately this definition severely restricts the possible space-time structure: e.g., every deterministic space-time contains compact Cauchy surfaces, whereas "there does not seem to be any physically compelling reason for believing that the universe admits a Cauchy surface" ([25], p.206). Furthermore ([58]), almost none of the standard cosmological models satisfy this condition, and one of the essential features of the observational approach is that one does not wish to exclude, a priori, a whole class of possible cosmological models.

One alternative is to impose, again a priori, initial data on some hypersurface intersecting $C^-(p)$. Again, this would be contrary to the spirit of the observational approach, since such data would, of course, be completely unobservable.

Assuming analaticity is also unsatisfactory, since the hyperbolic nature of the field equations seems to be a fundamental feature of general relativity. But even in the analytic case, the argument above is not yet complete. It is still necessary to show that the remaining equations which contain Δ-derivatives are now identically satisfied. The spin coefficient α, e.g., may be determined on $C^-(p')$ by integrating the system of equations (5.2.17) using the initial data set $I(w,v'^{**}_L)$. On the other hand, it is also possible, in the analytic case, to determine α on this light cone by expanding α in a Taylor series in w about points on $C^-(p)$ and using (3.3.16i) (which determines $\alpha_{,0}$). Unless (3.3.16i) is identically satisfied by virtue of previously derived relations, it is possible that these two methods of finding on $C^-(p')$ will lead to different results. The same method that was used in §5.2 to establish that the remaining non-propagation equations are automatically satisfied can be used to verify that no such consistency problems do, in fact, arise. However (see §5.1), this consistency proof is considerably simplified if one works only with the 10 field equations and contracted Bianchi identities, as we shall do in §5.7.

Anticipating the results of that discussion, it is clear that the observationally determined data is sufficient to determine the solution towards the interior of the light cone in the analytic case. There is no reason to think that the situation would be different in the non-analytic case, particularly in view of results by Muller zum Hagen and Seifert on the fluid characteristic initial value problem (they do not consider the case of a *cone*, but it should be possible to round off the vertex to cases they do consider). Thus while it would be useful to have the rigorous proofs available, general arguments from the hyperbolic nature of the field equations, together with the sufficiency of the data in the analytic case, suggest that the solution will be uniquely determined in some subset of the past Cauchy development of the set on which the data is known. If the results of §5.2 could be strengthened to show that, given data on $C^-(p,v^*)$, the solution is determined on $C^-(p,v^*)$ (rather than on $C^-(p,v_L^{**})$ with $v_L^{**} \leq v^*$), it seems likely that the solution will be determined in precisely the past Cauchy development of $C^-(p,v^*)$. From the structure of the field equations we may suspect that once the contracted Bianchi identities are satisfied, no further constraints will arise as we integrate towards the interior (not only will the data be sufficient, but there will be no superfluous data); i.e., not only uniqueness but also existence may be expected to hold in a suitable domain.

§5.4. Explicit integration for polynomial data

The integration procedure presented in §5.2 assumes that the maximal data set $D(w_o,z^*)$ is known to indefinite accuracy on $C^-(p,v^*)$. In this section we shall examine the slightly more realistic case where the initial data is specified in the form

$$r(z,x^A) = r_1 z + r_2 z^2 + \ldots + r_5 z^5 + o(z^6) \tag{5.4.1a}$$

$$(\sigma\frac{dv}{dz})(z,x^A) = (\sigma\frac{dv}{dz})_1 z + (\sigma\frac{dv}{dz})_2 z^2 + (\sigma\frac{dv}{dz})_3 z^3 + o(z^4) \tag{5.4.1b}$$

$$(M\frac{dv}{dz})(z,x^A) = (M\frac{dv}{dz})_0 + (M\frac{dv}{dz})_1 z + (M\frac{dv}{dz})_2 z^2 + o(z^3) \tag{5.4.1c}$$

$$u^A(z,x^B) = u_0^A + u_1^A z + o(z^2). \tag{5.4.1d}$$

It will also be assumed that two formal derivatives of (5.4.1) are valid; thus, e.g.,

$$\frac{d^2r}{dz^2} = 2r_2 + 6r_3 z + 12r_4 z^2 + 20r_5 z^3 + o(z^4).$$

(Throughout this section, subscripts denote expansion coefficients in an expansion in powers of z, while superscripts denote coefficients in an expansion in terms of v. Thus, e.g., r is given both by (5.4.1) and by

$$r = v + r^3v^3 + r^4v^4 + r^5v^5 + o(v^6).$$

At this stage, of course, the r^i are unknown, whereas the r_i are given functions of (z,x^A).)

The integration can now be carried out explicitly to determine the spin coefficients, metric variables and Riemann tensor components up to various orders in the affine parameter v. The form of the initial data (5.4.1) allows a slight simplification of the procedure used in §5.2: the rather complicated system (5.2.15) may be replaced by an equivalent but simpler system of equations, as will be shown below.

The first step in the integration is to determine $z = z(v)$ using (5.2.7a). From (5.4.1) it follows that the functions f_1, f_2 and f_3 defined by (5.2.7b-d) are given by

$$f_1 = z^{-1}\{1 + \frac{r_2}{r_1}z + [2\frac{r_3}{r_1} - (\frac{r_2}{r_1})^2]z^2 + [3\frac{r_4}{r_1} - 3\frac{r_2 r_3}{r_1^2} +$$

$$+(\frac{r_2}{r_1})^3]z^3 + [4\frac{r_5}{r_1} - 4\frac{r_2 r_4}{r_1^2} + 4\frac{r_2^2 r_3}{r_1^3} - 2(\frac{r_3}{r_1})^2 +$$

$$-(\frac{r_2}{r_1})^4]z^4 + o(z^5)\} \tag{5.4.2a}$$

$$f_2 = 2\frac{r_2}{r_1}z^{-1}\{1 + (3\frac{r_3}{r_2} - \frac{r_2}{r_1})z + [6\frac{r_4}{r_2} + (\frac{r_2}{r_1})^2 +$$

$$-4\frac{r_3}{r_1}]z^2 + [10\frac{r_5}{r_2} - 7\frac{r_4}{r_1} + 5\frac{r_2 r_3}{r_1^2} - 3\frac{r_3^2}{r_2 r_1} +$$

$$-(\frac{r_2}{r_1})^3]z^3\} + (\sigma\frac{dv}{dz})_1(\sigma\frac{dv}{dz})_1 z^2 + o(z^3) \tag{5.4.2b}$$

$$f_3 = \frac{1}{2}(M\frac{dv}{dz})_0 + [\frac{1}{2}(M\frac{dv}{dz})_1 + (M\frac{dv}{dz})_0]z + [\frac{1}{2}(M\frac{dv}{dz})_2 +$$

$$+(M\frac{dv}{dz})_1 + \frac{1}{2}(M\frac{dv}{dz})_0]z^2 + o(z^3) \tag{5.4.2c}$$

and now (5.2.7a) yields

$$z \equiv z^1 v + z^2 v^2 + \ldots + z^5 v^5 + o(v^6)$$

$$= (\frac{1}{r_1})v - (\frac{r_2}{r_1^3})v^2 + [2\frac{r_2^2}{r_1^5} - \frac{r_3}{r_1^4} - \frac{1}{12 r_1^2}(M\frac{dv}{dz})_0]v^3 +$$

$$+\{5\frac{r_2 r_3}{r_1^6} - \frac{r_4}{r_1^5} - 5\frac{r_2^3}{r_1^7} + \frac{1}{4r_1^3}[(M\frac{dv}{dz})_0(\frac{r_2}{r_1} - \frac{1}{3}) - \frac{1}{6}(M\frac{dv}{dz})_1]\}v^4 +$$

$$+\{6\frac{r_2 r_4}{r_1^7} - \frac{r_5}{r_1^6} - 21\frac{r_2^2 r_3}{r_1^8} + 3\frac{r_3^2}{r_1^7} + 14\frac{r_2^4}{r_1^9} - \frac{1}{40 r_1^4}(M\frac{dv}{dz})_2 +$$

$$+ \frac{1}{120 r_1^4}(3 + \frac{r_2}{r_1})(M\frac{dv}{dz})_1 + \frac{1}{120 r_1^4}(M\frac{dv}{dz})_0[39\frac{r_3}{r_1} + 35\frac{r_2}{r_1} - 98\frac{r_2^2}{r_1^2} +$$

$$-3] + \frac{1}{120 r_1^4}(M\frac{dv}{dz})_0^2 - \frac{1}{20 r_1^5}(\sigma\frac{dv}{dz})_1(\bar{\sigma}\frac{dv}{dz})_1\}v^5 + o(v^6) \tag{5.4.3}$$

Once $z = z(v)$ has been calculated, any quantity

$$a = a_0 + a_1 z + \ldots + a_5 z^5 + o(z^6) \tag{5.4.4a}$$

can be expressed in terms of v as

$$\begin{aligned} a = {} & a^0 + (a_1 z^1)v + \{a_1 z^2 + a_2 (z^1)^2\}v^2 + \{a_1 z^3 + \\ & + 2a_2 z^1 z^2 + a_3 (z^1)^3\}v^3 + \{a_1 z^4 + a_2[(z^2)^2 + 2z^1 z^3] + \\ & + 3a_3 (z^1)^2 z^2 + a_4 (z^1)^4\}v^4 + \{a_1 z^5 + 2a_2[z^2 z^3 + z^1 z^4] + \\ & + a_3[3z^1 (z^2)^2 + 3(z^1)^2 z^3] + 4a_4 (z^1)^3 z^2 + a_5 (z^1)^5\}v^5 + \\ & + o(v^6) \end{aligned} \tag{5.4.4b}$$

where $a^0 = a_0$, and z^i are given by (5.4.3).

Applying (5.4.4) to (5.4.1) gives

$$\begin{aligned} r = {} & v - \frac{1}{12r_1}(M\frac{dv}{dz})_0 v^3 - \frac{1}{24r_1^2}\{(M\frac{dv}{dz})_1 + 2(1 - \frac{r_2}{r_1})(M\frac{dv}{dz})_0\}v^4 + \\ & + \{\frac{1}{120r_1^2}(M\frac{dv}{dz})_0^2 - \frac{1}{20r_1^4}(\sigma\frac{dv}{dz})_1(\bar{\sigma}\frac{dv}{dz})_1 - \frac{1}{40r_1^3}[(M\frac{dv}{dz})_2 + \\ & + (3\frac{r_2}{r_1} - 1)(M\frac{dv}{dz})_1 + (M\frac{dv}{dz})_0(1 + 6(\frac{r_2}{r_1})^2 - 5\frac{r_2}{r_1} - 3\frac{r_3}{r_1})]\}v^5 + \\ & + o(v^6) \end{aligned} \tag{5.4.5a}$$

$$\begin{aligned} \sigma = {} & \frac{1}{r_1^2}(\sigma\frac{dv}{dz})_1 v + \frac{1}{r_1^3}\{(\sigma\frac{dv}{dz})_2 - 3\frac{r_2}{r_1}(\sigma\frac{dv}{dz})_1\}v^2 + \\ & + \frac{1}{r_1^4}\{(\sigma\frac{dv}{dz})_3 - 4\frac{r_2}{r_1}(\sigma\frac{dv}{dz})_2 + [10(\frac{r_2}{r_1})^2 - 4\frac{r_3}{r_1}](\sigma\frac{dv}{dz})_1 + \\ & - \frac{1}{3}r_1(\sigma\frac{dv}{dz})_1(M\frac{dv}{dz})_0\}v^3 \end{aligned} \tag{5.4.5b}$$

$$M = \frac{1}{r_1}(M\frac{dv}{dz})_0 + \frac{1}{r_1^2}\{(M\frac{dv}{dz})_1 - 2\frac{r_2}{r_1}(M\frac{dv}{dz})_0\}v + \frac{1}{r_1^3}\{(M\frac{dv}{dz})_2 +$$

$$- 3\frac{r_2}{r_1}(M\frac{dv}{dz})_1 + 3[2(\frac{r_2}{r_1})^2 - \frac{r_3}{r_1}](M\frac{dv}{dz})_0 - \frac{r_1}{4}(M\frac{dv}{dz})_0^2\}v^2 +$$

$$+ o(v^3) \quad (5.4.5c)$$

$$u^A = u_0^A + (z^1 u_1^A)v + o(v^2) \qquad (5.4.5d)$$

Since z and M are known, it follows from (5.2.2a) that Φ_{00} is determined up to $o(v^2)$:

$$\Phi_{00} = \frac{1}{2}\{M^0 + (M^1 + 2M^0 z^1)v + [M^2 + 2M^1 z^1 + M^0(2z^2 + (z^1)^2]v^2\} +$$
$$+ o(v^3) \qquad (5.4.6)$$

Next, (5.2.10) gives

$$\Psi_0 = (3\sigma^1) + (4\sigma^2)v + (5\sigma^3 - \frac{1}{6}M^0\sigma^1)v^2 + o(v^3) \qquad (5.4.7)$$

In what follows, it is convenient to express the results of the integration in terms of Φ_{00} and Ψ_0, rather than in terms of the observable quantities (5.4.1). Thus, e.g., (5.4.5a) may be written as

$$r = v\{1 - (\frac{1}{6}\Phi_{00}^0)v^2 - (\frac{1}{12}\Phi_{00}^1)v^3 + [\frac{1}{120}(\Phi_{00}^0)^2 +$$
$$- \frac{1}{180}\Psi_0^0\bar{\Psi}_0^0 - \frac{1}{20}\Phi_{00}^2]v^4 + o(v^5)\} \qquad (5.4.8)$$

It is therefore useful to express ρ and σ in terms of Φ_{00} and Ψ_0. This may be done either by using (5.4.5-7) or by integrating (3.3.15a,b). The latter method will be adopted, since it illustrates an integration technique (due to Newman and Unti) that will be used repeatedly in the sequel.

By (3.4.45)

$$\rho = -v^{-1} + g(v), \text{ where } g(v) = o(v),$$

and hence (3.3.15a) implies

$$-v^{-2} + Dg = [-v^{-2} - 2v^{-1}g + g^2] + \Phi_{00}^0 + o(v)$$

so that

$$g = v^{-2}\{\overset{\circ}{g} + \int_{0}^{v}[\overset{0}{\Phi}_{00}v^{2} + o(v^{3})]dv\}\,,\quad \overset{\circ}{g} = \overset{\circ}{g}(w,x^{A})$$
$$= \tfrac{1}{3}\overset{0}{\Phi}_{00}v + o(v^{2}),$$

since $\overset{\circ}{g}$ vanishes by virtue of the assumption $g = o(v)$. Thus

$$\rho = -v^{-1} + \tfrac{1}{3}\overset{0}{\Phi}_{00}v + h(v), \quad \text{where } h(v) = o(v^{2}). \tag{5.4.9}$$

Substituting (5.4.9) into (3.3.15a) then gives

$$D(v^{2}h) = \overset{1}{\Phi}_{00}v^{3}$$

and thus

$$\rho = -v^{-1} + \tfrac{1}{3}\overset{0}{\Phi}_{00}v + \tfrac{1}{4}\overset{1}{\Phi}_{00}v^{2} + k(v), \quad \text{where } k(v) = o(v^{3}) \tag{5.4.10}$$

Before additional terms can be calculated, it is necessary to integrate (3.3.15b). Using the same technique as above, we find

$$\sigma = \tfrac{1}{3}\overset{0}{\Psi}_{0}v + \tfrac{1}{4}\overset{1}{\Psi}_{0}v^{2} + (\overset{2}{\Psi}_{0} + \tfrac{1}{9}\overset{0}{\Phi}_{00}\overset{0}{\Psi}_{0})v^{3} + o(v^{4}). \tag{5.4.11}$$

Now substitution of (5.4.10,11) into (3.3.15a) gives a differential equation for $k(v)$ which is easily solved to yield

$$\rho = -v^{-1} + \tfrac{1}{3}\overset{0}{\Phi}_{00}v + \tfrac{1}{4}\overset{1}{\Phi}_{00}v^{2} + \tfrac{1}{5}[\overset{2}{\Phi}_{00} + \tfrac{1}{9}(\overset{0}{\Phi}_{00})^{2} + \\ + \tfrac{1}{9}\overset{0}{\Psi}_{0}\overset{0}{\bar{\Psi}}_{0}]v^{3} + o(v^{4}). \tag{5.4.12}$$

No further terms can be calculated, since these would involve $\overset{3}{\Phi}_{00}$ and $\overset{3}{\Psi}_{0}$, and hence $(M\frac{dv}{dz})_{3}$ and $(\sigma\frac{dv}{dz})_{4}$, neither of which is known from observations. Thus (5.4.11,12) represent the maximum information that can be extracted from (3.3.15a, b), given the initial data specified in (5.4.1).

Equation (3.3.13a) may now be integrated to give ξ^{A}, and then (3.3.10c) is used to find g^{AB}. We obtain

$$\xi^A = \xi^{AO}v^{-1} + \frac{1}{6}\{\xi^{AO}\Phi^0_{00} + \bar{\xi}^{AO}\Psi^0_0\}v + \frac{1}{12}\{\xi^{AO}\Phi^1_{00} + \bar{\xi}^{AO}\Psi^1_0\}v^2 + \{\frac{1}{20}[\xi^{AO}\Phi^2_{00} + \bar{\xi}^{AO}\Psi^2_0] + \frac{7}{360}\xi^{AO}[(\Phi^0_{00})^2 + \Psi^0_0\bar{\Psi}^0_0] + \frac{1}{30}\bar{\xi}^{AO}\Phi^0_{00}\Psi^0_0\}v^3 + o(v^4) \tag{5.4.13}$$

$$g^{22} = -2P^2v^{-2}\{1 + \frac{1}{3}[\Phi^0_{00}+Re\Psi^0_0]v^2 + \frac{1}{6}[\Phi^1_{00} + Re\Psi^1_0]v^3 + [\frac{1}{10}(\Phi^2_{00} + Re\Psi^2_0) + \frac{1}{15}(\Phi^0_{00})^2 + \frac{1}{15}\Psi^0_0\bar{\Psi}^0_0 + \frac{11}{90}\Phi^0_{00}Re\Psi^0_0]v^4 + o(v^5)\} \tag{5.4.14a}$$

$$g^{23} = 2P^2\{\frac{1}{3}Im\Psi^0_0 + (\frac{1}{6}Im\Psi^1_0)v + (\frac{1}{10}Im\Psi^2_0 + \frac{11}{90}\Phi^0_{00}Im\Psi^0_0)v^2 + o(v^3)\} \tag{5.4.14b}$$

$$g^{33} = -2P^2v^{-2}\{1 + \frac{1}{3}[\Phi^0_{00} - Re\Psi^0_0]v^2 + \frac{1}{6}[\Phi^1_{00} - Re\Psi^1_0]v^3 + [\frac{1}{10}(\Phi^2_{00} - Re\Psi^2_0) + \frac{1}{15}(\Phi^0_{00})^2 + \frac{1}{15}\Psi^0_0\bar{\Psi}^0_0 - \frac{11}{90}\Phi^0_{00}Re\Psi^0_0]v^4 + o(v^5)\} \tag{5.4.14c}$$

where ξ^{AO} and P are defined by (3.4.46b).

Next g_{AB} is calculated using (3.3.8e), which implies

$$g_{22} = (\det g^{AB})^{-1}g^{33}, \quad g_{23} = -(\det g^{AB})^{-1}g^{23},$$

$$g_{33} = (\det g^{AB})^{-1}g^{22}.$$

Now

$$\det g^{AB} = 4P^4v^{-4}\{1 + \frac{2}{3}\Phi^0_{00}v^2 + \frac{1}{3}\Phi^1_{00}v^3 + [\frac{1}{5}\Phi^2_{00} + \frac{11}{45}(\Phi^0_{00})^2 + \frac{1}{45}\Psi^0_0\bar{\Psi}^0_0]v^4 + o(v^5)\} \tag{5.4.15}$$

and hence

$$g_{22} = -\frac{1}{2}P^{-2}v^2\{1 - \frac{1}{3}[\Phi^0_{00} + Re\Psi^0_0]v^2 - \frac{1}{6}[\Phi^1_{00} + Re\Psi^1_0]v^3 + [-\frac{1}{10}(\Phi^2_{00} + Re\Psi^2_0) + \frac{2}{45}(\Phi^0_{00})^2 + \frac{2}{45}\Psi^0_0\bar{\Psi}^0_0 + \frac{1}{10}\Phi^0_{00}Re\Psi^0_0]v^4 + o(v^5)\} \tag{5.4.16a}$$

$$g_{23} = \frac{1}{2}P^{-2}v^4\{\tfrac{1}{3}Im\Psi_0^0 + \tfrac{1}{6}(Im\Psi_0^1)v + \tfrac{1}{10}(Im\Psi_0^2 - \Phi_{00}^0 Im\Psi_0^0)v^2 + $$
$$+ o(v^3)\} \qquad (5.4.16b)$$

$$g_{33} = -\frac{1}{2}P^{-2}v^2\{1 + \tfrac{1}{3}[Re\Psi_0^0 - \Phi_{00}^0]v^2 + \tfrac{1}{6}[Re\Psi_0^1 - \Phi_{00}^1]v^3 +$$
$$+ [\tfrac{1}{10}(Re\Psi_0^2 - \Phi_{00}^2) + \tfrac{2}{45}(\Phi_{00}^0)^2 + \tfrac{2}{45}\Psi_0^0\bar{\Psi}_0^0 +$$
$$- \tfrac{1}{10}\Phi_{00}^0 Re\Psi_0^0]v^4 + o(v^5)\}. \qquad (5.4.16c)$$

The leading terms in the expansions of the Φ_{IJ}'s may now be expressed as follows. Let

$$F \equiv Y^2 + iY^3 = (u^{20} + iu^{30}) + [(u^{21} + X^{21}) +$$
$$+ i(u^{31} + X^{31})]v + o(v^2) \qquad (5.4.17)$$

(cf.(5.2.3)). Then

$$Y^A\xi^B g_{AB} = \frac{v}{2P}[\bar{F}^0 + \bar{F}^1 v + o(v^2)] \qquad (5.4.18a)$$

and

$$Y^A Y^B g_{AB} = -\frac{v^2}{2P^2}[F^0\bar{F}^0 + 2(ReF^0\bar{F}^1)v + o(v^2)]. \qquad (5.4.18b)$$

Thus, by (5.2.2),

$$\Phi_{01} = -\frac{v}{4P}\{M^0\bar{F}^0 + [M^0\bar{F}^1 + (M^1 + M^0 z^1)\bar{F}^0]v + o(v^2)\} \qquad (5.4.19a)$$

$$\Phi_{02} = \frac{v^2}{8P^2}\{M^0(\bar{F}^0)^2 + [M^1(\bar{F}^0)^2 + 2M^0\bar{F}^0\bar{F}^1]v + o(v^2)\} \qquad (5.4.19b)$$

$$\Phi_{11} = \frac{1}{8}\{M^0 + M^1 v + [M^2 + P^{-2}M^0F^0\bar{F}^0]v^2 + o(v^3)\} \qquad (5.4.19c)$$

$$\Phi_{12} = -\frac{v}{8P}\{M^0\bar{F}^0 + [M^0\bar{F}^1 + (M^1 - M^0 z^1)\bar{F}^0]v + o(v^2)\} \qquad (5.4.19d)$$

$$\Phi_{22} = \frac{1}{8}\{M^0 + [M^1 - 2M^0 z^1]v + [M^2 - 2M^1 z^1 + 3M^0(z^1)^2 + \qquad (5.4.19e)$$
$$- 2z^2M^0 + P^{-2}M^0F^0\bar{F}^0]v^2 + o(v^3)\} \qquad (5.4.19e)$$

(Recall that Φ_{00} is given by (5.4.6).)

At this stage of the integration, only F^0 is known; thus Φ_{01}, Φ_{02}, Φ_{11}, Φ_{12} and Φ_{22} are known up to orders v, v^2, v^2, v

and v^2, respectively. This allows one to integrate sequentially (3.3.15d), (3.3.17a), (3.3.15c), (3.3.13c) and (3.3.13b) up to orders v^2, v, v^2, v^2 and v^3, respectively.

By (3.3.11), equation (3.3.15d) may be written as

$$D\alpha = \rho\alpha + (\tau - \bar{\alpha})\bar{\sigma} + \bar{\Phi}_{01} \ . \tag{5.4.20}$$

But, by (3.4.45),

$$\alpha = \alpha^0 v^{-1} + g(v), \quad \text{where } g(v) = o(v),$$

and thus (5.4.20) gives

$$Dg + v^{-1}g = \frac{1}{3}(\alpha^0\Phi^0_{00} - \bar{\alpha}^0\bar{\Psi}^0_0)$$

so that

$$\alpha = \alpha^0 v^{-1} + \frac{1}{6}(\alpha^0\Phi^0_{00} - \bar{\alpha}^0\bar{\Psi}^0_0)v + [\frac{1}{3}\bar{\Phi}^1_{01} +$$
$$+ \frac{1}{12}(\alpha^0\Phi^1_{00} - \bar{\alpha}^0\Psi^1_0)]v^2 + o(v^3) \tag{5.4.21}$$

This is as far as we can proceed with (5.4.20) at the moment, since the term in $o(v^3)$ involves $\bar{\Phi}^2_{01}$, which requires knowledge of $\bar{F}^1$ and hence (by (5.4.17)) of X^{A1}. But in order to determine X^{A1}, τ^1 is required, and this presupposes knowledge of Ψ^0_1. It is therefore clear why the equations need to be integrated in precisely the sequence described above. It will also be seen that the integration scheme used here is simpler than the general algorithm described in §5.2. This is possible only because (i) the data is given in the polynomial form (5.4.1), so that the integration can proceed term by term and (ii) the central conditions imply $X^A = o(v)$. If, instead, $X^A = o(1)$, one could not effectively decouple the equations in this way, since F^0 and hence Φ^1_{01} would not be known until (3.3.13c) had been solved.

The details of the integration to determine Ψ_1, τ, X^A and ω will not be given; the procedure is the same as that used to find α. The relations (3.4.55) are used frequently (the spin-weights of all quantities of interest are listed in Appendix II). One obtains

$$\Psi_1 = -\left(\frac{1}{4\sqrt{2}}\bar{\eth}\Psi_0^0\right) + \frac{1}{5}\left(3\Phi_{01}^1 - \frac{1}{\sqrt{2}}\bar{\eth}\Psi_0^1 + \frac{1}{\sqrt{2}}\eth\Phi_{00}^1\right)v + o(v^2) \qquad (5.4.22a)$$

$$\tau = -\left(\frac{1}{8\sqrt{2}}\bar{\eth}\Psi_0^0\right)v + \frac{1}{15}\left(8\Phi_{01}^1 - \frac{1}{\sqrt{2}}\bar{\eth}\Psi_0^1 + \frac{1}{\sqrt{2}}\eth\Phi_{00}^1\right)v^2 + o(v^3) \qquad (5.4.22b)$$

$$X^A = -\left[\frac{1}{4\sqrt{2}}Re(\xi^{AO}\eth\bar{\Psi}_0^0)\right]v + \frac{1}{15}Re\Big\{\xi^{AO}\Big[8\bar{\Phi}_{01}^1 + \frac{1}{\sqrt{2}}\left(\bar{\eth}\Phi_{00}^1 - \eth\bar{\Psi}_0^1\right)\Big]\Big\}v^2 + o(v^3) \qquad (5.4.22c)$$

$$\omega = \left(\frac{1}{24\sqrt{2}}\bar{\eth}\Psi_0^0\right)v^2 - \frac{1}{60}\left[8\Phi_{01}^1 + \frac{1}{\sqrt{2}}\left(\eth\Phi_{00}^1 - \bar{\eth}\Psi_0^1\right)\right]v^3 + o(v^4) \qquad (5.4.22d)$$

Once X^{A1} has been calculated, it follows from (5.4.17) that Φ_{01}^2 is determined and thus the term in $o(v^3)$ in the expansion of α may be obtained. Then (5.4.22) is substituted back into (3.3.17a),(3.3.15c),(3.3.13c) and (3.3.13d), and one more term in each expansion is calculated (further terms require that the initial data be specified to higher order than that given in (5.4.1)). The remaining equations are integrated analogously.

From (5.4.17,22c) it follows that

$$F^1 = Q + \frac{1}{4\sqrt{2}}\eth\Psi_0^0 \qquad (5.4.23a)$$

where

$$Q \equiv u^{21} + iu^{31} \qquad (5.4.23b)$$

The results of the integration may now be expressed as follows:

A. Ricci tensor components:

$$\Phi_{00} = \frac{1}{2}\{M^0 + (M^1 + 2M^0 z^1)v + [M^2 + 2M^1 z^1 + \quad (5.4.24a)$$
$$+ M^0(2z^2 + (z^1)^2)]v^2 + o(v^3)\}$$

$$\Phi_{01} = -\frac{v}{4P}\{M^0\bar{F}^0 + [M^0(\bar{Q} + \bar{F}^0 z^1 + \frac{1}{4\sqrt{2}}P\bar{\eth}\Psi_0^0) + M^1\bar{F}^0]v +$$
$$+ o(v^2)\} \quad (5.4.24b)$$

$$\Phi_{02} = \frac{v^2}{8P^2}\{M^0(\bar{F}^0)^2 + [M^1(\bar{F}^0)^2 + 2M^0\bar{F}^0(\bar{Q} + \frac{1}{4\sqrt{2}}P\bar{\eth}\Psi_0^0)]v +$$
$$+ o(v^2)\} \quad (5.4.24c)$$

$$\Phi_{11} = \frac{1}{8}\{M^0 + M^1 v + [M^2 + P^{-2}M^0 F^0\bar{F}^0]v^2 + o(v^3)\} \quad (5.4.24d)$$

$$\Phi_{12} = -\frac{v}{8P}\{M^0\bar{F}^0 + [M^0(\bar{Q} + \frac{1}{4\sqrt{2}}P\bar{\eth}\Psi_0^0 - \bar{F}^0 z^1) + M^1\bar{F}^0]v +$$
$$+ o(v^2)\} \quad (5.4.24e)$$

$$\Phi_{22} = \frac{1}{8}\{M^0 + [M^1 - 2M^0 z^1]v + [M^2 - 2M^1 z^1 +$$
$$+ M^0(3(z^1)^2 - 2z^2 + P^{-2}F^0\bar{F}^0)]v^2 + o(v^3)\} \quad (5.4.24f)$$

$$\Lambda = \frac{1}{24}\{M^0 + M^1 v + M^2 v^2 + o(v^3)\} \quad (5.4.24g)$$

where M^0, M^1, M^2, F^0, F^1, z^1 and z^2 are given in terms of observable quantities by (5.4.3,5,23).

B. Weyl tensor components:

$$\Psi_0 = \Psi_0^0 + \Psi_0^1 v + \Psi_0^2 v^2 + o(v^3) \quad (5.4.25a)$$

$$\Psi_1 = -\frac{1}{4\sqrt{2}}\bar{\eth}\Psi_0^0 + \frac{1}{5}(3\Phi_{01}^1 - \frac{1}{\sqrt{2}}\bar{\eth}\Psi_0^1 + \frac{1}{\sqrt{2}}\eth\Phi_{00}^1)v +$$
$$+ \frac{1}{6}(4\Phi_{01}^2 - \frac{1}{\sqrt{2}}\bar{\eth}\Psi_0^2 + \frac{1}{\sqrt{2}}\eth\Phi_{00}^2 - \frac{3}{4\sqrt{2}}\Phi_{00}^0\bar{\eth}\Psi_0^0 +$$

$$- \frac{1}{6\sqrt{2}}\bar{\Psi}^0_0 \eth \Psi^0_0) v^2 + o(v^3) \tag{5.4.25b}$$

$$\begin{aligned}
\Psi_2 = {} & \frac{1}{24}\bar{\eth}^2 \Psi^0_0 + \frac{1}{4}[\frac{1}{\sqrt{2}}(\eth \bar{\Phi}^1_{01} - \frac{3}{5}\bar{\eth}\Phi^1_{01}) + \frac{1}{10}(\bar{\eth}^2 \Psi^1_0 - \bar{\eth}\eth\Phi^1_{00}) + \\
& + \frac{1}{6}M^1 - \frac{1}{2}M^0 z^1]v + \frac{1}{5}\{\frac{1}{\sqrt{2}}(\eth\bar{\Phi}^2_{01} - \frac{2}{3}\bar{\eth}\Phi^2_{01}) + \frac{1}{12}(\bar{\eth}^2\Psi^2_0 + \\
& - \bar{\eth}\eth\Phi^2_{00}) + \frac{7}{48}\Phi^0_{00}\bar{\eth}^2\Psi^0_0 + \frac{1}{48}\bar{\Psi}^0_0(\eth\bar{\eth}\Psi^0_0 + 4\Psi^0_0) + \frac{1}{72}\bar{\Psi}^0_0\bar{\eth}\eth\Psi^0_0 + \\
& + \frac{1}{72}(\bar{\eth}\Psi^0_0)(\eth\Psi^0_0) + \frac{1}{4}M^2 - \frac{1}{2}M^1 z^1 + \frac{1}{2}M^0[z^2 + (z^1)^2 + \\
& + P^{-2}F^0\bar{F}^0]\}v^2 + o(v^3)
\end{aligned} \tag{5.4.25c}$$

$$\begin{aligned}
\Psi_3 = {} & - \frac{1}{48\sqrt{2}}\bar{\eth}^3\Psi^0_0 + \frac{1}{3}[- \frac{1}{\sqrt{2}}\bar{\eth}(\Psi^1_2 + \frac{1}{12}M^1) + 3\Phi^1_{21} - \bar{\Phi}^1_{01}]v + \\
& + \frac{1}{4}[- \frac{1}{\sqrt{2}}\bar{\eth}(\Psi^2_2 + \frac{1}{12}M^2) + 4\Phi^1_{21} - \bar{\Phi}^2_{01} - \frac{1}{6\sqrt{2}}\Phi^0_{00}\bar{\eth}\Psi^0_2 + \\
& - \frac{1}{6\sqrt{2}}\bar{\Psi}^0_0\eth\Psi^0_2 + \frac{1}{\sqrt{2}}\eth\Phi^2_{20} + \frac{1}{6}\bar{\Psi}^0_0\Psi^0_1 + \frac{2}{3}\Phi^0_{00}\Psi^0_3] v^2 + o(v^3)
\end{aligned} \tag{5.4.25d}$$

$$\begin{aligned}
\Psi_4 = {} & \frac{1}{96}\bar{\eth}^4\Psi^0_0 - \frac{1}{2\sqrt{2}}\bar{\eth}(\Psi^1_3 + \Phi^1_{21})v + \frac{1}{3}[- \frac{1}{\sqrt{2}}\bar{\eth}(\Psi^2_3 + \Phi^2_{21}) + \\
& - \frac{1}{6\sqrt{2}}\Phi^0_{00}\bar{\eth}\Psi^0_3 - \frac{1}{6\sqrt{2}}\bar{\Psi}^0_0\eth\Psi^0_3 + \frac{1}{3}(\Phi^0_{00}\Psi^0_4) + \frac{1}{4}\bar{\Psi}^0_0\Psi^0_2 + \\
& + \frac{1}{8}\bar{\Psi}^0_0\Phi^0_{00} + \frac{1}{2}\Phi^0_{20} - (\Phi^0_{00})^{-1}(\bar{\Phi}^1_{01})^2]v^2 + o(v^3)
\end{aligned} \tag{5.4.25e}$$

where Ψ^n_0 (n=0,1,2) are given in terms of observable quantities by (5.4.5b,7).

It will be noticed that not all the expressions given above have been reduced to their most primitive form. Ψ^1_3, e.g., contains the term $\frac{1}{\sqrt{2}}\bar{\eth}(\Psi^1_2 + \frac{1}{12}M^1)$; by substituting for Ψ^1_2 from (5.4.25b) and for Φ^1_{01} from (5.4.24), this may be expressed in terms of M^0, M^1, F^0, z^1 and angular derivatives of these quantities. The reason for not undertaking this kind of simplification in all cases is simply that the expressions soon become completely unmanageable (Ψ_4^2 alone, if completely reduced, runs into several pages). The basic idea underlying the integration is that the equations should be integrated sequentially; the solution is given here in the same way: in order to determine Ψ_4, say, from (5.4.25e), it is necessary first to calculate Ψ_1, Ψ_2 and Ψ_3 from (5.4.25b-d).

Similar remarks hold for the rest of the solution given below; the reduction will be carried out only when this does not vastly increase the complexity of the expressions involved.

C. Spin coefficients:

$$\rho = -v^{-1} + \tfrac{1}{3}\Phi^0_{00}v + \tfrac{1}{4}\Phi^1_{00}v^2 + \tfrac{1}{5}[\Phi^2_{00} + \tfrac{1}{9}(\Phi^0_{00})^2 + + \tfrac{1}{9}\Psi^0_0\bar{\Psi}^0_0]v^3 + o(v^4) \qquad (5.4.26a)$$

$$\sigma = \tfrac{1}{3}\Psi^0_0 v + \tfrac{1}{4}\Psi^1_0 v^2 + \tfrac{1}{5}[\Psi^2_0 + \tfrac{1}{9}\Psi^0_0\Phi^0_{00}]v^3 + o(v^4) \qquad (5.4.26b)$$

$$\alpha = \alpha^0 v^{-1} + \tfrac{1}{6}[\alpha^0\Phi^0_{00} - \bar{\alpha}^0\bar{\Psi}^0_0]v + \tfrac{1}{12}[4\bar{\Phi}^1_{01} + \alpha^0\Phi^1_{00} + -\bar{\alpha}^0\bar{\Psi}^1_0]v^2 + \{\tfrac{1}{20}[\alpha^0\Phi^2_{00} - \bar{\alpha}^0\bar{\Psi}^2_0] + \tfrac{7}{360}\alpha^0[(\Phi^0_{00})^2 + + \Psi^0_0\bar{\Psi}^0_0] - \tfrac{1}{30}\bar{\alpha}^0\Phi^0_{00}\Psi^0_0 - \tfrac{1}{96\sqrt{2}}\bar{\Psi}^0_0\eth\bar{\Psi}^0_0 + \tfrac{1}{4}\bar{\Phi}^2_{01}\}v^3 + + o(v^4) \qquad (5.4.26c)$$

$$\tau = -(\tfrac{1}{8\sqrt{2}}\bar{\eth}\Psi^0_0)v + \tfrac{1}{15}[8\Phi^1_{01} + \tfrac{1}{\sqrt{2}}(\eth\Phi^1_{00} - \bar{\eth}\Psi^1_0)]v^2 + + \tfrac{1}{24\sqrt{2}}[10\sqrt{2}\Phi^2_{01} + (\eth\Phi^2_{00} - \bar{\eth}\Psi^2_0) - \Phi^0_{00}\bar{\eth}\Psi^0_0 - \tfrac{1}{4}\Psi^0_0\eth\bar{\Psi}^0_0 + - \tfrac{1}{6}\bar{\Psi}^0_0\eth\Psi^0_0]v^3 + o(v^4) \qquad (5.4.26d)$$

$$
\begin{aligned}
\mu = & -\frac{1}{2}v^{-1} + \left(\frac{1}{48}\bar{\eth}^2\Psi_0^0\right)v + \frac{1}{12}\Big[\frac{1}{\sqrt{2}}\left(\eth\bar{\Phi}_{01}^1 - \frac{3}{5}\bar{\eth}\Phi_{01}^1\right) + \\
& + \frac{1}{10}\left(\bar{\eth}^2\Psi_0^1 - \bar{\eth}\eth\Phi_{00}^1\right) + \frac{1}{4}M^1 - M^0 z^1\Big]v^2 + \\
& + \frac{1}{4}\Big\{\frac{1}{144}\Phi_{00}^0\bar{\eth}^2\Psi_0^0 + \frac{1}{12}M^2 + \Psi_2^2 - \frac{1}{10}\Big[\Phi_{00}^2 + \frac{1}{9}(\Phi_{00}^0)^2 + \\
& + \frac{7}{18}\Psi_0^0\bar{\Psi}_0^0\Big]\Big\}v^3 + o(v^4) \qquad (5.4.26e)
\end{aligned}
$$

$$
\begin{aligned}
\lambda = & -\frac{1}{12}\bar{\Psi}_0^0 v - \frac{1}{24}\bar{\Psi}_0^1 v^2 + \frac{1}{4}\Big[\bar{\Phi}_{02}^2 - \frac{1}{10}\bar{\Psi}_0^2 + \frac{7}{180}\Phi_{00}^0\bar{\Psi}_0^0 + \\
& + \frac{1}{144}\bar{\Psi}_0^0\bar{\eth}^2\Psi_0^0\Big]v^3 + o(v^4) \qquad (5.4.26f)
\end{aligned}
$$

$$
\begin{aligned}
\gamma = & \Big[-\frac{i}{4\sqrt{2}}\mathit{Im}(\alpha^0\bar{\eth}\Psi_0^0) + \frac{1}{24}\bar{\eth}^2\Psi_0^0 + \frac{1}{12}M^0\Big]v + \\
& + \Big\{\frac{i}{15\sqrt{2}}\mathit{Im}(8\sqrt{2}\Phi_{01}^1 + \eth\Phi_{00}^1 - \bar{\eth}\Psi_0^1) + \frac{1}{8}\Big[\frac{1}{\sqrt{2}}\Big(\eth\bar{\Phi}_{01}^1 + \\
& - \frac{3}{5}\bar{\eth}\Phi_{01}^1\Big) + \frac{1}{10}(\bar{\eth}^2\Psi_0^1 - \bar{\eth}\eth\Phi_{00}^1) + \frac{1}{8}(M^1 - M^0 z^1)\Big]\Big\}v^2 + \\
& + \frac{1}{3}\Bigg(i\mathit{Im}\Big\{\alpha^0\Big[\frac{5}{6}\Phi_{01}^2 + \frac{1}{12\sqrt{2}}(\eth\Phi_{00}^2 - \bar{\eth}\Psi_0^2) - \frac{1}{8\sqrt{2}}\Phi_{00}^0\bar{\eth}\Psi_0^0 + \\
& - \frac{1}{16\sqrt{2}}\Psi_0^0\eth\bar{\Psi}_0^0 - \frac{1}{144}\bar{\Psi}_0^0\eth\Psi_0^0\Big]\Big\} + \frac{1}{128}(\bar{\eth}\Psi_0^0)^2 + \frac{1}{12}M^2 + \\
& + \frac{1}{8}P^{-2}M^0F^0\bar{F}^0\Bigg)v^3 + o(v^4) \qquad (5.4.26g)
\end{aligned}
$$

$$
\begin{aligned}
\nu = & \frac{1}{48\sqrt{2}}(3\eth\bar{\Psi}_0^0 - \bar{\eth}^3\Psi_0^0)v + \Big[\frac{1}{5}\bar{\Phi}_{01}^1 - \frac{1}{60\sqrt{2}}(\bar{\eth}\Phi_{00}^1 - \eth\bar{\Psi}_0^1) + \\
& - \frac{1}{6\sqrt{2}}\bar{\eth}\Big(\Psi_2^1 + \frac{1}{12}M^1\Big)\Big]v^2 + \frac{1}{3\sqrt{2}}\Big[\frac{1}{48}\Phi_{00}^0\Big(\eth\bar{\Psi}_0^0 - \frac{1}{4}\bar{\eth}^3\Psi_0^0\Big) + \\
& - \frac{1}{48}(\bar{\eth}\Phi_{00}^2 - \eth\bar{\Psi}_0^2) - \frac{1}{64}\bar{\Psi}_0^0\Big(\bar{\eth}\Psi_0^0 + \frac{1}{9}\eth\bar{\eth}^2\Psi_0^0\Big) + \frac{1}{288}\Psi_0^0\bar{\eth}\bar{\Psi}_0^0 + \\
& - \frac{1}{384}(\eth\bar{\Psi}_0^0)(\bar{\eth}^2\Psi_0^0) - \frac{1}{4}\bar{\eth}\Big(\Psi_2^2 + \frac{1}{12}M^2\Big) + 2\bar{\Phi}_{12}^1 - \frac{11}{24}\bar{\Phi}_{01}^2 + \\
& + \frac{1}{4}\eth\bar{\Phi}_{02}^2\Big]v^3 + o(v^4) \qquad (5.4.26h)
\end{aligned}
$$

where α^0 is defined by (3.4.45b).

The spin coefficients not listed above are given by

$$\kappa = \varepsilon = \pi = 0, \qquad \beta = \tau - \bar{\alpha} \tag{5.4.26i}$$

D. Metric variables:

$$\begin{aligned} U = &-\frac{1}{2} - \frac{1}{12}[M^0 + \frac{1}{4}(\bar{\eth}^2\Psi^0_0 + \eth^2\bar{\Psi}^0_0)]v^2 - \frac{1}{6}[\frac{1}{10\sqrt{2}}(\eth\bar{\Phi}^1_{01} + \\ &+ \bar{\eth}\Phi^1_{01}) + \frac{1}{10}(\bar{\eth}^2\Psi^1_0 + \eth^2\Psi^1_0) - \frac{1}{5}\eth\bar{\eth}\Phi^1_{00} + \frac{1}{2}M^1 - M^0 z^1]v^3 + \\ &- \frac{1}{4}[\frac{1}{128}(\bar{\eth}\Psi^0_0)(\eth\bar{\Psi}^0_0) + \frac{1}{3}(\Psi^2_2 + \bar{\Psi}^2_2) + \frac{1}{9}M^2 + \\ &+ \frac{1}{12}P^{-2}M^0F^0\bar{F}^0]v^4 + o(v^5) \end{aligned} \tag{5.4.27a}$$

$$\begin{aligned} \omega = &\frac{1}{24\sqrt{2}}(\bar{\eth}\Psi^0_0)v^2 + \frac{1}{60}[2M^0\frac{\bar{F}^0}{P} + \frac{1}{\sqrt{2}}(\bar{\eth}\Psi^1_0 - \eth\Phi^1_{00})]v^3 + \\ &+ \frac{1}{240}[5M^0\frac{\bar{Q}}{P} + 5(M^1 + M^0z^1)\frac{\bar{F}^0}{P} + \frac{31}{6\sqrt{2}}\Phi^0_{00}\bar{\eth}\Psi^0_0 + \frac{7}{6\sqrt{2}}\Psi^0_0\eth\bar{\Psi}^0_0 + \\ &+ \frac{1}{3}\bar{\Psi}^0_0\eth\Psi^0_0 + \sqrt{2}(\bar{\eth}\Psi^2_0 - \eth\Phi^2_{00})]v^4 + o(v^5) \end{aligned} \tag{5.4.27b}$$

$$\begin{aligned} X^A = &-\frac{1}{4\sqrt{2}}Re(\xi^{A0}\eth\bar{\Psi}_0)v - \frac{1}{15}Re\{\xi^{A0}[\frac{1}{\sqrt{2}}(\eth\bar{\Psi}^1_0 - \bar{\eth}\Phi^1_{00}) + \\ &+ 2M^0\frac{F^0}{P}]\}v^2 - \frac{1}{72}Re\{\xi^{A0}[5M^0\frac{Q}{P} + 5(M^1 + M^0z^1)\frac{F^0}{P} + \\ &+ \sqrt{2}(\eth\bar{\Psi}^2_0 - \bar{\eth}\Phi^2_{00}) + \frac{11}{2\sqrt{2}}\Phi^0_{00}\eth\bar{\Psi}^0_0 + \frac{3}{2\sqrt{2}}\bar{\Psi}^0_0\bar{\eth}\Psi^0_0 + \\ &+ \frac{1}{3}\Psi^0_0\bar{\eth}\bar{\Psi}^0_0]\}v^3 + o(v^4) \end{aligned} \tag{5.4.27c}$$

$$\begin{aligned} \xi^A = &\xi^{A0}v^{-1} + \frac{1}{6}(\xi^{A0}\Phi^0_{00} + \bar{\xi}^{A0}\Psi^0_0)v + \frac{1}{12}(\xi^{A0}\Phi^1_{00} + \\ &+ \bar{\xi}^{A0}\Psi^1_0)v^2 + \{\frac{1}{20}(\xi^{A0}\Phi^2_{00} + \bar{\xi}^{A0}\Psi^2_0) + \\ &+ \frac{7}{360}\xi^{A0}[(\Phi^0_{00})^2 + \Psi^0_0\bar{\Psi}^0_0] + \frac{1}{30}\bar{\xi}^{A0}\Phi^0_{00}\Psi^0_0\}v^3 + o(v^4) \end{aligned} \tag{5.4.27d}$$

where ξ^{A0} is defined by (3.4.46b).

E. Metric tensor components:

$$g^{11} = -1 - \frac{1}{6}[M^0 + \frac{1}{4}(\bar{\eth}^2\Psi_0^0 + \eth^2\bar{\Psi}_0^0)]v^2 - \frac{1}{3}[\frac{1}{10\sqrt{2}}(\eth\bar{\Phi}_{01}^1 + \bar{\eth}\Phi_{01}^1) + \frac{1}{10}(\bar{\eth}^2\Psi_0^1 + \eth^2\bar{\Psi}_0^1) - \frac{1}{5}\eth\bar{\eth}\Phi_{00}^1 + \frac{1}{2}M^1 + - M^0z^1]v^3 - \frac{1}{2}[\frac{1}{9}M^2 + \frac{1}{12}P^{-2}M^0F^0\bar{F}^0 + \frac{1}{3}(\Psi_2^2 + \bar{\Psi}_2^2) + + \frac{13}{2304}(\bar{\eth}\Psi_0^0)(\eth\bar{\Psi}_0^0)]v^4 + o(v^5) \tag{5.4.28a}$$

$$g^{1A} = -\frac{1}{3\sqrt{2}}Re(\xi^{A0}\eth\bar{\Psi}_0^0)v - \frac{1}{5}Re\{\xi^{A0}[\frac{1}{2\sqrt{2}}(\eth\bar{\Psi}_0^1 - \eth\Phi_{00}^1) + + M^0\frac{F^0}{P}]\}v^2 - \frac{1}{15}Re\{\xi^{A0}[\frac{2}{3\sqrt{2}}(\eth\bar{\Psi}_0^2 - \eth\Phi_{00}^2) + + 2\Phi_{00}^0\eth\bar{\Psi}_0^0 + \frac{1}{4}\bar{\Psi}_0^0\bar{\eth}\Psi_0^0 + \frac{1}{9}\Psi_0^0\bar{\eth}\bar{\Psi}_0^0 + \frac{5}{3}M^0\frac{Q}{P} + + \frac{5}{3}(M^1 + M^0z^1)\frac{F^0}{P}]\}v^3 + o(v^4) \tag{5.4.28b}$$

The remaining metric tensor components g^{AB} are given by (5.4.14).

§5.5. *Angular dependence of the initial data*

In this section we shall derive the explicit angular behaviour of the expansion coefficients Ψ_0^0, z^1, M^1 and F^0. In addition, it will be shown how similar expressions can be derived for higher order expansion coefficients.

The explicit angular dependence of Ψ_0^0 may be found by examining the leading terms of (3.3.17,18). From (3.3.17) it follows that

$$\Psi_Q^0 = [\sqrt{2}(Q-5)]^{-1}\,\bar{\eth}\Psi_{Q-1}^0 \quad , \quad (Q=1,2,3,4) \tag{5.5.1a}$$

(cf. the leading terms in (5.4.25)), while (3.3.18) implies

$$\eth\Psi_Q^0 = \frac{1}{\sqrt{2}}Q\Psi_{Q-1}^0. \tag{5.5.1b}$$

In particular, for the case $Q = 1$, (5.5.1) gives

$$\eth\bar{\eth}\Psi_0^0 + 4\Psi_0^0 = 0. \tag{5.5.2}$$

This differential equation for Ψ_0^0 is easily solved using (3.4.56) with $s = 2$:

$$\Psi_0^0 = a^m\,{}_2Y_{2m} \;, \quad (m=-2,-1,0,1,2) \tag{5.5.3}$$

where the a^m are complex functions only of w.

An important consequence of (5.5.2) which simplifies considerably some of the expressions in (5.4.24-28), is the fact that, by (3.4.56c) and (5.5.3),

$$\eth\Psi_0^0 = 0. \tag{5.5.4}$$

This can also be derived directly from (5.5.2): since ([74])

$$(\bar{\eth}\eth - \eth\bar{\eth})\eta = 2s\eta \tag{5.5.5}$$

for any spin-s quantity η, it follows from (5.5.2) that $\bar{\eth}\eth\Psi_0^0 = 0$. Now, if η has spin weight $s > 0$, then ([73]) $\bar{\eth}\eta = 0$ implies $\eta = 0$. But Ψ_0^0 has $s = 2$ and thus $\eth\Psi_0^0$ is a spin-3 quantity; hence (5.5.4) holds.

The physical significance of the 5 complex functions a^m is established by comparing (3.2.4a) and (5.5.3): they are simply linear combinations of the 10 (real) independent orthonormal tetrad components of the Weyl tensor C_{abcd} on C.

In order to find the angular dependence of z , we shall use various properties of the fluid kinematic quantities (see, e.g., [31]). It is well-known that for a pressure-free perfect fluid, the acceleration $u_{a;b}u^b$ vanishes, so that

$$u_{a;b} = h^c{}_a h^d{}_b u_{c;d} \tag{5.5.6a}$$

$$= \theta_{ab} + \omega_{ab} , \tag{5.5.6b}$$

where

$$h_{ab} \equiv g_{ab} - u_a u_b \tag{5.5.6c}$$

projects into the instantaneous rest space of $\underline{u}$, while

$$\theta_{ab} \equiv u_{(a;b)} , \tag{5.5.6d}$$

$$\omega_{ab} \equiv u_{[a;b]} \tag{5.5.6e}$$

are the fluid expansion and vorticity tensors, respectively.

The expansion tensor may be split further as

$$\theta_{ab} = \frac{1}{3}\theta h_{ab} + \sigma_{ab} \tag{5.5.6f}$$

where

$$\theta = \theta_{ab} g^{ab} , \tag{5.5.6g}$$

and σ_{ab} is the fluid shear tensor.

An important property of all these quantities is the fact that

$$\theta_{ab} u^b = \sigma_{ab} u^b = \omega_{ab} u^b = 0 , \tag{5.5.6h}$$

as is obvious from (5.5.6a) and the fact that $h_{ab}u^b = 0$.

Now (4.2.3b) implies

$$D(1+z) = -D(u^a k_a) = -u_{a;b}k^a k^b$$

$$= -[\tfrac{1}{3}\theta(g_{ab} - u_a u_b) + \sigma_{ab}]k^a k^b$$

$$= \tfrac{1}{3}\theta(1+z)^2 - \sigma_{ab}k^a k^b. \tag{5.5.7}$$

But

$$D(1+z) = z^1 + 2z^2 v + o(v^2),$$

and hence

$$[D(1+z)]_C = z^1. \tag{5.5.8a}$$

The trick is now to evaluate the right hand side of (5.5.7) by using the orthonormal basis $\{\underline{E}_a\}$ defined in §3.4. Note from (3.4.35a,36a) and (3.3.1) that

$$\underline{k}|_C = -\underline{E}_o|_C + (k^i \underline{E}_i)_C \qquad (i=1,2,3)$$

$$= -\underline{u}|_C + (k^i \underline{E}_i)_C \tag{5.5.8b}$$

so that, by (5.5.6b,7,8),

$$z^1 = \tfrac{1}{3}\theta|_C - (\sigma_{ij})_C (k^i k^j)_C$$

where σ_{ij} are the orthonormal tetrad components of the fluid shear. Since the tetrad $\{\underline{E}_a\}$ is perfectly well-behaved on C and C is regular, these components are functions only of proper time w on C. Thus (dropping the subscript $|_C$).

$$z^1 = \tfrac{1}{3}\theta(w) - \sigma_{ij}(w)k^i k^j, \tag{5.5.9}$$

where k^i is given by (3.4.59b).

Similarly the proper motions at C, represented by $F^0 \equiv u^{20} + iu^{30}$, are found by evaluating

$$D(Y^A \xi^B g_{AB}) = D(u^a m_a) = (\tfrac{1}{3}\theta h_{ab} + \sigma_{ab} + \omega_{ab})m^a k^b. \tag{5.5.10}$$

Using (5.4.18a), we find

$$\frac{\bar{F}^0}{P} = 2(\sigma_{ij} + \omega_{ij})m^i k^j \tag{5.5.11}$$

where, as before, the terms on the right hand side are understood to be evaluated on C, and m^i is defined by (3.4.59a).

In order to find M^1, we evaluate $DM = M_{;a}k^a$ on C:

$$M^1 = -\dot{M}^0 + (M_{;i})k^i \tag{5.5.12}$$

where "." denotes $\frac{\partial}{\partial w}$ and $(M_{;i})$ are the spatial, orthonormal tetrad derivatives of M on C.

By using (3.4.58)

$$\eth m^i = 0, \qquad \eth k^i = -\sqrt{2}m^i \ , \tag{3.4.58}$$

which implies

$$\bar{\eth} m^i = \sqrt{2}k^i, \tag{5.5.13}$$

as many angular derivatives of (5.5.9,11,12) as required are easily calculated. These may be used to simplify (5.4.25-28) if desired.

The angular dependence of higher order expansion coefficients may be derived similarly. For example, differentiation of (5.5.7) gives

$$z^2 = \frac{1}{3}\theta z^1 + \frac{1}{6}(\theta_{;a})(k^a) - \frac{1}{2}(\sigma_{ab;c}k^a k^b k^c), \tag{5.5.14a}$$

from which it follows that

$$z^2 = z^2(w) + z^2_i(w)k^i + z^2_{ij}(w)k^i k^j + z^2_{il\ell}(w)k^i k^j k^\ell. \tag{5.5.14b}$$

Similarly, it can be shown that

$$\frac{\bar{F}^1}{P} = (\frac{\bar{F}^1}{P})_{ij}(w)m^i k^j + (\frac{\bar{F}^1}{P})_{ij\ell}(w)m^i k^j k^\ell \tag{5.5.15}$$

and

$$M^2 = M^2(w) + M^2_i(w)k^i + M^2_{ij}(w)k^i k^j \tag{5.5.16}$$

In order to find Ψ_0^1, it is necessary to examine the next term in the expansion of (3.3.18a). This gives

$$\dot{\Psi}_0^0 + \Psi_0^1 = \frac{1}{\sqrt{2}}\eth(\Phi_{01}^1 + \Psi_1^1)$$

$$= \frac{1}{\sqrt{2}}\eth[\frac{8}{5}\Phi_{01}^1 + \frac{1}{5\sqrt{2}}(\eth\Phi_{00}^1 - \bar{\eth}\Psi_0^1)] \tag{5.5.17}$$

where we have substituted for Ψ_1^1 from (5.4.25b).

Now

$$\eth\Phi_{01} = -\frac{1}{4}M^0\eth(\frac{\bar{F}^0}{P}) = \frac{1}{\sqrt{2}}M^0\sigma_{ij}m^i m^j$$

while

$$\eth^2\Phi_{00}^1 = \frac{1}{2}\eth^2 M^1 + M^0\eth^2 z^1$$

$$= -8M^0\sigma_{ij}m^i m^j$$

Hence

$$\eth\bar{\eth}\Psi_0^1 + 10(\Psi_0^1 + \dot{\Psi}_0^0) = 0. \tag{5.5.18}$$

Again, this differential equation is easily solved to give

$$\Psi_0^1 = b^p(w)\,{}_2Y_{3p} - \frac{5}{3}\dot{a}^m(w)\,{}_2Y_{2m} \quad (p=-3,-2,\ldots,3),$$
$$(m=-2,-1,\ldots,2) \tag{5.5.19}$$

The explicit angular dependence calculated in this section may be used in at least two ways. Firstly, equations (5.4.24-28) may be simplified; secondly, the explicit angular dependence of observable quantities may be calculated.

By way of illustration, note from (5.4.24a,b) and (5.4.25b) that

$$5\Psi_1^1 = -\frac{1}{3}M^0(\frac{\bar{F}^0}{P}) - \frac{1}{\sqrt{2}}\bar{\eth}\Psi_0^1 + \frac{1}{2\sqrt{2}}\eth(M^1 + 2M^0 z^1) \tag{5.5.20}$$

Substituting (5.5.9,11,12) into (5.2.20) and using (3.4.56c, d), (3.4.58) and (5.5.13) yields

$$\Psi_1^1 = \frac{1}{5}[(\frac{1}{2}M^0\sigma_{ij} - \frac{3}{2}M^0\omega_{ij})m^i k^j - \frac{1}{2}(M_{;i})m^i + \sqrt{5}b^p\,{}_1Y_{3p} +$$

$$-\frac{5}{3\sqrt{2}}a^m{}_1Y_{2m}. \qquad (5.5.21)$$

Similar expressions may be derived for all the terms in (5.4.24-28) if desired. However, since no new insights appear to be gained from this rather tedious calculation, we shall not carry out the procedure here.

As an example of the second application, we see from (5.4.1b,11) and (5.5.3,9) that the observable quantity $(\sigma\frac{dv}{dz})$ is given by

$$(\sigma\frac{dv}{dz}) = [\sigma^1(\frac{dv}{dz})]v + o(v^2)$$

$$= [\tfrac{1}{3}\Psi^0_0(z^1)^{-1}][(z^1)^{-1}z + o(z^2)] + o(z^2)$$

from which (5.1.4) follows immediately:

$$(\sigma\frac{dv}{dz})_1 = \frac{1}{3}[\frac{1}{3}\Theta(w) - \sigma_{ij}(w)k^ik^j]^{-2}a^m(w)\,{}_2Y_{2m}$$

Thus, assuming that z^1 has already been measured, so that Θ and σ_{ij} are known, measurement of $(\sigma\frac{dv}{dz})_1$ determines the 5 complex constants a^m. (These, as has been remarked before, are simply linear combinations of the orthonormal tetrad components of the Weyl tensor at the centre.) The significance of (5.1.4) and similar expressions which may be derived for all the expansion coefficients in (5.4.1) is that the general angular behaviour of the observable quantities is completely determined in terms of the spin-weighted spherical harmonics; what needs to be found by observation are the constants a^m and their analogues for other quantities.

This underlines again the importance of measurements of the *angular variations* of cosmologically significant quantities (such as the Hubble parameter, z^1) to the observational approach being advocated here. Only by measuring these angular variations and Fourier analysing the results using the complete set of functions (3.4.56a) is it possible to determine the parameters (such as a^m) which determine the solution (5.4.24-28). Even if these angular measurements produce nothing more than an upper limit on the anisotropy of the various observables, the results could still be used to place bounds on the magnitude and anisotropy of these parameters, and then (5.4.24-28) would imply bounds on the possible space-time anisotropy.

§5.6. The non-radial equations

During the course of the integration performed in §5.5, all the radial equations (3.3.13,15,17) have been satisfied up to the order considered. The non-radial equations (3.3.14,16,18,19) are all satisfied to lowest order by virtue of the central conditions (3.4.45,46,47) and (5.5.1,3). It remains to be shown that the non-radial equations are either identically satisfied up to the order in v considered, or that they determine the w-derivatives of a consistent initial data set. Unless this can be proved, it is possible that new relations between the expansion coefficients listed in (5.4.24-28) can arise, thereby restricting even further the angular dependence of these quantities. It is even conceivable that these equations may show that some of the elements of the initial data set (5.4.1) are redundant. In this section we shall demonstrate that this is not the case: in addition to being sufficient to determine the solution (5.4.24-28), the initial data given by (5.4.1) is also necessary - it contains no redundant elements and the non-radial equations imply no angular relations not already derived (or implicit in (5.4.24-28)).

We consider firstly the contracted Bianchi identities (3.3.19). The first of these, (3.3.19a), may be written as

$$\begin{aligned}
\left(-\frac{\partial}{\partial w} + U\frac{\partial}{\partial v} + X^A\frac{\partial}{\partial x^A}\right)\Phi_{00} &= \left(\omega\frac{\partial}{\partial v} + \xi^A\frac{\partial}{\partial x^A}\right)\bar{\Phi}_{01} + \\
&+ \left(\bar{\omega}\frac{\partial}{\partial v} + \bar{\xi}^A\frac{\partial}{\partial x^A}\right)\Phi_{01} - \frac{\partial}{\partial v}(\Phi_{11} + 3\Lambda) + \\
&+ [2(\gamma+\bar{\gamma}) - (\mu+\bar{\mu})]\Phi_{00} - 2(\alpha+\bar{\tau})\Phi_{01} - 2(\bar{\alpha}+\tau)\bar{\Phi}_{01} \\
&+ 4\rho\Phi_{11} + \bar{\sigma}\Phi_{02} + \sigma\bar{\Phi}_{02} .
\end{aligned}$$

From the term in O(1), it follows that

$$\dot{M}^0 + 3M^0 z^1 + \frac{1}{2\sqrt{2}}[\eth(\frac{F^0}{P}) + \bar{\eth}(\frac{\bar{F}^0}{P})] = 0, \tag{5.6.1a}$$

where, as before (see §5.5), "$\cdot$" denotes $\frac{\partial}{\partial w}$. Similarly, (3.3.19b) implies

$$(\bar{F}^0)^{\cdot} + 2\bar{F}^0 z^1 + \frac{1}{2\sqrt{2}}\bar{F}^0[\eth(\frac{F^0}{P}) - \bar{\eth}(\frac{\bar{F}^0}{P})] +$$

$$+ \frac{P}{24\sqrt{2}}(\eth^3\bar{\Psi}^0_0 - 6\bar{\eth}\Psi^0_0) = 0. \tag{5.6.1b}$$

while the leading term in (3.3.19c) again implies (5.6.1a). (This is because we have assumed the fluid to be pressure-free; in the case of non-vanishing pressure, $p \neq 0$, equations (3.3.19a,c) determine $\dot{M}^0$ and $\dot{p}^0$.)

From the terms in $o(v)$ and $o(v^2)$, it is possible to deduce $\dot{M}^1$, $\dot{M}^2$, $\dot{F}^1$, $\dot{z}^1$ and $\dot{z}^2$; e.g.,

$$\dot{z}^1 + (z^1)^2 - 2P^{-2}F^0\bar{F}^0 + \frac{1}{6}M^0 + \frac{1}{2\sqrt{2}}(\frac{F^0}{P}\eth z^1 + \frac{\bar{F}^0}{P}\bar{\eth}z^1) +$$

$$+ \frac{1}{24}(\bar{\eth}^2\Psi^0_0 + \eth^2\bar{\Psi}^0_0) = 0. \tag{5.6.1c}$$

Thus, up to second order in v, the contracted Bianchi identities are either identically satisfied (this is the case, e.g., for the $o(v^{-1})$ term in (3.3.19a)), or they determine the w-derivatives of M^0, M^1, M^2, F^0, F^1, z^1 and z^2.

Next consider the Bianchi identities (3.3.18). The terms in $o(v^{-1})$ are satisfied by virtue of (5.5.1b,3) and (5.4.25). The terms in o(1) are satisfied iff

$$\dot{\Psi}^0_0 + \Psi^1_0 = \frac{1}{\sqrt{2}}\eth(\Phi^1_{01} + \Psi^1_1) \tag{5.6.2a}$$

$$\dot{\Psi}^0_1 + \frac{3}{2}\Psi^1_1 = \frac{1}{\sqrt{2}}\eth(\Psi^1_2 + \Phi^1_{11} - \Lambda^1) + \frac{1}{2}\Phi^1_{01} \tag{5.6.2b}$$

$$\dot{\Psi}^0_2 + 2\Psi^1_2 + \frac{1}{12}\dot{M}^0 = \frac{1}{\sqrt{2}}\eth(\Phi^1_{21} + \Psi^1_3) + 2\Phi^1_{22} - \Phi^1_{11} - \frac{1}{24}M^1 \tag{5.6.2c}$$

$$\dot{\Psi}_3^0 + \frac{5}{2}\Psi_3^1 = \frac{1}{\sqrt{2}}\eth\Psi_4^1 - \frac{1}{\sqrt{2}}\bar{\eth}\Phi_{22}^1 + \frac{3}{2}\Phi_{21}^1 . \tag{5.6.2d}$$

(In order to simplify (5.6.2c), we have used (3.3.19c).)

Equation (5.6.2a) determines $\dot{\Psi}_0^0$, since all the other terms in this equation are known; in fact, from (5.5.19), (5.4.24,25a) and (5.5.11) it follows that $\dot{a}^m$ is determined, where (see (5.5.3)) $\Psi_0^0 = a^m{}_2Y_{2m}$. But if one now compares (5.6.2) with (5.5.1a), it is clear that consistency problems may arise. For example, (5.5.1a) implies

$$\dot{\Psi}_1^0 = -\frac{1}{\sqrt{2}}\bar{\eth}\dot{\Psi}_0^0. \tag{5.6.3}$$

(That $\eth$ and $\frac{\partial}{\partial w}$ commute is a direct consequence of (3.4.30c) and (3.4.53).) Thus, by differentiating (5.6.2a) w.r.t. $\bar{\eth}$, and eliminating the term $\bar{\eth}\dot{\Psi}_0^0$ by using (5.6.2b,3), one obtains a differential equation involving quantities which are already known at this stage of the integration:

$$\frac{3}{2}\Psi_1^1 + \frac{1}{4\sqrt{2}}\bar{\eth}\Psi_0^1 = \frac{1}{8}\bar{\eth}\eth(\Phi_{01}^1 + \Psi_1^1) + \frac{1}{\sqrt{2}}\eth(\Psi_2^1 + \Phi_{11}^1 - \Lambda^1) + + \frac{1}{2}\Phi_{01}^1. \tag{5.6.4}$$

We need to show that this is identically satisfied by virtue of previously derived relations. Substituting for Ψ_1^1 and Ψ_2^1 from (5.4.25) and using the fact that

$$\bar{\eth}\eth(\Phi_{01}^1 + \Psi_1^1) = \eth\bar{\eth}(\Phi_{01}^1 + \Psi_1^1) + 2(\Psi_1^1 + \Phi_{01}^1)$$

(as follows from (5.5.5)), it is possible to show that (5.6.4) is identically satisfied iff

$$3M^0\eth z^1 = \sqrt{2}(\eth^2\Phi_{10}^1 + \eth\bar{\eth}\Phi_{01}^1), \tag{5.6.5a}$$

or, using (5.4.24b), iff

$$\eth\{3M^0z^1 + \frac{1}{2\sqrt{2}}[\eth(\frac{F^0}{P}) + \bar{\eth}(\frac{\bar{F}^0}{P})]\} = 0. \tag{5.6.5b}$$

But (5.6.5b) is identically satisfied by virtue of (5.6.1a) and the fact that $\eth\hat{M}^0 = 0$, M^0 being a function only of w. (Alternatively, one may substitute for z^1 and $\frac{F^0}{P}$ from (5.5.9,11) and then use (5.5.13) and (3.4.58) to verify that (5.6.5) is identically satisfied.)

The calculation above therefore shows that (5.6.2b) yields no new information once (5.6.2a) has been satisfied. In almost exactly the same way it is readily verified that (5.6.2c,d) are both identically satisfied by virtue of (5.6.1a,2a) and previously derived relations. Similarly it may be shown that the terms in o(v) and o(v^2) in (3.3.18a) determine $\dot{\Psi}^1_0$ and $\dot{\Psi}^2_0$, while the remaining equations (3.3.18b-d) are identically satisfied up to second order in v.

Next, consider the remaining non-radial equations (3.3.14,16). The calculations involved are again extremely tedious; as an example we shall show how the term in o(1) in (3.3.14a) is treated. The other equations and higher orders are handled analogously.

From (3.3.9) and the central conditions (3.4.45,46) it follows that (3.3.14a) is satisfied up to o(1) iff

$$\xi^{BO}X^{A1}{}_{,B} + \xi^{A1} + U^2\xi^{AO} - X^{B1}\xi^{AO}{}_{,B} = (\mu^1+\bar{\gamma}^1-\gamma^1)\xi^{AO} + \bar{\lambda}^1\bar{\xi}^{AO} \quad ,$$

which holds iff

$$X^{21}{}_{,2} - iX^{21}{}_{,3} - \frac{1}{P}\xi^{21} - X^{21}(\log P)_{,2} - X^{31}(\log P)_{,3} + U^2 = (\mu^1+\bar{\gamma}^1-\gamma^1) + \bar{\lambda}^1 \qquad (5.6.6a)$$

and

$$-iX^{31}{}_{,2} - X^{31}{}_{,3} + \frac{i}{P}\xi^{31} + X^{21}(\log P)_{,2} + X^{31}(\log P)_{,3} +$$

$$- U^2 = -(\mu^1+\bar{\gamma}^1-\gamma^1) + \bar{\lambda}^1. \tag{5.6.6b}$$

Adding, we find,

$$2\frac{\partial}{\partial\zeta}(X^{21}-iX^{31}) - \frac{1}{P}(\xi^{21}-i\xi^{31}) = 2\bar{\lambda}^1 , \tag{5.6.7}$$

where, as before, $\zeta \equiv x^2 + ix^3$. But, by (5.4.27),

$$X^{21} - iX^{31} = \frac{1}{4\sqrt{2}}P(\bar{\partial}\Psi^0_0) ,$$

while

$$\xi^{21} - i\xi^{31} = -\frac{1}{3}P\Psi^0_0.$$

Thus (5.6.7) holds iff

$$\frac{1}{2\sqrt{2}}\frac{\partial}{\partial\zeta}(P\bar{\partial}\Psi^0_0) + \frac{1}{3}\Psi^0_0 = 2\bar{\lambda}^1,$$

which, in turn, holds iff

$$\frac{1}{8}\partial\bar{\partial}\Psi^0_0 + \frac{1}{3}\Psi^0_0 = 2\bar{\lambda}^1. \tag{5.6.8}$$

But this is identically satisfied by virtue of (5.4.26f) and (5.5.2).

On the other hand, subtraction of (5.6.6a) and (5.6.6b) gives

$$2\frac{\partial}{\partial\zeta}(X^{21} + iX^{31}) - \frac{1}{P}(\xi^{21}+i\xi^{31}) + 2U^2 = 2X^{21}(\log P)_{,2} +$$

$$+ 2X^{31}(\log P)_{,3} + 2(\mu^1+\bar{\gamma}^1-\gamma^1). \tag{5.6.9}$$

Substituting for μ^1, γ^1, U^2, ξ^{A1} and X^{A1} from (5.4.26,27), we find that (5.6.9) holds iff

$$\frac{\partial}{\partial\zeta}(P\partial\bar{\Psi}^0_0) = \alpha^0\bar{\partial}\Psi^0_0 - \bar{\alpha}^0\partial\bar{\Psi}^0_0 + \frac{1}{2\sqrt{2}}\partial^2\bar{\Psi}^0_0 + (\partial\bar{\Psi}^0_0)\frac{\partial P}{\partial\zeta} + (\bar{\partial}\Psi^0_0)\frac{\partial P}{\partial\bar{\zeta}} ,$$

which, according to (3.4.40), holds iff

$$\partial^2\bar{\Psi}^0_0 = 2\sqrt{2}\,[P\frac{\partial}{\partial\zeta}(\partial\bar{\Psi}^0_0) - (\partial\bar{\Psi}^0_0)\frac{\partial P}{\partial\zeta}] \tag{5.6.10}$$

Since Ψ_0^0 is a spin-2 quantity, $\bar{\Psi}_0^0$ and $\eth\bar{\Psi}_0^0$ have spin-weights -2 and -1, respectively, and thus (5.6.10) is an identity by virtue of (3.4.53).

Similarly, it is found that the terms in $o(v)$ and $o(v^2)$ are identically satisfied. These results hold also for the other non-radial equations (3.3.14b-d) and (3.3.16).

Collecting results, we have therefore shown (§5.4) that, given initial data in the form (5.4.1), all the spin coefficients, metric variables and the Riemann and metric tensors are uniquely determined on the initial past light cone up to the order specified in (5.4.24-28). At this stage in the integration, all the radial equations (3.3.13,15,17) are satisfied up to the order in v considered. Next (§5.6), the non-radial equations (3.3.18a,19) are used to determine $\dot{\Psi}_0^n$, $\dot{M}^n$ (n=0,1,2), $\dot{z}^1$, $\dot{z}^2$, $\dot{F}^0$ and $\dot{F}^1$. These equations are now satisfied up to $o(v^2)$, while all the remaining non-radial equations are satisfied up to the relevant order in v.

But now it follows immediately from (5.4.1,3,5,7) that the w-derivatives of the observables are known up to the same order in v as the observables themselves are known.

Hence, at least in the analytic case, the solution may be propagated off the initial past light cone as discussed in §§5.1,3. (When the initial data is given in the form (5.4.1), it is not necessary to consider an alternative initial data set; the w-derivatives of all the elements of $D(w_o, z^*)$ are determined directly.)

Finally, it should be noted that calculations of the kind performed in §§5.4-6 were first carried out by Kristian and Sachs ([54]). It is readily verified that our results are in agreement with theirs, up to the order they consider (which is lower than that given in (5.4.1) and (5.4.24-28; cf. their equation (85) and our (5.4.3).)

§5.7. *The Bondi-Sachs method*

In §5.6 we showed that, once the hypersurface integration of §5.4 has been performed to determine the solution on the initial past light cone, the remaining equations are either identically satisfied, or they determine the w-derivatives of the elements of an initial data set. The proof involves a series of long calculations, because the full set of NP equations is used.

In this section we show how consistency proofs of this kind are simplified if one works instead only with the ten independent tetrad components of the Einstein field equations

$$R_{ab} = -Mu_a u_b + \tfrac{1}{2}Mg_{ab} \ . \tag{5.7.1}$$

The method is a simple generalisation to the perfect fluid case of the vacuum analysis of Bondi et. al. ([59]) and Sachs ([60]; see also Chellone and Williams, [83]), and will only be outlined very briefly.

We shall use the null observational coordinate system (w,v,θ,ϕ), where w and v are chosen as in §3.3, while (θ,ϕ) are related to the previously used angular coordinates (x^A) by

$$e^{i\phi}\cot(\tfrac{1}{2}\theta) = x^2 + ix^3$$

It is convenient to work with a different null tetrad $\{\underline{n},\underline{k},\underline{m},\underline{\bar{m}}\}$ from that employed thus far. Let $\underline{k}$ be defined as in §3.3 and consider the tangent space $T_q\mathcal{M}$ of some point q in M. Choose $\underline{n}|_q$ to be the (unique) null vector which lies in the time-like 2-plane spanned by $\frac{\partial}{\partial w}|_q$ and $\underline{k}|_q$, and which satisfies $\underline{n}|_q \cdot \underline{k}|_q = 1$. Let $\underline{X}|_q$ and $\underline{Y}|_q$ be two (real) unit space-like vectors, mutually orthogonal and orthogonal to both $\underline{n}|_q$ and $\underline{k}|_q$. These are defined only up to a rigid rotation in the space-like 2-plane which they span; we make a definite choice by (arbitrarily) demanding that the component X^3 of $\underline{X}$ with respect to the coordinate basis should vanish. Then

$$\underline{k} = \frac{\partial}{\partial v} \tag{5.7.2a}$$

$$\underline{n} = -\frac{\partial}{\partial w} + \frac{1}{2}g_{00}\frac{\partial}{\partial v} \tag{5.7.2b}$$

$$\underline{X} = (\sqrt{g_{22}})^{-1}(-g_{02}\frac{\partial}{\partial v} + \frac{\partial}{\partial \theta}) \tag{5.7.2c}$$

$$\underline{Y} = (\sqrt{hg_{22}})^{-1}[(g_{02}g_{23} - g_{03}g_{22})\frac{\partial}{\partial v} - g_{23}\frac{\partial}{\partial\theta} + g_{22}\frac{\partial}{\partial\phi}] \tag{5.7.2d}$$

where

$$h \equiv \det g_{AB} .$$

By setting

$$\underline{m} = \frac{1}{\sqrt{2}}(\underline{X} + i\underline{Y}) \tag{5.7.3}$$

we thus obtain a null tetrad satisfying (3.2.1).

Again, this tetrad is not well-defined globally, but it is well-defined at least in some simply convex normal neighbourhood U of the point $w = w_0$ on C (except, of course, at points on C).

With this choice of tetrad, the spin coefficients can be shown to satisfy the following conditions:

$$\kappa = 0, \quad Re(\varepsilon) = 0, \quad \rho - \bar{\rho} = 0, \quad \tau = \bar{\alpha} + \beta, \quad \bar{\tau} + \pi = 0,$$
$$Im(\sigma) = 2Im(\varepsilon), \quad Im(\lambda + \mu) = 2Im(\gamma). \tag{5.7.4}$$

The non-zero spin coefficients can now be expressed in terms of the metric components (3.3.8c). It is convenient to use the notation

$$\partial_X \equiv \sqrt{2}\mathit{Re}\,\delta \equiv \sqrt{2}\mathit{Re}(m^a\nabla_a), \quad \partial_Y \equiv \sqrt{2}\mathit{Im}\,\delta \equiv \sqrt{2}\mathit{Im}(m^a\nabla_a).$$

Then

$$\sigma = \frac{1}{2}\{D[\log(\frac{\sqrt{h}}{g_{22}})]\} - \frac{i}{2}[\frac{g_{22}}{\sqrt{h}} D(\frac{g_{23}}{g_{22}})] \tag{5.7.5a}$$

$$\rho = -\frac{1}{4}D(\log h) \tag{5.7.5b}$$

$$\tau = \frac{1}{2\sqrt{2}}(\frac{1}{\sqrt{g_{22}}}Dg_{02}) + \frac{i}{2\sqrt{2h}}(-\frac{g_{23}}{\sqrt{g_{22}}}Dg_{02} + \sqrt{g_{22}}Dg_{03}) \tag{5.7.5c}$$

$$\nu = -\frac{1}{\sqrt{2}}(\frac{1}{\sqrt{g_{22}}}\Delta g_{02} + \frac{1}{2}\partial_X g_{00}) - \frac{i}{\sqrt{2h}}(\frac{g_{23}}{\sqrt{g_{22}}}\Delta g_{02} +$$

$$- \sqrt{g_{22}}\,\Delta g_{03} - \frac{1}{2}\sqrt{h}\,\partial_Y g_{00}) \tag{5.7.5d}$$

$$\lambda = -\frac{1}{2}\{\Delta[\log(\frac{\sqrt{h}}{g_{22}})]\} - [\frac{i}{2}\frac{g_{22}}{\sqrt{h}}\Delta(\frac{g_{23}}{g_{22}})] \tag{5.7.5e}$$

$$\mu = \frac{1}{4}[\Delta(\log h)] - \frac{i}{2}(\frac{g_{23}}{\sqrt{hg_{22}}}\partial_X g_{02} - \frac{\sqrt{g_{22}}}{\sqrt{h}}\partial_X g_{03} +$$

$$+ \frac{1}{\sqrt{g_{22}}}\partial_Y g_{02}) \tag{5.7.5f}$$

$$\pi = -\frac{1}{2\sqrt{2}g_{22}}Dg_{02} - \frac{i}{2\sqrt{2h}}(\frac{g_{23}}{\sqrt{g_{22}}}Dg_{02} - \sqrt{g_{22}}Dg_{03}) \tag{5.7.5g}$$

$$\beta = \frac{1}{4\sqrt{2}}\{\frac{1}{\sqrt{g_{22}}}Dg_{02} - \partial_X[\log(\frac{g_{22}}{h})]\} + \frac{i}{4\sqrt{2h}}[-\frac{g_{23}}{\sqrt{g_{22}}}Dg_{02} +$$

$$+ \sqrt{g_{22}}Dg_{03} + \sqrt{h}\,\partial_Y(\log g_{22}) - 2g_{22}\partial_X(\frac{g_{23}}{g_{22}})] \tag{5.7.5h}$$

$$\alpha = \frac{1}{4\sqrt{2}}\{\frac{1}{\sqrt{g_{22}}} Dg_{02} + \partial_X[\log(\frac{g_{22}}{h})]\} + \frac{i}{4\sqrt{2h}}[\frac{g_{23}}{\sqrt{g_{22}}} Dg_{02} +$$

$$- \sqrt{g_{22}}\, Dg_{03} - \sqrt{h}\partial_Y(\log g_{22}) + 2g_{22}\partial_X(\frac{g_{23}}{g_{22}})] \tag{5.7.5i}$$

$$\varepsilon = -\frac{i}{4}\frac{g_{22}}{\sqrt{h}} D(\frac{g_{23}}{g_{22}}) \tag{5.7.5j}$$

$$\gamma = -\frac{1}{4} Dg_{00} - \frac{i}{4\sqrt{h}}[\frac{g_{23}}{\sqrt{g_{22}}} \partial_X g_{02} - \sqrt{g_{22}}\partial_X g_{03} + \frac{\sqrt{h}}{\sqrt{g_{22}}} \partial_Y g_{02} +$$

$$+ \frac{g_{22}}{\sqrt{h}} \Delta(\frac{g_{23}}{g_{22}})] \tag{5.7.5k}$$

and

$$\Phi_{00} = \frac{1}{2}M(1+z)^2 \tag{5.7.6a}$$

$$\Phi_{01} = \frac{-1}{2\sqrt{2g_{22}}}M(1+z)(u^A g_{A2} - i\sqrt{h}u^3) \tag{5.7.6b}$$

$$\Phi_{02} = \frac{1}{4g_{22}}M\{[(u^A g_{A2})^2 - h(u^3)^2] - 2i(\sqrt{h}u^3 u^A g_{A2})\} \tag{5.7.6c}$$

$$\Phi_{11} = \frac{1}{8}M(1 - 2u^A u^B g_{AB}) \tag{5.7.6d}$$

$$\Phi_{12} = \frac{1}{4\sqrt{2g_{22}}}M(1+z)^{-1}(u^A u^B g_{AB} - 1)(u^A g_{A2} - i\sqrt{h}u^3) \tag{5.7.6e}$$

$$\Phi_{22} = \frac{1}{8}M(1+z)^{-2}(u^A u^B g_{AB} - 1)^2 \tag{5.7.6f}$$

$$\Lambda = \frac{1}{24}M \tag{5.7.6g}$$

The 36 (real) NP equations (3.3.15,16) may be regarded as consisting of (i) the 10 tetrad components of the field equations (5.7.1), (ii) 10 equations determining the Weyl curvature Ψ_Q and (iii) the 16 components of the Jacobi identities for the tetrad vectors. By taking linear combinations of the NP equations, one can eliminate the Ψ_Q's to obtain 10 *field equations*. These may be divided into 3 groups:

I. Main Equations.

(i) Hypersurface equations:

$$D\rho = \rho^2 + \sigma\bar{\sigma} + \Phi_{00} \tag{5.7.7a}$$

$$D\tau = (\rho+2\varepsilon)\tau + \bar{\delta}\sigma + (4\bar{\beta}-3\bar{\tau})\sigma - \delta\rho + 2\Phi_{01} \tag{5.7.7b}$$

(ii) Standard equations:

$$\begin{aligned} D(Re\mu) = {} & 2Re[\rho\mu + \delta(\bar{\beta}-\bar{\tau})] + 3(\tau-\beta)(\bar{\tau}-\bar{\beta}) + \beta\bar{\beta} + \\ & + \varepsilon(\mu-\bar{\mu}) + (\Phi_{11}+3\Lambda) \end{aligned} \tag{5.7.7c}$$

$$D\lambda = (\rho-4\varepsilon)\lambda + \bar{\sigma}\mu + 2\bar{\beta}\bar{\tau} - \bar{\delta}\bar{\tau} + \Phi_{20} \tag{5.7.7d}$$

II. Trivial Equation.

$$D[Re(\mu-\gamma)] = Re(\rho\mu - 2\bar{\beta}\tau - \bar{\delta}\tau + \sigma\lambda) + 3\tau\bar{\tau} - (\Phi_{11}-3\Lambda) \tag{5.7.7e}$$

III. Supplementary Conditions.

$$\Delta\beta = \delta\gamma - \mu\tau - \beta(\mu+\bar{\gamma}-\gamma) + \varepsilon\bar{\nu} + \sigma\nu + \bar{\lambda}(\bar{\beta}-\bar{\tau}) - \Phi_{12} \tag{5.7.7f}$$

$$\Delta(Re\mu) = Re[\delta\nu - \mu^2 - \lambda\bar{\lambda} - \mu(\gamma+\bar{\gamma}) - \bar{\nu}\bar{\tau} + 2\beta\nu] - \Phi_{22} \tag{5.7.7g}$$

(This terminology was first introduced by Bondi ([59]).)

Using the commutators (3.3.12) together with the simplifying conditions (5.7.4), it can be shown that

$$\Delta\sigma = (\gamma+\bar{\gamma})\sigma - D\lambda + 2(\bar{\sigma}Re\lambda - \bar{\lambda}Re\sigma) \tag{5.7.8a}$$

$$\Delta\rho = (\gamma+\bar{\gamma})\sigma - D(Re\mu) \ . \tag{5.7.8b}$$

It is useful to introduce some notation (see [83]) to label the field equations. Let

$$E_{ab} = R_{ab} - \frac{1}{2}Rg_{ab} + T_{ab} \ . \tag{5.7.9}$$

By comparing the definitions (3.2.4b) of the Φ_{IJ}'s with the last term in each of (5.7.7), it is seen that the *hypersurface equations* are $E_{ab}k^ak^b = 0$ and $E_{ab}k^am^b = 0$, respectively; the *standard equations* are $E_{ab}n^ak^b = 0$ and $E_{ab}m^am^b = 0$, respectively, while the *trivial equation* is $E_{ab}m^a\bar{m}^b = 0$ and the *supplementary* *conditions* are $E_{ab}n^a m^b = 0$ and $E_{ab}n^a n^b = 0$, respectively.

Bondi's lemma concerning the relationship between the various groups of field equations becomes, in our case

Lemma 5.7.1.

Let U be a simply convex normal neighbourhood of the point $w = w_0$ on C. Suppose the main equations and the contracted Bianchi identities are satisfied in U, and $\rho \neq 0$. Then the trivial equation is identically satisfied in U, while the supplementary conditions are satisfied on $C^-(p) \cap U$ iff for each null generator Σ of $C^-(p)$, there exists a point $q \in \Sigma \cap U$ such that the supplementary conditions are satisfied at q.

Proof.

(Throughout this proof, indices refer to components with respect to the null tetrad $\{\underline{n},\underline{k},\underline{m},\underline{\bar{m}}\}$. Thus, e.g., $E_{13} = E_{ab}k^a\bar{m}^b$. The tetrad components of the metric tensor are given by (3.3.2).)

Suppose the main equations and the contracted Bianchi identities are satisfied in U. The contracted Bianchi identities may be written as

$$g^{cb}E_{ab;c} = 0. \tag{5.7.10}$$

For $a = 1$, this reads,

$$g^{cb}(E_{1b,c} - E_{db}\Gamma^d_{c1} - E_{1d}\Gamma^d_{cb}) = 0,$$

(where a comma now denotes a tetrad derivative).

The first and last terms in this equation vanish by hypothesis, while the surviving term gives

$$E_{02}\Gamma^0_{31} + \bar{E}_{02}\bar{\Gamma}^0_{03} + E_{23}(\Gamma^2_{21} + \bar{\Gamma}^2_{21}) = 0.$$

Since

$$\Gamma^{0}_{31} = \Gamma_{131} = k^{b}\bar{m}^{a}\nabla_{a}k_{b} = 0,$$

and

$$\Gamma^{2}_{21} = \Gamma_{321} = \bar{m}^{b}m^{a}\nabla_{a}k_{b} = \bar{\rho} = \rho,$$

this is simply

$$\rho E_{23} = 0.$$

Since $\rho \neq 0$, it follows that the trivial equation is satisfied. Similarly, for $a = 2,3$, (5.7.10) implies

$$D[\sqrt{hg_{22}}(E_{02} + \bar{E}_{02})] = 0, \tag{5.7.11a}$$

$$D[\frac{h}{\sqrt{g_{22}}}(E_{02} - \bar{E}_{02}) + 2ig_{03}\frac{\sqrt{h}}{\sqrt{g_{22}}}(E_{02} + \bar{E}_{02})] = 0 \tag{5.7.11b}$$

$$D(\sqrt{h}E_{00}) = \sqrt{h}[\delta\bar{E}_{02} + \bar{\delta}E_{02} - (\tau - 2\beta)\bar{E}_{02} - (\bar{\tau} - 2\bar{\beta})E_{02}] \tag{5.7.11c}$$

(where we have used (5.7.5)).

Thus, if E_{02} and E_{00} vanish at some point q on Γ, then they vanish everywhere as long as (5.7.11) is valid, i.e. they vanish on $\Gamma \cap U$. The result follows. □

This lemma shows that, once the main equations and the contracted Bianchi identities have been satisfied in U, one need merely check that the supplementary conditions are satisfied on a hypersurface $\{v = \text{constant}\}$ in order to ensure that they are satisfied everywhere in U.

A formal iterative procedure similar to that of §5.2 is now readily established. It should be noted that the initial data set $D(w_0, z^*)$ was derived under the assumption that the null tetrad used satisfies the conditions (3.3.22). Since the tetrad (5.7.2) does *not* satisfy these conditions, the distortion is not simply represented by the spin coefficient σ. However, since g_{AB} is measurable (see §4.2.3), we can take as initial data the set

$$D'(w_o, z^*) \equiv \{(u^o, u^A, g_{AB}, M\tfrac{dv}{dz}) \mid w = w_o,\ 0 \le \theta < \pi,\ 0 \le \phi < 2\pi\}.$$

Using the central conditions

$$g_{00} = 1 + g^2_{00}v^2 + o(v^3)$$

$$g_{AB} = \mathrm{diag}(1, \sin^2\theta)v^2 + \begin{pmatrix} g^4_{22} & g^4_{23} \\ g^4_{23} & g^4_{33} \end{pmatrix} v^4 + o(v^5),$$

$$g_{0A} = g^3_{0A}v^3 + o(v^4),$$

the hypersurface integration is easily performed to yield z, g_{AB}, u^a and M as functions of (v,θ,ϕ) on $C^-(p, v^{**}_L)$.

As before, the contracted Bianchi identities are written as a system of linear algebraic equations which determine the w-derivatives of z, M and u^A. A distinct advantage of the tetrad (5.7.2) is that the tetrad components of the trace-free Ricci tensor are known as soon as the hypersurface equation (5.7.7a) has been solved to give $z = z(v)$. (Cf. the discussion following (5.2.10).)

The standard equations are now integrated to give λ and μ, and now (5.7.5a,b) and (5.7.8) determine $g_{AB,0}$. The solution may then be propagated off $C^-(p)$ as discussed in §5.3.

When the results of this integration are substituted into (5.7.7f,g) it is found that

$$\lim_{v\to 0} E_{ab}n^a m^b = 0 = \lim_{v\to 0} E_{ab}n^a n^b,$$

and it therefore follows from lemma 5.7.1 that the supplementary conditions are satisfied (since they are satisfied at the vertex of the past light cone $C^-(p)$.)

Details of this procedure, for a perfect fluid in the axially symmetric case, may be found in [83]. The general case sketched here presents no new problems.

6. SPHERICALLY SYMMETRIC SPACE-TIME: AN EXAMPLE

§6.1. Introduction

It was noted in §1.4 that the observational approach to cosmology can be used in conjunction with more conventional methods. In this chapter we shall assume, a priori, that the space-time under consideration is spherically symmetric about a regular geodesic world line, and then undertake an observational analysis. These space-times are of interest for a number of reasons. Firstly, Maartens ([5]) has shown that the class of pressure-free, isotropic ideal observational space-times (see §4.5) is simply the Bondi-Tolman solution (see, e.g., [81]). Secondly, spherically symmetric models have recently been investigated as possible alternatives to the FRW models, but these investigations have all proceeded along conventional lines (see, e.g., [21]). Finally, as was explained in §1.3, it would be useful to have some observational test of the spatial homogeneity assumption underlying the FRW models. This chapter provides such a test in the context of spherically symmetric space-times: in §6.4 it will be shown that a spherically symmetric pressure-free model is spatially homogeneous iff the $r(z)$ and $(M \frac{dv}{dz})(z)$ relations take precisely the form (see (6.4.12,14) below) predicted by the FRW models. (In this connection, see also [1].)

In §6.2, the simplification of the metric and Riemann tensor components and the spin coefficients induced by the assumption of spherical symmetry will be derived, and the non-trivial NP-equations and Bianchi identities isolated. The integration scheme of §5.2. will then be examined in §6.3.

Finally, in §6.4, we shall derive the form of the r(z) and $(M \frac{dv}{dz})(z)$ relations under the assumption that the space-time is, in addition, spatially homogeneous. Then the observational integration will be carried out with exactly these observational relations as initial data, and it will be shown that the resulting space-time is spatially homogeneous.

§6.2. Simplification of the field equations

Suppose the space-time under consideration is spherically symmetric about a regular geodesic world line C (see, e.g., [25]). The metric is then of type D and ([85]) the null directions $\underline{n}$ and $\underline{k}$ as specified in (3.3.9), where now $x^A = 0$, are repeated principal null directions, so that

$$\Psi_0 = \Psi_1 = \Psi_3 = \Psi_4 = 0 \tag{6.2.1a}$$

while the remaining Weyl tensor component satisfies

$$\Psi_2 = \bar{\Psi}_2 \tag{6.2.1b}$$

Clearly the proper motions must vanish, for otherwise they would define preferred directions on the celestial sphere, contrary to the assumption of spherical symmetry. Hence, by (5.2.2,3,5a)

$$\Phi_{01} = \Phi_{02} = \Phi_{12} = 0 \tag{6.2.2a}$$

$$\Phi_{00} = \tfrac{1}{2}M(1+z)^2 \tag{6.2.2b}$$

$$\Phi_{11} = \tfrac{1}{8}M \tag{6.2.2c}$$

$$\Phi_{22} = \tfrac{1}{8}M(1+z)^{-2} \tag{6.2.2d}$$

Also, it follows from the symmetry assumed that all physical quantities must be functions only of w and v, so that

$$r = r(w,v), \quad M = M(w,v), \quad z = z(w,v) \tag{6.2.3}$$

As before, the spin coefficients satisfy (3.3.22).

Equation (3.3.15b) now reads:

$$D\sigma = 2\rho\sigma$$

or, since, by (4.2.17a), $\rho = -\frac{Dr}{r}$,

$$D(r^2\sigma) = 0. \tag{6.2.4a}$$

But, by (3.4.45a) and (4.2.18b),

$$\sigma = o(v), \quad r = v + o(v^3)$$

and hence (6.2.4a) implies

$$\sigma = 0. \tag{6.2.4b}$$

Thus, by (3.3.13a)

$$D(r\xi^A) = 0$$

so that

$$\xi^A = r^{-1}\ F^A(w,x^B), \text{ some functions } F^A. \tag{6.2.5}$$

Evaluating (6.2.5) in the limit $v \to 0$ using (3.4.14) and (4.2.18b) gives

$$\xi^A = \xi^{A0}\ r^{-1} \tag{6.2.6}$$

where, as before, ξ^{A0} is defined by (3.4.17).

Since $X^A = 0$, (3.3.13c) implies

$$\tau = 0 \tag{6.2.7a}$$

and hence, by (3.3.11)

$$\alpha = -\bar{\beta} \tag{6.2.7b}$$

while (3.3.13b) gives

$$D(r\omega) = 0. \tag{6.2.8}$$

Since $\omega = o(v^2)$, this means that

$$\omega = 0. \tag{6.2.9}$$

Similarly, (3.3.15g,i) imply

$$\lambda = \nu = 0, \tag{6.2.10}$$

while (3.3.15d) may be integrated to yield

$$\alpha = \alpha^0\ r^{-1} \tag{6.2.11}$$

where α^0 is defined by (3.4.45b).

Next, we note from (3.3.14d) and (6.2.9) that

$$\mu = \bar{\mu} \tag{6.2.12}$$

and from (3.3.15f) and (6.2.1) that

$$D(\gamma - \bar{\gamma}) = 0$$

and thus

$$\gamma = \bar{\gamma} \tag{6.2.13}$$

(since $\gamma = o(v)$).

From (3.3.14a) it now follows that

$$\Delta\xi^A = -\mu\xi^A$$

whence, by (6.2.6),

$$\mu = \frac{\Delta r}{r} \tag{6.2.14}$$

(It is interesting that, in this case, the expansion of the $\underline{k}$ - congruence is $\rho = -\nabla_{\underline{k}}(\log r)$, while the expansion of the $\underline{n}$ - congruence is $\mu = \nabla_{\underline{n}}(\log r)$. This is related to the fact (see below) that r contains all the information about both the intrinsic and extrinsic curvature of the past light cones of points on C.)

Before considering the non-radial equations, let us examine the meaning of the equation

$$\delta\eta = 0.$$

Note that, by (3.3.9) and (6.2.6,9)

$$\delta\eta = 0 \quad \text{iff} \quad r^{-1}\,\xi^{AO}\eta_{,A} = 0.$$

Thus, if η is real, then

$$\delta\eta = 0 \quad \text{iff} \quad \eta_{,A} = 0 \quad \text{iff} \quad \eta = \eta(w,v) \tag{6.2.15}$$

while, if η is complex, then

$$\delta\eta = 0 \quad \text{iff} \quad [Re(\eta)]_{,2} + [Im(\eta)]_{,3} = 0 \tag{6.2.16a}$$

$$\text{and} \quad [Re(\eta)]_{,3} - [Im(\eta)]_{,2} = 0 \tag{6.2.16b}$$

which means that η is analytic.

Applying (6.2.15) to (3.3.14b) gives

$$U = U(w,v) \tag{6.2.17}$$

and, similarly, (3.3.17c) shows

$$\Psi_2 = \Psi_2(w,v). \tag{6.2.18}$$

The non-radial equations (3.3.16a,b,d,f,g,i), (3.3.18a, b,d) and the radial equations (3.3.17a,c,d) are now all trivially satisfied, and it is not difficult to show that (3.3.14c) is satisfied iff

$$\alpha = -r^{-1}\frac{\partial P}{\partial \bar{\zeta}},$$

which is an identity by virtue of (6.2.11) and (3.4.45b). Thus the non-radial equations (3.3.14) are all satisfied.

The remaining Bianchi identities (3.3.17b) and (3.3.18c) may now be integrated as follows: from (3.3.15a,b), (3.3.16e,h), (3.3.17b) and (3.3.18c) it follows that

$$D(\rho\mu - \Psi_2 + \Lambda + \Phi_{11}) = 2\rho(\rho\mu - \Psi_2 + \Lambda + \Phi_{11}) \tag{6.2.19a}$$

$$\Delta(\rho\mu - \Psi_2 + \Lambda + \Phi_{11}) = -2\mu(\rho\mu - \Psi_2 + \Lambda + \Phi_{11}) \tag{6.2.19b}$$

But, since $\rho = -D(\log r)$ and $\mu = \Delta(\log r)$, this gives

$$\rho\mu - \Psi_2 + \Lambda + \Phi_{11} = G(x^a)r^{-2} \tag{6.2.19c}$$

where, as usual, $G(x^a)$ is determined by evaluating (6.2.19c) at the centre. This yields

$$\rho\mu - \Psi_2 + \Lambda + \Phi_{11} = \tfrac{1}{2}r^{-2} \tag{6.2.20}$$

(The Penrose complex curvature invariant K of a 2-surface S (see, e.g., [56],[76]) is defined by

$$K = \rho\mu - \Psi_2 + \Lambda + \Phi_{11} \tag{6.2.21}$$

Choosing S to be the 2-surface $\{w = \text{constant},\ v = \text{constant}\}$ gives the simple relationship

$$K = \tfrac{1}{2}(r|_S)^{-2} \tag{6.2.22}$$

The real part of K is related to the Gaussian curvature, ${}^{(2)}R$, of S by

$$^{(2)}R = (K + \bar{K}) = r^{-2} \tag{6.2.23}$$

while the imaginary part of K describes an extrinsic curvature invariant of S in the manifold in which it is embedded (in this case, the only non-trivial extrinsic curvature invariant). Thus the area distance r determines completely both the intrinsic and extrinsic geometry of the 2-dimensional cross-sections S of the past light cones of points on C.)

From (6.2.20) it is readily shown that (3.3.16c) is satisfied iff

$$P \frac{\partial^2 P}{\partial\zeta\partial\bar{\zeta}} - \frac{\partial P}{\partial\zeta}\frac{\partial P}{\partial\bar{\zeta}} = \frac{1}{8} ,$$

an identity by virtue of (3.4.46b).

Collecting results, the spin coefficients, metric variables, metric and Riemann tensor components have therefore been shown to satisfy the following conditions:

$$\kappa = \varepsilon = \pi = \tau = \lambda = \nu = 0 ; \tag{6.2.24a}$$

$$\alpha = - \frac{\partial P}{\partial\bar{\zeta}} r^{-1} , \beta = -\bar{\alpha} , \rho = - D(\log r),$$
$$\mu = \Delta(\log r); \tag{6.2.24b}$$

$$\omega = 0 = X^A, \ U = U(w,v), \ \xi^A = \xi^{A0} r^{-1}; \tag{6.2.24c}$$

where

$$\xi^{20} = -P, \ \xi^{30} = iP, \ P = \frac{1}{2\sqrt{2}}(1+\zeta\bar{\zeta}); \tag{6.2.24d}$$

$$\Psi_0 = \Psi_1 = \Psi_3 = \Psi_4 = 0; \tag{6.2.24e}$$

$$\Psi_2 = \Psi_2(w,v), \ \Psi_2 = \bar{\Psi}_2; \tag{6.2.24f}$$

$$\Phi_{01} = \Phi_{02} = \Phi_{12} = 0; \tag{6.2.24g}$$

$$\Phi_{00} = \frac{1}{2}M(1+z)^2, \ \Phi_{11} = \frac{1}{8}M, \ \Phi_{22} = \frac{1}{8}M(1+z)^{-2}; \tag{6.2.24h}$$

$$\Lambda = \frac{1}{24}M; \tag{6.2.24i}$$

$$g^{11} = 2U,\ g^{1A} = 0,\ g^{AB} = \mathrm{diag}(\ -r^{-2},\ -r^{-2}\mathrm{cosec}^2\theta) \qquad (6.2.24j)$$

the last equation being a consequence of (3.3.10c), (3.4.43) and (6.2.24c,d).

Of the 37 equations (3.3.13-19), only the following still need to be satisfied:

$$DU = -2\gamma \qquad (6.2.25)$$

$$D\rho = \rho^2 + \Phi_{00} \qquad (6.2.26a)$$

$$D\mu = \rho\mu + \Psi_2 + 2\Lambda \qquad (6.2.26b)$$

$$D\gamma = \Psi_2 - \Lambda + \Phi_{11} \qquad (6.2.26c)$$

$$\Delta\mu = -(\mu+2\gamma)\rho + \Phi_{22} \qquad (6.2.27a)$$

$$\Delta\rho = (2\gamma-\mu)\mu - \Psi_2 - 2\Lambda \qquad (6.2.27b)$$

$$\Delta\Phi_{00} + D(\Phi_{11}+3\Lambda) = (4\gamma-2\mu)\Phi_{00} + 4\rho\Phi_{11} \qquad (6.2.28a)$$

$$\Delta(\Phi_{11}+3\Lambda) + D\Phi_{22} = -4\mu\Phi_{11} + 2\rho\Phi_{22} \qquad (6.2.28b)$$

§6.3. Observational integration scheme

The observational integration scheme presented in §5.2 simplifies considerably in the case of spherical symmetry. The only non-trivial elements of the maximal data set $D(w_o,z^*)$ are now the area distance- and number counts-redshift relations, so that we may define

$$D_S(w_o,z^*) = \{r(z),\ (M\frac{dv}{dz})(z)\,|\,w = w_o,\ 0 < z \le z^*\} \qquad (6.3.1)$$

(where S denotes spherical symmetry) as the initial data set in this case.

As before, (6.2.26a) is used to determine the relation between the redshift z and the affine distance v. Again, one rewrites (6.2.26a) as

$$f_1 \frac{d}{dz}(\frac{dz}{dv}) + f_2(\frac{dz}{dv}) + f_3 = 0 \qquad (6.3.2)$$

where now

$$f_1 = r^{-1}\frac{dr}{dz} \tag{6.3.3a}$$

$$f_2 = r^{-1}\frac{d^2r}{dz^2} \tag{6.3.3b}$$

$$f_3 = \frac{1}{2}(M\frac{dv}{dz})(1+z)^2. \tag{6.3.3c}$$

Equation (6.3.2) is now easily integrated to give

$$\frac{dv}{dz} = (\frac{dr}{dz})\{1 - \frac{1}{2}\int_o^z r(M\frac{dv}{dz})(1+z)^2\, dz\}^{-1} \tag{6.3.4}$$

and then a simple quadrature determines $v=v(z)$ and hence $z=z(v)$.

Once (6.3.4) has been solved, both r and M are known as functions of v on the initial light cone, and thus ρ, ξ^A, α, β and all the non-trivial components of the Ricci tensor may be determined.

This leaves μ and γ as the only spin coefficients still to be found. But, by (6.2.20,26b),

$$D\mu = 2\rho\mu + \Phi_{11} + 3\Lambda - \frac{1}{2}r^{-2}$$

so that

$$\mu = \frac{1}{r^2}\int_o^v (\frac{1}{4}Mr^2 - \frac{1}{2})\, dv \tag{6.3.5a}$$

or

$$\mu = \frac{1}{r^2}\int_o^z (\frac{1}{4}M\frac{dv}{dz}r^2 - \frac{1}{2}\frac{dv}{dz})\, dz \tag{6.3.5b}$$

Once μ has been determined, Ψ_2 follows immediately from (6.2.20), while γ is obtained by integrating (6.2.26c):

$$\gamma = \int_o^v (\rho\mu - \frac{1}{4}M - \frac{1}{2}r^{-2})\, dv \tag{6.3.6}$$

and now (6.2.25) yields

$$U = -\frac{1}{2} - 2\int_o^v \gamma\, dv \tag{6.3.7}$$

This completes the radial integration: all the spin coefficients, metric variables, metric and Riemann tensor components are now determined on the initial past light cone $C^{-}(p)$. But Bondi has shown ([81]) that a pressure-free, spherically symmetric space-time is completely determined by specification of 2 functions of a single variable: thus the required existence and uniqueness of the solution off $C^{-}(p)$ follow from the fact that the initial data set $D_S(w_o, z^*)$ contains precisely two functions of one variable (see (6.3.1). In fact, the entire radial integration above may be carried out in the co-moving coordinate system used by Bondi, and it can then be shown explicitly how $D(w_o, z^*)$ determines the two arbitrary functions of the Bondi solution. (The derivation, which is not entirely trivial, will not be given here.)

Even though the Bondi solution is reasonably simple when expressed in co-moving coordinates, we have not been able to reproduce it by integrating the NP-equations or by transforming from co-moving to null observational coordinates. The basic problem is that, even in co-moving coordinates, it is not possible to find explicitly the relation between z (or r) and the affine distance v: the differential equation governing this relationship has thus far proved intractable. It is precisely this relation that is necessary in order to write down explicitly the required coordinate transformation.

It is not surprising that the propagation equations (6.2.27,28) are difficult to integrate - the null coordinates we have chosen are tied into the null geometry (and hence to the observations) rather than to fluid. Similarly, in co-moving coordinates, the observational integration is far more complicated than the scheme presented here.

§6.4 *Observational integration with FRW initial data*

In this section we show, by explicit integration, that a pressure-free, spherically symmetric space-time is, in addition, spatially homogeneous iff the $r(z)$ and $(M\frac{dv}{dz})(z)$ relations take exactly the form predicted by the standard FRW models (see equations (6.4.12,14)).

Suppose, firstly, that the universe is both spherically symmetric and spatially homogeneous. It is well-known (see, e.g., [73]) that the space-time is then conformally flat, so that

$$\Psi_2 = 0. \tag{6.4.1}$$

From (3.3.17b,18c,19c) and (6.2.24g,h,i) it then follows that

$$\frac{4}{3}\, D\Phi_{11} = -\mu\Phi_{00} + 2\rho\Phi_{11} \tag{6.4.2a}$$

$$2\, D\Phi_{22} = -2\mu\Phi_{11} + \rho\Phi_{22}. \tag{6.4.2b}$$

But (6.2.24h) implies

$$\Phi_{00}\Phi_{22} = 4(\Phi_{11})^2 \tag{6.4.2c}$$

and hence

$$\Phi_{11}\, D\Phi_{22} - \frac{1}{3}\Phi_{22}\, D\Phi_{11} = 0, \tag{6.4.3}$$

which integrates to give

$$\Phi_{22} = \frac{2}{a}(\Phi_{11})^{1/3}, \tag{6.4.4a}$$

where

$$a = 8(M_o)^{-2/3} \tag{6.4.4b}$$

and now (6.4.2c) shows that

$$\Phi_{00} = 2a(\Phi_{11})^{5/3} \tag{6.4.4c}$$

Thus, by (6.2.24h) and (6.4.4),

$$(1+z)^4 = \frac{1}{4}\Phi_{00}(\Phi_{22})^{-1} = \frac{1}{4}a^2(\Phi_{11})^{4/3} = \left(\frac{M}{M_o}\right)^{4/3},$$

which means that

$$8\Phi_{11} = M = M_o(1+z)^3. \tag{6.4.5}$$

In order to find $(\frac{dz}{dv})$ and $(M\frac{dv}{dz})$, we differentiate (6.4.2a) and substitute for $D\rho$ and $D\mu$ from (6.6.26a,b). This gives

$$3\Phi_{11}\, D^2\Phi_{11} - 5(D\Phi_{11})^2 - 6a(\Phi_{11})^{1/3} = 0 \tag{6.4.6}$$

which, integrated once, gives

$$(D\Phi_{11})^2 = (\Phi_{11})^{10/3}\{b+12a(\Phi_{11})^{1/3}\} \tag{6.4.7}$$

where $b = b(w)$ is a "constant" of integration whose value will be determined later by evaluating (6.4.7) at the centre.

By substituting for $\Phi_{11} = \frac{1}{8}M$ from (6.4.5) and for $D\Phi_{11}$ from

$$D\Phi_{11} = \left(\frac{d\,\Phi_{11}}{dz}\right)\left(\frac{dz}{dv}\right) = \frac{3}{8}M_o(1+z)^2\left(\frac{dz}{dv}\right) \tag{6.4.8}$$

into (6.4.7), and evaluating the result at the centre, one obtains

$$\frac{dz}{dv} = (1+z)^3\left(H_o + \frac{1}{3}M_o z\right)^{1/2} \tag{6.4.9}$$

where

$$H_o = \left.\frac{dz}{dv}\right|_o \tag{6.4.10a}$$

is just the usual Hubble parameter (in the notation of the previous chapter, $H_o = z^1$).

It is customary to define the so-called deceleration parameter q_o by

$$q_o = \frac{M_o}{6H_o^2}\ . \tag{6.4.10b}$$

so that (6.4.7) assumes the more familiar form

$$\frac{dz}{dv} = H_o(1+z)^3(1+2q_oz)^{1/2}. \tag{6.4.11}$$

Thus, by (6.4.5), the $(M\frac{dv}{dz})(z)$ relation becomes in this case

$$M\frac{dv}{dz} = 6q_oH_o(1+2q_oz)^{-1/2}. \tag{6.4.12}$$

The r(z) relation may now be derived by integrating (6.2.26a). Since

$$\rho = -\frac{1}{r}Dr = -\frac{1}{r}(\frac{dr}{dz})(\frac{dz}{dv}),$$

while Φ_{00} is given by (6.4.4c) and (6.4.5), equation (6.2.26a) now reads

$$\frac{d^2r}{dz^2} + \{3(1+z)^{-1} + q_o(1+2q_oz)^{-1}\}\frac{dr}{dz} + $$
$$+ 3q_o(1+z)^{-1}(1+2q_oz)^{-1}r = 0 \tag{6.4.13a}$$

with initial conditions

$$r(0) = 0, \quad (\frac{dr}{dz})(0) = (\frac{dz}{dv})(0) = H_o \tag{6.4.13b}$$

Now (6.4.13a) is a linear, homogeneous differential equation which is easily solved to give

$$r(z) = H_o^{-1}q_o^{-2}(1+z)^{-2}\{q_oz + 1 - q_o +$$
$$(q_o - 1)(1+2q_oz)^{1/2}\} \tag{6.4.14}$$

Thus (6.4.12,14) are the observational relations predicted if the universe is both spatially homogeneous and isotropic, ie. if the space-time is FRW.

Suppose now, conversely, that the space-time is isotropic about C, and that the maximal data set $D_S(w_o,z^*)$ is given by (6.4.12,14). Then, as shown in §6.3,

$$\frac{dv}{dz} = (\frac{dr}{dz})\{1 - \frac{1}{2}\int_0^z r(M\frac{dv}{dz})(1+z)^2\,dz\}^{-1}$$

$$= H_o^{-1}q_o^{-2}(1+z)^{-3}(1+2q_oz)^{-1/2}\{(3q_o-2-q_oz)(1+2q_oz)^{1/2} + $$
$$+ (q_o-1)(q_o-2-3q_oz)\}\{1-3q_o^{-1}(q_o-1)z - q_o^{-2}(q_oz+2-3q_o)\times$$
$$\times(1+2q_oz)^{1/2}+q_o^{-2}(2-3q_o)\}^{-1}$$

$$= H_o^{-1}(1+z)^{-3}(1+2q_oz)^{-1/2} \qquad (6.4.15)$$

(cf. (6.4.11)). By (6.3.5)

$$2\mu r^2 = \int_o^z (\tfrac{1}{2}M\frac{dv}{dz}r^2 - \frac{dv}{dz})\,dz$$

Substituting for $M\frac{dv}{dz}$ from (6.4.12), for r from (6.4.14) and for $\frac{dv}{dz}$ from (6.4.15) gives

$$2\mu r^2 = 3q_o^{-3}H_o^{-1}\{(1-2q_o)^2I_4 + 2q_o(1-2q_o)I_3 + q_o^2I_2 +$$
$$-\tfrac{1}{3}q_o^3I_3 + (q_o-1)^2\int_o^z(1+z)^{-4}(1+2q_oz)^{1/2}\,dz +$$
$$-q_o(q_o-1)(1+z)^{-2} - \tfrac{2}{3}(q_o-1)(1-2q_o)(1+z)^{-3}\} \qquad (6.4.16a)$$

where

$$I_n \equiv \int_o^z(1+z)^{-n}(1+2q_oz)^{-1/2}\,dz\ . \qquad (6.4.16b)$$

Using

$$I_n = \frac{(1+2q_oz)^{1/2}}{(n-1)(2q_o-1)(1+z)^{n-1}} + \frac{(2n-3)q_o}{(n-1)(2q_o-1)}I_{n-1}\,,$$
$$(n\geq 2;\ q_o\neq\tfrac{1}{2}) \qquad (6.4.17a)$$

and

$$\int_o^z(1+z)^{-4}(1+2q_oz)^{1/2}\,dz = -\tfrac{1}{3}(1+2q_oz)^{1/2}(1+z)^{-3} +$$
$$+\tfrac{1}{3}q_oI_3 \qquad (6.4.17b)$$

the integration is readily performed and yields

$$2\mu r^2 = -q_o^{-3}H_o^{-1}(1+z)^{-3}\{[q_oz + (q_o-1)(q_o-2)]\times$$
$$\times(1+2q_oz)^{1/2} + 3q_o(q_o-1)z + (q_o-1)(2-q_o)\} \qquad (6.4.18)$$

since the terms in I_1, which involve logarithms or hyperbolic tangents of z (depending on the sign of $(1-2q_o)$) cancel. Even though (6.4.17a) is defined only for $2q_o \neq 1$, so that the case $2q_o = 1$ needs to be considered seperately, (6.4.18) is still valid for both cases.

Hence, finally

$$\mu = -\frac{1}{2}q_0H_0(1+z)\{(q_0-1) + (1+2q_0z)^{1/2}\}\{q_0z+1-q_0 + +(q_0-1)(1+2q_0z)^{1/2}\}^{-1}. \qquad (6.4.19)$$

Note that, for small values of z,

$$\mu = -\frac{1}{2}H_0 z^{-1} + o(z)$$

as required by the central limit (3.3.45a) and the fact that

$$z = z^1v + o(v^2) = H_0v + o(v^2).$$

Now, by (6.2.20),

$$\Psi_2 = \rho\mu + \frac{4}{3}\Phi_{11} - \frac{1}{2}r^{-2}$$

so that

$$\Psi_2 r^2 = -\mu r\frac{dr}{dz}\frac{dz}{dv} + \frac{4}{3}\Phi_{11}r^2 - \frac{1}{2}.$$

But, by (6.4.12,14,15,19),

$$\mu r\frac{dr}{dz}\frac{dz}{dv} = -\frac{1}{2}q_0^{-3}(1+z)^{-1}\{(3q_0-2-q_0z)(1+2q_0z) + +(q_0-1)^2(q_0-2-3q_0z) + (q_0-1)(4q_0-4-4q_0z)\times \times(1+2q_0z)^{1/2}\} \qquad (6.4.20)$$

$$= \frac{4}{3}\Phi_{11}r^2 - \frac{1}{2}$$

and hence, by (6.4.20),

$$\Psi_2 r^2 = 0.$$

This proves that Ψ_2 vanishes everyhere, except possibly at the centre. But that Ψ_2 vanishes at $v = 0$ is an immediate consequence of (5.5.1a) and the fact that (see (6.2.1a)) $\Psi_0 = 0$. Once this has been established, it becomes a simple matter, using Bondi's co-moving coordinates, to show that the two functions that determine the solution take precisely their FRW values, and thus the space-time is spatially homogeneous. This result is valid in the region of space-time into which the initial data is dragged by the fluid flow lines, and thus also to the future of $C^{-}(p)$. It therefore follows that in this case, a "no-interference" condition of the kind discussed in §5.3. is automatically satisfied. This is because the condition that the fluid be both pressure-free and spherically symmetric appears to break the characteristic property of the field equations.

7. SUMMARY AND CONCLUSIONS

We have examined a number of alternatives to the standard Friedman-Robertson-Walker models of the universe. In chapter 1 the basic assumptions underlying most modern cosmological theories were investigated, and it was argued that such alternatives need to be considered because of the dubious observational status of the space-time symmetry assumptions on which the FRW models are based.

The possible consequences of weakening these assumptions was studied in chapter 2 and Appendix III. In chapter 2 we presented a systematic method for obtaining explicit solutions of the field equations in the case of models which are spatially homogeneous (but anisotropic). This method allows one to regain, in a unified manner, many of the known solutions for this class of model. In addition, we derived some (apparently) new solutions. (See also Appendix IV.)

In Appendix III, we considered a static, spherically symmetric (but inhomogeneous) model, in which the observer is assumed to be located near a regular center of symmetry. It was shown that, cosmographically, such a model appears to be as consistent with currently available cosmological observations as the FRW models. When the field equations are imposed, it becomes difficult to fit the predicted area distance-redshift relation to the observed data, but this difficulty can be overcome by dropping the restriction that the space-time should be static.

A more radical alternative was considered in the remainder of the thesis, in which the conventional approach based on a priori symmetry assumptions is abandoned altogether.

Instead, one attempts to discover what can be deduced directly from observations made at one point on our galactic world line C. In chapter 3, we introduced a system of null coordinates, proved that they were observational coordinates, and deduced the limiting behaviour of the metric and curvature tensors in the vicinity of C by integrating the NP equations under the assumption that C is a regular, geodesic world line.

In chapter 4 we described a number of cosmologically significant astronomical observations and used these to derive a maximal data set $D(w_o,z^*)$, which was taken to represent the maximal information that could be obtained at the point of observation by observations of distant galaxies. It was noted that not all the elements of $D(w_o,z^*)$ could be expected to be measured within the near future, and various ways of circumventing this problem were discussed. We pointed out that those observations which were not currently available could simply be assumed to have a particular functional form, thereby giving rise to the concept of an observational space-time, i.e. a space-time in which the observational relations predicted by the field equations are precisely those assumed. This is analogous to the conventional approach, in which a variety of space-time symmetry assumptions can be made, each leading to a different class of space-times. The crucial difference here is that any assumptions made about $D(w_o,z^*)$ can, in principle, be verified observationally at one space-time point.

Under the assumption that the universe may be modelled by a pressure-free perfect fluid, we proved in §5.2 that $D(w_o,z^*)$ determines a unique solution of the field equations on some part of the past light cone $C^-(p)$. Less rigorously, we indicated in §5.3 how the solution may be propagated towards the interior of $C^-(p)$. In §5.4 we performed the integration explicitly for initial data of the from (5.4.1), and in §5.5 showed how the general angular behaviour of the initial data could be determined. The main conclusion to be drawn from the results of these two sections is that it is extremely important to attempt to measure the angular variation of observable quantities. Given such data, the approximate structure of the model in the vicinity of our galactic world line is determined explicitly by (5.4.4,5, 24-28). These results will be discussed in more detail elsewhere (see [3],[4]).

A number of problems that arose during the course of our discussion of the cosmological initial value problem in chapter 5 deserve to be investigated in more detail. Firstly, it would be useful to have a rigorous proof for the existence and uniqueness of the solution off $C^-(p)$ in the general (non-analytic) case. This would shed some light on the question of what additional data needs to be specified, or what conditions can be imposed in order to ensure that the solution is determined to the future as well as to the past of $C^-(p)$. Such a proof would also allow one to undertake a stability analysis of the kind discussed in chapter 1, i.e. to determine how observational errors are propagated through the field equations.

Possibly the most fruitful area for future research would be a detailed investigation of the properties of the various classes of observational space-time defined in §4.5. Of these, the most promising appears to be the case of Simplified Ideal Observations, in which both the proper motions and distortion are assumed to vanish, but the (r,z) and $(M\frac{dv}{dz},z)$ relations are arbitrary. Thus far, no definite characterisation of these space-times has been found (but see [5]).

The observational approach can also be used in conjunction with more conventional methods, as was demonstrated in chapter 6, where we showed how the integration scheme presented earlier could be simplified in the case of a spherically symmetric space-time. As an application, we showed that such a space-time is, in addition, spatially homogeneous if and only if the observed (r,z) and $(M\frac{dv}{dz},z)$ relations take on $C^-(p)$ exactly the form predicted by the FRW models. In principle, this provides an observational method of verifying that the universe is spatially homogeneous, once isotropy about our world line has been established.

Although we considered only the case of a pressure-free perfect fluid, no problems arise when one extends these results by assuming, e.g., that the stress-energy tensor consists of two separately conserved radiation and matter components, both represented by the same 4-velocity field u^a, and satisfying the equations of state

$$p_r = \frac{1}{3}M_r, \quad p_m = 0.$$

Instead of being related by (5.2.7), the affine distance v and the redshift z are now related by an equation of the form

$$\left(\frac{dz}{dv}\right)\frac{d}{dz}\left(\frac{dz}{dv}\right) + h_1\left(\frac{dz}{dv}\right)^2 + h_2\left(\frac{dz}{dv}\right) + h_3 = 0,$$

where $h_i = h_i(z,x^A)$ are all measurable, provided one assumes that the background radiation can adequately be described by a blackbody spectrum with temperature T varying as $(1+z)^4$ (see §4.4). The rest of the integration scheme remains essentially unchanged.

APPENDIX I

Transformation of the spin coefficients and curvature tensor components under null and spatial rotations.

Consider a null tetrad $\underline{n},\underline{k},\underline{m},\underline{\bar{m}}$ satisfying (3.2.1) and (3.3.22). A null rotation about the tetrad vector $\underline{k}$ is given by

$$k^{a'} = k^a, \quad n^{a'} = n^a + B\bar{B}k^a + \bar{B}m^a + B\bar{m}^a,$$

$$m^{a'} = m^a + Bk^a. \tag{I.1}$$

We consider only the case where B is independent of v (since this preserves the conditions (3.3.22)).Then the spin coefficients transform as ([86])

$$\rho' = \rho, \quad \alpha' = \alpha + \bar{B}\rho \ , \quad \lambda' = \lambda + 2\bar{B}\alpha + \rho\bar{B}^2 + \bar{\delta}\bar{B}$$

$$\kappa' = \kappa, \quad \varepsilon' = \varepsilon, \quad \pi' = \pi, \quad \sigma' = \sigma, \quad \beta' = \beta + \bar{B}\sigma,$$

$$\mu' = \mu + 2\bar{B}\beta + \bar{B}^2\sigma + \delta\bar{B}, \quad \tau' = \tau + \bar{B}\sigma + B\rho,$$

$$\gamma' = \gamma + \bar{B}(\tau+\beta) + B\alpha + \bar{B}^2\sigma + B\bar{B}\rho$$

$$\nu' = \nu + \bar{B}(\mu+2\gamma) + B\lambda + \bar{B}^2(\tau+2\beta) + 2B\bar{B}\alpha + B\bar{B}^2\rho + \bar{B}^3\sigma +$$

$$+ \Delta\bar{B} + \bar{B}\delta\bar{B} + B\bar{\delta}\bar{B}. \tag{I.2}$$

The tetrad components of the Weyl tensor transform as ([86])

$$\Psi_0' = \Psi_0, \quad \Psi_1' = \bar{B}\Psi_0 + \Psi_1, \quad \Psi_2' = \bar{B}^2\Psi_0 + 2\bar{B}\Psi_1 + \Psi_2,$$

$$\Psi_3' = \bar{B}^3\Psi_0 + 3\bar{B}^2\Psi_1 + 3\bar{B}\Psi_2 + \Psi_3,$$

$$\Psi_4' = \bar{B}^4\Psi_0 + 4\bar{B}^3\Psi_1 + 6\bar{B}^2\Psi_2 + 4\bar{B}\Psi_3 + \Psi_4. \tag{I.3}$$

The tetrad components of the trace-free Ricci tensor transform as ([86])

$$\Phi_{00}' = \Phi_{00}, \quad \Phi_{01}' = B\Phi_{00} + \Phi_{01}, \quad \Phi_{02}' = B^2\Phi_{00} + 2B\Phi_{01} + \Phi_{02},$$

$$\Phi_{11}' = B\bar{B}\Phi_{00} + \bar{B}\Phi_{01} + B\Phi_{10} + \Phi_{11},$$

$$\Phi_{12}' = B^2\bar{B}\Phi_{00} + 2B\bar{B}\Phi_{01} + \bar{B}\Phi_{02} + 2B\Phi_{11} + B^2\Phi_{10} + \Phi_{12},$$

$$\Phi'_{22} = B^2\bar{B}^2\Phi_{00} + 2B\bar{B}^2\Phi_{01} + \bar{B}^2\Phi_{02} + 2B^2\bar{B}\Phi_{10} + 4B\bar{B}\Phi_{11} +$$

$$+ 2\bar{B}\Phi_{12} + B^2\Phi_{20} + 2B\Phi_{21} + \Phi_{22}. \qquad \text{(I.4)}$$

Under a spatial rotation in the ($\underline{m}$-$\bar{\underline{m}}$)-plane:

$$k^{a'} = k^a, \quad n^{a'} = n^a, \quad m^{a'} = e^{iC}m^a, \qquad \text{(I.5)}$$

the spin coefficients transform as ([86]; again $C = C(w,x^A)$)

$$\rho' = \rho, \quad \alpha' = \alpha + i\bar{\delta}C, \quad \lambda' = e^{-2iC}\lambda, \quad \kappa' = e^{2iC}\kappa,$$

$$\pi' = \pi, \quad \varepsilon' = \varepsilon, \quad \sigma' = e^{2iC}\sigma, \quad \beta' = \beta + i\bar{\delta}C,$$

$$\mu' = \mu, \quad \tau' = e^{iC}\tau, \quad \gamma' = \gamma + i\Delta C, \quad \nu' = e^{-iC}\nu. \qquad \text{(I.6)}$$

The tetrad components of the Weyl tensor transform as ([86])

$$\Psi'_0 = e^{2iC}\Psi_0, \quad \Psi'_1 = e^{iC}\Psi_1, \quad \Psi'_2 = \Psi_2, \quad \Psi'_3 = e^{-iC}\Psi_3,$$

$$\Psi'_4 = e^{-2iC}\Psi_4. \qquad \text{(I.7)}$$

The tetrad components of the trace-free Ricci tensor transform as ([86])

$$\Phi'_{00} = \Phi_{00}, \quad \Phi'_{01} = e^{iC}\Phi_{01}, \quad \Phi'_{02} = e^{2iC}\Phi_{02},$$

$$\Phi'_{11} = \Phi_{11}, \quad \Phi'_{12} = e^{-iC}\Phi_{12}, \quad \Phi'_{22} = \Phi_{22}. \qquad \text{(I.8)}$$

APPENDIX II

Spin-weights of the Newman-Penrose variables.

The spin-weights of the NP variables (see §3.4.4) are as follows:

Spin coefficients

$$\kappa : 2, \quad \sigma : 2, \quad \rho : 0, \quad \tau : 1, \quad \nu : -1, \quad \lambda : -2,$$
$$\mu : 0, \quad \pi : -1. \tag{II.1}$$

The spin coefficients α, β, γ and ε do not, in general, have well-defined spin-weights. However, if one restricts attention to spatial rotations in which the group parameter C (see I.5) is independent of v, then ε can be assigned a spin-weight of 0.

Tetrad components of the Weyl tensor

$$\Psi_0 : 2, \quad \Psi_1 : 1, \quad \Psi_2 : 0, \quad \Psi_3 : -1, \quad \Psi_4 : -2. \tag{II.2}$$

Tetrad components of the trace-free Ricci tensor

$$\Phi_{00} : 0, \quad \Phi_{01} : 1, \quad \Phi_{02} : 2, \quad \Phi_{11} : 0, \quad \Phi_{12} : 1,$$
$$\Phi_{22} : 0. \tag{II.3}$$

Note that if η has weight s, then $\bar{\eta}$ has weight -s. Thus, since $\Phi_{IJ} = \bar{\Phi}_{JI}$, it follows that, e.g., Φ_{10} has weight -1.

APPENDIX III

Mon. Not. R. astr. Soc. (1978) **184**, 439–465

The expansion of the Universe

G. F. R. Ellis, R. Maartens and S. D. Nel *Department of Applied Mathematics, University of Cape Town, Rondebosch, Cape Town, South Africa*

Received 1978 January 13; in original form 1977 September 19

Summary. We show that, in principle, the observed galactic redshifts and microwave background radiation in the Universe can be explained by a static spherically symmetric (or SSS) universe model, in which there is a singularity which continually interacts with the Universe (rather than the once for all interaction that occurs in the Friedmann–Robertson–Walker, or FRW, universe models). However, when Einstein's field equations are introduced with a plausible equation of state, we cannot obtain a good fit to the observed (m, z) relation in a SSS universe model with all the properties we require.

The standard discussions of the expansion of the Universe proceed from an unverified *a priori* assumption – the Cosmological Principle – and deduce that the Universe is expanding, on the basis of that assumption. Our result constitutes a proof that the Universe is expanding, which is not based on *a priori* assumptions about space–time geometry. Such a proof is needed if one is seriously to attempt to make cosmology an observationally-based subject.

In addition, we use the SSS models to point out various features of interest that could arise in expanding, inhomogeneous universe models. Particularly, we show that an observer cannot deduce from the isotropy of the microwave background radiation, that the Universe is spherically symmetric about him; and that there may exist plausible cosmological models with a 'continuously active' singularity, such as is envisaged for the SSS universe models. This could have interesting implications for 'Machian' discussions.

In such models, the observer would be situated somewhere near the centre. We point out that there could be good reasons for such a situation.

1 Introduction

It is probable that the most significant features of the Universe that are currently observed are the systematic redshifts of the galaxies, and the spectrum and isotropy of the cosmic microwave background radiation (see, e.g. the surveys by Sciama [1]; Peebles [2]; Weinberg [3]).

The isotropy of the microwave background radiation is usually used as evidence for the isotropy of the Universe about the observer (see [1–4]). The galactic redshifts are usually interpreted as evidence for the expansion of the Universe, and this is supported by the interpretation of the microwave background radiation as the relic radiation from a hot big bang at the beginning of the Universe.

Given our current understanding of the nature of space–time, one may regard this expansion from an initial big bang as the most important single property of the Universe. Accordingly it is important to see how uniquely the observational evidence leads to this conclusion.

The assumption of spatial homogeneity (the 'cosmological principle') together with assuming isotropy of the space–time about a single observer, leads unambiguously to the Friedmann–Robertson–Walker (or FRW) universe models. The interpretation of the redshift as evidence for the expansion of the Universe, is then unique. However, if the assumption that the Universe is spatially homogeneous is dropped, the situation is not so clear.

The problem is that while isotropy is directly observable, homogeneity (on a cosmological scale) is not. In the standard discussions the assumption of homogeneity is made *a priori*, either directly, or in some equivalent form (e.g. as the assumption that the Universe is isotropic for *all* observers [4–6]), and so is not subjected to observational verification. Accordingly the standard 'proof' of the expansion of the Universe is based on an unverified *a priori* assumption.

In this paper we consider what happens when the homogeneity assumption is dropped. We shall show that, in principle, present observations can also be explained by a static spherically-symmetric universe model with two centres, and our Galaxy near one of the centres. In this case the systematic redshifts of the galaxies are interpreted as cosmological gravitational redshifts, while the microwave background radiation originates from a hot gas surrounding a singularity situated at the second centre of the Universe. That is, a non-expanding cosmographic [3] universe model could explain current observations. Within the context of such a model, the philosphic objections one might have about our Galaxy being near the centre of the Universe, can be met in a reasonably plausible way.

However, we shall also show that it is difficult to find a viable cosmological model of this kind. More precisely, when Einstein's field equations are imposed and a suitable equation of state chosen, we have been unable to obtain a close fit to the observed magnitude–redshift relation in such a model with all the desired properties. Our methods are not completely conclusive, but they make it seem unlikely that such models exist. (More optimistic statements in the Gravity Foundation Essay have had to be revised in view of the more detailed analysis carried out, particularly regarding the implications of the centrality condition.) Although our investigation suggests that an exactly static inhomogeneous cosmology is not viable, it points out certain interesting features of such cosmologies which might remain in realistic expanding inhomogeneous universe models; in particular, the singularity structure in such universes could be completely different from that in a FRW universe model.

2 The assumptions

2.1 PHYSICAL ASSUMPTIONS

We assume that normal local physical theory is valid at all points of the Universe to a good approximation. In accordance with the usual cosmological description, we represent the large-scale distribution of matter and radiation in the Universe by a continuum approximation; that is, we assume space–time is filled by a family (a 'congruence') of world lines representing the average motion of matter at each point (see, e.g. [7]). The unit tangent vector **u** to these world lines is then the average 4-velocity of the matter. This cosmic fluid will have associated with it well-defined macroscopic properties, represented by the scalar fields n (the number density of particles), μ (the total energy density), and p (the fluid pressure). The exact meaning of these symbols will vary depending on the physical situation (for a 'gas' of galaxies, n would be the number density of galaxies, but the Universe could be currently dominated by an intergalactic gas composed of hydrogen and helium ions, in which case n would be the number density of these ions). It seems that the 'perfect fluid' approximation may be sufficiently accurate, that is, it may be permissible to ignore viscosity, heat conduction, etc. [7].

The kinematic or cosmographic approach [3] proceeds by specifying a space–time, and the fluid 4-velocity and galaxy number density in that space–time. It is then possible to work out the observational relations that would be seen in such a space–time by a 'fundamental observer', i.e. an observer moving with 4-velocity **u**. One can then attempt to deduce the geometry of the Universe by comparing actual observations with these theoretical relations and (without invoking the field equations), adjusting the space–time geometry to reproduce the observed relations.

By contrast, in the cosmological approach [1–3] we prescribe a dynamical behaviour for the cosmological fluid by prescribing equations of state relating n, μ and p; and we determine the dynamical behaviour of the space–time itself by adopting Einstein's field equations. Then the observational relations seen by a fundamental observer can be predicted from these dynamical assumptions, plus suitable initial conditions; one can compare these predictions with the actual observations to determine the remaining parameters, which will then determine the cosmological model.

2.2 GEOMETRICAL ASSUMPTIONS

In either case we restrict the space–time geometries considered by noting that there is no agreed evidence that, on a cosmological scale, the Universe is anisotropic about us. We will therefore construct our cosmological models on the assumptions that the Universe is precisely isotropic about some particular observer.

The second assumption usually made is that the Universe is spatially homogeneous (or equivalently, that it appears isotropic to *all* observers). However, this is precisely the assumption we wish to query; in the main body of this paper, we replace it by the assumption that the Universe is static. The geometries we consider are therefore the static, spherically symmetric (or SSS) space–times. We further restrict the situation by supposing that the fundamental observers move so as to *see* the Universe as static and spherically symmetric, and that the space–time possesses at least one regular centre (corresponding to the world line of the observer who sees the Universe to be isotropic). (Strictly, they *observe* the Universe to be *rotationally* symmetric [32], and *deduce* that it is *spherically* symmetric (except for the central observer, who actually observes the spherical symmetry; *cf.* [6]).) This is the situation we will consider in the subsequent sections.

3 Static, spherically-symmetric space–time geometry

A space–time is spherically symmetric if it admits the orthogonal group $O(3, R)$ as a symmetry group of two-dimensional space-like trajectories (Schmidt [8]; box 23.3 of [9]). If, in addition, it is static, coordinates $\{t, r, \theta, \phi\}$ can be chosen ([8], [9]; see also [4], [10]) so that the metric takes the form

$$ds^2 = -g^2(r)\,dt^2 + dr^2 + f^2(r)\;\{d\theta^2 + \sin^2\theta\,d\phi^2\} \tag{1}$$

where $-\infty < t < \infty$, and θ, ϕ are the usual spherical coordinates ($0 < \theta < \pi$, $0 < \phi < 2\pi$).

The Killing vector generating the static symmetry orthogonal to the group trajectories is $\partial/\partial t$; so observers moving with 4-velocity **u** will see the Universe as static and spherically symmetric if

$$u^a = g^{-1}\,\delta^a_0 (\Rightarrow u_a u^a = -1). \tag{2}$$

(There is a 4-parameter family of time-like Killing vector fields, but only a 1-parameter family which is orthogonal to the orbits of the rotation group. For convenience, we select from this family that Killing vector field with group parameter t.) The world lines of matter are then the lines $\{r, \theta, \phi \text{ constant}\}$, which are non-geodesic where $g'(r) \neq 0$. The scalar quantities n, μ, p will depend only on r, because they are invariant under the group of space–time symmetries (the observers moving with 4-velocity **u** see them to be static and spherically-symmetric quantities).

Let the central world line, denoted by C, be at $r = 0$; because this is a regular centre of the space–time one has [8], on using the notation $' \equiv d/dr$,

$$\left.\begin{array}{l} f(r) \to 0, f'(r) \to 1, f''(r)/f(r) \to \text{finite limit} \\ g(r) \to \text{finite non-zero limit}, g'(r) \to 0 \end{array}\right\} \text{as } r \to 0. \tag{3}$$

The space–time is then spherically symmetric about the world line C. We will normalize $g(r)$ by the condition that $g(0) = 1$.

We assume that $f(r)$ and $g(r)$ are regular, non-zero functions of r on the interval $0 < r < R$, where R may be finite or infinite.

4 Cosmography

4.1 OBSERVATIONAL RELATIONS

Having specified the metric tensor, the matter 4-velocity, and the number density $n(r)$ of galaxies, the basic observational relations are now determined (*cf.* [7]) in terms of the functions $g(r)$, $f(r)$ and $n(r)$. The goal of the kinematic or cosmographic approach would be to choose these functions so that relations predicted by the model fit the actual observations.

Because $\partial/\partial t$ is a Killing vector, $\mathbf{k}\cdot(\partial/\partial t)$ is constant along any null geodesic with tangent vector **k**; consequently the redshift z of a source at $r = r_e$ as measured by an observer at C is given by

$$1 + z = (k^a u_a)_{r_e}/(k^b u_b)_C = g^{-1}(r_e)/g^{-1}(0) = g^{-1}(r_e). \tag{4}$$

(For details of the principles of derivation of these relations, see, e.g. [7].) Also the area spanned at $r = r_e$ by a tube of radial null geodesics through the origin is $f^2(r_e)\,d\Omega_0$, where $d\Omega_0 = \sin\theta\,d\theta\,d\phi$ is the solid angle subtended by these geodesics at the origin. Accordingly the 'observer area distance' r_0 of the source at r_e, as measured by C, is

$$r_0 = f(r_e). \tag{5}$$

Therefore the flux of radiation F measured by C for a source of luminosity L at r_e is given by

$$F = (L/4\pi)(1 + z)^{-4}(r_0)^{-2} = (L/4\pi)\,g^4(r_e)/f^2(r_e), \tag{6}$$

and the (m, z) relation is given by

$$m = -2.5 \log_{10} L + 5 \log_{10}\{f(r_e)(1+z)^2\} + \text{constant}. \tag{7}$$

Finally, the number $N(r_e)$ of sources seen up to distance r_e in a solid angle $d\Omega_0$ will be given by

$$N = d\Omega_0 \int_0^{r_e} f^2(r)\, n(r)\, dr. \tag{8}$$

Because of the spherical symmetry of the metric (1) and the matter distribution (see (2)), together with the centrality condition (3), cosmological observations made by an observer on the central world line C will be exactly isotropic, i.e. the same (m, z) and (N, z) relations will be seen in all directions by such an observer. By continuity, observations made by an observer near C will be nearly isotropic, except for very small redshifts.

In the cosmographic approach, $f(r)$, $g(r)$ and $n(r)$ are to be determined from the observational relations. If we take as our basic data the observed (m, z) and (N, z) relations, there are a large number of ways this can be done. (Equivalently, the (N, S) curve for radio sources could be considered.) One should note here, that first the distance scales in such models could be quite different from those in the FRW universe models, for the build up of z relative to proper distance r can be very different. Secondly, the intrinsic luminosity L of the galaxies observed may in general be expected to depend on r, i.e. in (6) and (7) we may expect that $L = L(r_e)$ where $L(r)$ is a function to be determined from astrophysical considerations (see below). Thirdly, some caution must be exercised because of the 'central' relations (3) which constrain the way the functions f and g behave as $r \to 0$; in fact they imply $z \to 0$, $dz/dr \to 0$ and $dr_0/dr \to 1$ as $r \to 0$. If we could observe the (m, z) relation down to $z = 0$, this would cause serious difficulties at this point; however, in fact the (m,z) relation cannot be used for distances closer than those corresponding to a redshift of about 0.004 (the fluid approximation does not hold for close galaxies, as random velocities dominate the redshift). Accordingly one can indeed adjust the available arbitrary functions to reproduce the observed curves in their range of validity.

That is, one can in this way make cosmographic spherically-symmetric universe models which precisely reproduce any spherically-symmetric (m, z) and (N, z) observations that are made of the Universe, by supposing our Galaxy is near the centre C of such a space–time.

However, agreement with these observational relations, while necessary conditions, are by no means adequate criteria for a universe model to be a realistic cosmology. We will first consider criteria of an astrophysical nature, and then those of a philosophical kind.

4.2 BLACKBODY RADIATION

As well as matter and other fields that may be present in the universe model, blackbody radiation will be present with a temperature [7]

$$T(r) = T_0(1+z) = T_0 g(r)^{-1} \tag{9}$$

in a region surrounding the central world line, where T_0 is the temperature of this radiation on the central world line (approximately 3 K). As in the FRW universes, this radiation could be cosmological in origin. More precisely, consider the situation when $g(r)$ decreases monotonically to zero on $(0, R)$. Then at some value r_d of r, the radiation temperature will reach a value T_d corresponding to ionization (with T_d approximately 3000 K). For $0 < r < r_d$, (9) will be valid – in this region the matter and radiation would be weakly interacting; beyond r_d, (9) might still be a good approximation (as in the FRW universes), although the matter

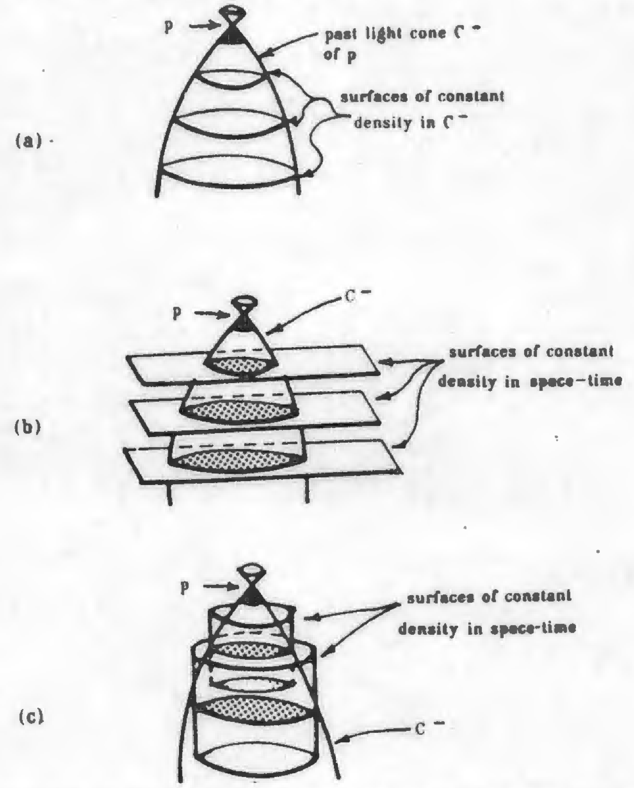

Figure 1. How observations on our past light cone can be extended in either a spatially homogeneous or a static manner. (a) Surfaces of constant density in the past light cone C^- of a point. (b) Extension of these surfaces into space−time as space-like surfaces (all the variation is time variation). (c) Extension of these surfaces into space−time as time-like surfaces (all the variation is space-like variation).

and radiation would be strongly interacting. Just as in the standard universe models, the region beyond (for $r_d < r < R$) would be occupied by a hot cosmic plasma; and this could be the source of the blackbody radiation.

In this situation, the whole radial variation of the matter and radiation content of space−time would be in very close correspondence with the time dependence in the standard models. Thus there would be successive values of r defining spheres around C, where decoupling takes place; nucleosynthesis takes place; pair production takes place; and so on, as the temperature rises (*cf.* [1−3]). In fact each variation that was previously ascribed to a time variation would still take place, but would now be ascribed to spatial variation. The situation is that in general, our observations are made on the past light cone [6], and one can extend these variations off the light cone in a spatially homogeneous way as in the FRW universes, but one can equally do so in a static way, see Fig. 1. Which of these extensions is more viable depends on detailed astrophysical arguments, or the use of specific field equations.

4.3 THE SINGULARITY

Now suppose that R is finite, so that $g(r)$ goes monotonically to zero in a finite proper distance from C. Then the temperature of the background radiation will diverge at this radius, as will the density of this radiation; so one encounters here a matter singularity S,

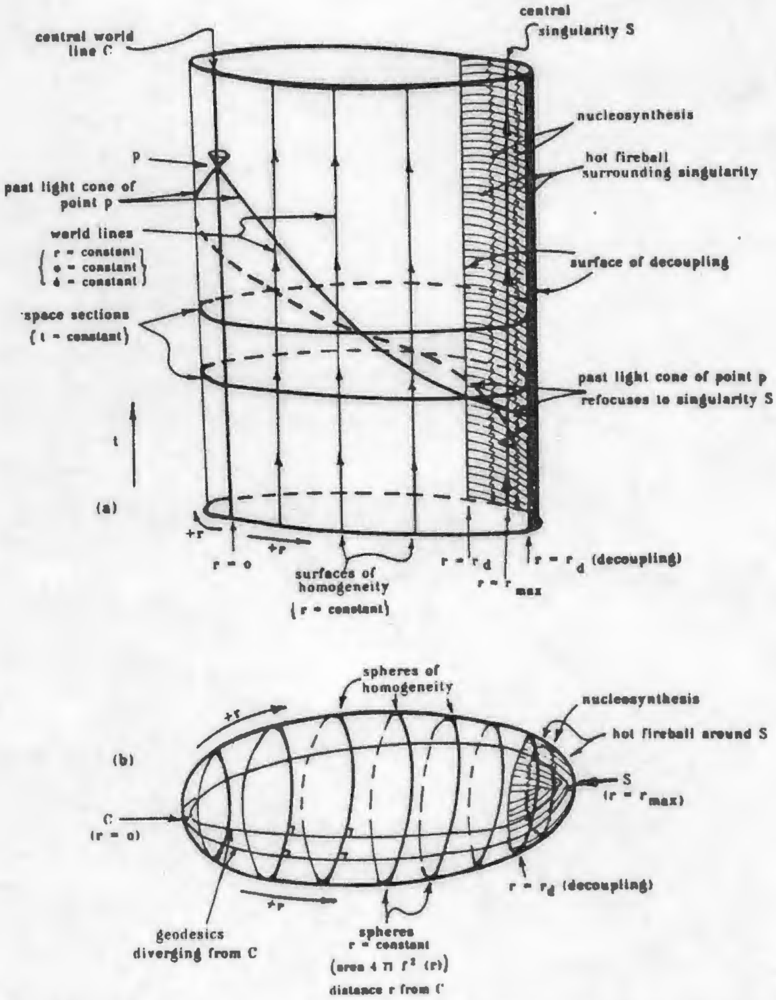

Figure 2. The static, spherically-symmetric (SSS) universe concept. (a) Space–time diagram of a SSS universe. This is a section $\{\theta = \text{constant}, \phi = \text{constant}\}$ of the full space–time. It is a cylinder with a (singular) line S removed. (b) Space section $\{t = \text{constant}, \theta = \pi/2\}$ of a SSS universe. It is eggshell shaped with a (singular) point S removed.

corresponding precisely to the initial singularity in the FRW universes. Just as in those universes, this singularity could be the source of the matter and radiation in the Universe. All past radial null geodesics intersect this singularity, as in the FRW cases; however, in this case all space-like radial geodesics intersect it, whereas in the FRW case, all past directed time-like geodesics intersect it. That is, it surrounds C rather than bounding the past of C. Further (under the conditions we have postulated) it lies within a finite distance from C; so the Universe is spatially finite and bounded, each space section $\{t = \text{constant}\}$ having diameter $2R$. Traversing a radial null geodesic from the singularity, one goes through C after a distance R and arrives back at the singularity after travelling a distance $2R$.

The limiting surface area of the 2-spheres of symmetry, as $r \to R$, is $4\pi f^2(R)$. The most attractive models arise if $f(r)$ is bounded and $f(r) \to 0$ as $r \to R$, $g \to 0$; for in this case, one can regard the singularity as a point singularity in each space section. The space–time would then be spherically symmetric about both the regular centre C and the singular centre S. Space and space–time diagrams of such space–times are given in Fig. 2. While in all cases the fireball around the singularity can be thought of as a sphere *surrounding* the cool central region (in space–time, a cylinder surrounding the central world line C), in this situation it can also be thought of as *surrounded* by the rest of the Universe; from nearby, it would resemble a giant star enclosing the singularity.

A brief digression: the cosmographic models we are discussing arise if the space–time terminates where $g(r)$ goes to zero at a finite value R. There are many other possibilities. First, $g(r)$ could diverge as $r \rightarrow R$; or secondly, $g(r)$ could be finite and positive but $f(r)$ could diverge or go to zero. In the first case the matter and radiation density would go to zero as $r \rightarrow R$ (if normal momentum-conservation equations hold), and continuation beyond this surface would be problematic; the matter-occupied part of the space–time could then be part of a larger space–time (as in the Milne universe). In the second situation, in almost all cases a second singularity would occur here but would be associated with a finite matter density. The exception would be the case where the centrality conditions were again fulfilled (*cf.* (3)), i.e. if as $r \rightarrow R$ one had $f \rightarrow 0, f' \rightarrow -1, f''/f \rightarrow$ finite limit, $g \rightarrow$ finite non-zero limit and $g' \rightarrow 0$. In this case one would arrive not at a singularity but at a second regular centre of the space–time. Then the space-sections $\{t = \text{constant}\}$ would be complete, compact surfaces, just as in the Einstein static universe. Finally R could be infinite: i.e. the space-sections could be complete and non-compact. While these are all cosmographically possible situations, they do not provide an astrophysically plausible setting; so we return to consideration of the case when $g(r) \rightarrow 0$ as $r \rightarrow R$ for a finite value of R. (We will specifically have in mind the case when g goes to zero monotonically; obvious variations would occur if g was not monotonic.)

4.4 NON-EQUILIBRIUM PROCESSES

For the universe model to be viable, as well as permitting a natural explanation of the microwave background radiation, it is necessary that local non-equilibrium processes be able to take place continually. At first sight this presents a major problem in any steady-state universe, and in particular in a static universe, for one might suppose that at most one generation of galaxies could form and die in such backgrounds. However, just as in the original Steady State discussion, this supposition is not correct. The key idea here is that while **u** is the average velocity of matter, this does not mean that systematic motion cannot take place relative to **u**. In fact a continual circulation of matter could take place, with light elements drifting in from near the singularity S towards C, being built up to heavier elements in the galaxies, and then drifting back towards S and being broken down to lighter elements by the fireball surrounding the singularity. The average 4-velocity of the matter could still be **u**; and this process would enable one (as in the Steady State universe) to envisage a continual process of galaxy birth, evolution and decay.

The astrophysical situation in the Universe would be dominated by this process, for it would determine the galaxy formation conditions and so the number density $n(r)$ and average luminosity $L(r)$ of galaxies. There are various possible mechanisms that could cause such a preferential drift of lighter particles 'inwards' (from S towards C) and heavier particles outwards, for example the competition between radiation pressure and gravitational force acting on the particles.

It is therefore possible to continually have non-equilibrium processes in such a universe, although it is static. An essential role would be played by the singularity S; this would have to be time-asymmetric, emitting radiation (and perhaps matter) in the future direction of time and absorbing radiation (and perhaps matter) arriving there from the past. This asymmetric role would be partially expressed in the boundary values that would be imposed on differential equations in this universe at the singularity S (*cf.* [11]). That is, the master arrow of time for local processes in the Universe would be built into the singularity.

4.5 PHILOSOPHICAL CONSIDERATIONS

Models of the sort considered here have not been considered previously because of the assumption – made right at the beginning in setting up the standard models – of a principle of uniformity (the Cosmological or Copernican principle, see, e.g. [1–4, 6, 12, 13]). This is assumed *a priori*, and not tested by observations.

This assumption is made because it is believed to be unreasonable that we should be near the centre of the Universe.

It is certainly unreasonable for our group of galaxies to be near the centre of the Universe, if the implication is taken to be that the Universe has been centred on our presence. However, there is no need for this implication. Rather one should ask: given a universe model of the type proposed, where would one be likely to find life like that we know on Earth? The answer must be, where conditions are favourable for life of this kind. But in the model we are considering, the conditions for life would be most favourable near the centre C where the Universe is cool. Accordingly, it could be highly probable that if life occurred in such a universe, it would occur near the (cool) centre of the Universe. The situation would not be that the Universe had been created in an anthropocentric way; rather, the Universe being in existence, our life would have evolved in the most probable region for life, namely near the centre (this is just the spatial analogue of Carter's statement [14] that life only occurs at favoured times in the history of FRW universes). Accordingly it is eminently reasonable that we should observe such a static universe from somewhere near its centre. (If the model had quite different background radiation temperatures, it would be possible for there to be *no* place with low enough temperatures for life, or for there to be a band of temperatures suitable for life only *outside* the central region (if the temperature is too low, this may also inhibit life).)

If this point is granted, then these universes have many attractive features. They are, like the Steady State universe and the Einstein static universe, unchanging in time; like the Einstein static universe, they are finite in the sense that at any particular time there are only a finite number of galaxies present, and the space sections are bounded. They have no particle horizons, so the well-known communication limitations in the FRW universes near the singularity do not arise. The associated question as to why the FRW universes are spatially homogeneous, is replaced by the question as to why the SSS universes should be static. The answer would be similar, namely it has this symmetry because the singularity dictates it to be so; as one does not have here the causality problem which arises in the FRW universes, it is at least as convincing an answer.

In the FRW universes, the singularity is hidden away inaccessibly in the past; in these universes, it is sitting 'over there' (in a sense, surrounding the Universe), where it can influence, and be influenced by, the Universe continually. (There are no global Cauchy surfaces.) Its continuous interaction with the Universe – it might emit and absorb gravitational waves, neutrinos, electromagnetic radiation, particles – gives a different emphasis to the concept 'creation of the Universe'; for this continuing interaction might be envisaged as the process which keeps the Universe in existence. Questions of a Machian nature (*cf.* [11]) take on a new light in this context, where the boundary conditions for differential equations are governed by the behaviour at the singularity a finite spatial distance away. The question of the concept 'the origin of the Universe' may be amenable to a more meaningful discussion in the SSS situation, than in the usual FRW framework.

4.6 CONCLUSION

It is not our intention to claim that the Universe is actually like such a static spherically-symmetric model. What is claimed is that there are no overwhelming arguments which immediately show such a cosmographic model could not reproduce all the current observations. Such arguments might be obtainable by closer investigation of the astrophysical aspects of these models (for example, whether stratification might take place in the fireball around the singularity, leading to difficulty in producing a blackbody spectrum); but until they are made explicit, cosmographic models which interpret the current evidence as showing that the Universe is expanding are based on the assumption of spatial homogeneity, which is made on philosophical rather than observational ground. The one kind of observation that would rule out a model of this type immediately is the observation of a time variation in some cosmological quantity (e.g. the Hubble constant). We are of course very far from having such observations available.

When Einstein's field equations are used, the situation is quite different. We now turn to this question.

5 Hydrostatic support and equations of state

5.1 HYDROSTATIC SUPPORT

When the field equations of General Relativity are imposed, the energy–momentum conservation equations $T^{ab}_{;b} = 0$ are integrability conditions for these equations. In this section we consider the consequences of the conservation equations.

Energy and momentum conservation only entails specific consequences when one has stated the dynamic properties of the matter considered, by specifying the form of the stress tensor T_{ab} and of the equations of state. We assume (*cf.* Section 2) a perfect fluid stress tensor, i.e.

$$T_{ab} = (\mu + p)\, u_a u_b + p g_{ab}, \quad u^a u_a = -1. \tag{10}$$

In a space–time with metric (1), stress tensor (10) and matter 4-velocity (2), the only non-trivial conservation-equation component is the radial equation of hydrostatic support:

$$g'/g + p'/(\mu + p) = 0, \tag{11}$$

showing how the pressure gradient balances the gravitational potential gradient (the flow lines are non-geodesic when $g' \neq 0$, and the pressure gradient is needed to hold the matter at rest in the potential field). Using (4) this can be rewritten as

$$dz/(1 + z) = dp/(\mu + p) \tag{12}$$

showing how the corresponding gravitational redshifts relate to the change in pressure.

5.2 EQUATIONS OF STATE

The usual approach to the equation of state of the cosmic fluid is to regard the stress tensor T_{ab} as representing the stresses due to matter and radiation, i.e. we take $T_{ab} = T_{M\,ab} + T_{R\,ab}$, where the matter stress tensor $T_{M\,ab}$ can be taken as that of a pressure-free fluid ('dust'), or more accurately as a fluid obeying a relativistic γ law; and the radiation stress tensor $T_{R\,ab}$ as that of a Fermi gas.

If the matter and radiation interact only weakly, one can take each of the stress tensors $T_{M\,ab}$ and $T_{R\,ab}$ as separately conserved, so that equations (11), (12) are obeyed both with

$\mu = \mu_M$ and $p = p_M$ (the matter values), and with $\mu = \mu_R, p = p_R = \mu_R/3$ (the radiation values). If they are strongly interacting, this need no longer be true; one might, for example, have the situation envisaged in the early FRW universes where the radiation variables obey the conservation equation very closely but the matter variables do not (the matter interacts strongly with radiation which completely dominates its behaviour).

In the SSS models one cannot use the non-interacting dust representation for matter, as then the fluid acceleration would be zero and consequently the matter could not remain static in the space–time (equations (11) and (12) could not be satisfied). Rather one can contemplate using the relativistic γ law, where (*cf.* [32])

$$p_M = An^\gamma, \quad \mu_M = Bn^\gamma + p_M/(\gamma - 1), \quad T_M = Cn^{\gamma-1} \tag{13}$$

relate the matter pressure p_M, energy density μ_M, number density n and temperature T_M, and A, B, C are constants. On using equation (12) these equations imply

$$n = \left\{\left(\frac{1-\gamma}{\gamma}\right)\left(\frac{B}{A}\right) + \bar{D}(1+z)\right\}^{1/(\gamma-1)}, \quad T_M = \left(\frac{\gamma-1}{\gamma}\right)\left(\frac{CB}{A}\right)\{Dz + (D-1)\} \tag{14}$$

where $\bar{D}$ and D are constants. To evaluate the constants, we take $B = mc^2$ where m is the particle rest mass and c the speed of light, and $(BC/A) = mc^2/k$, where k is the Boltzmann constant (assuming one can evaluate these constants as for an ordinary gas); then one can evaluate D, by using equation (14), in terms of the value $T_M(0)$ of T_M when $z = 0$. One obtains $D = 1 + T_M(0)/[(\gamma - 1)/\gamma](BC/A)$. For any plausible values, the second term is negligible compared with the first (e.g. take $\gamma = 5/3$ and $m = m_H$ = mass of the hydrogen atom; then $[(\gamma-1)/\gamma](BC/A) = 4 \times 10^{12}$ K). Accordingly, $D \simeq 1$; so $T_M \simeq [(\gamma-1)/\gamma](BC/A)z$. This gives impossibly small values of the redshift when we set $T_M \simeq 10^3$ K, which we might take as corresponding to decoupling. Accordingly – assuming this method of evaluating the constants would be adequate for this 'gas' – the matter pressure by itself cannot produce the required redshifts (unless one could have an equation of state where $\gamma = 1 + \epsilon$, where $\epsilon \simeq 10^{-12}$; we reject this as implausible).

This implies that the matter by itself cannot satisfy equation (12). The alternative would be that the Universe is currently radiation dominated. For the case of radiation, (12) implies the relation

$$p_R = p_R(0)(1+z)^4, \quad \mu_R = 3p_R \tag{15}$$

where $p_R(0)$ is the pressure at the centre (and this also follows from (9)).

5.3 RADIATION-DOMINATED UNIVERSES

Could the Universe be dominated by radiation at the present time? The density of observed electromagnetic radiation is far too low, and there is little prospect of there being large densities of electromagnetic radiation which has escaped detection. However, there could easily be large amounts of other kinds of radiation. In fact an examination of the literature on possible gravitational radiation or neutrino background radiation in the Universe, shows that the most stringent limit on these energy densities is that arrived at by using the predictions of the FRW universe models, i.e. these limits are arrived at by use of the assumption that the Universe is spatially homogeneous. (Essentially, by using the Raychaudhuri equation at the present time, see, e.g. [7].) If we drop this assumption, we can no longer use these limits. The direct observational limits are quite different: one could have energy densities of 10^{-20} g/cm^3 in the form of neutrinos and 10^{-27} g/cm^3 in the form of gravitational waves, without these fields having been detected by direct experiment. Details of the spectral bands where this radiation might reside, and associated references, are given in Appendix A.

It is therefore possible that the Universe is dominated by gravitational waves or neutrinos (*cf.* [15, 16]), leading to an effective equation of state $p = \mu/3$, and corresponding relation (15). In such a universe model, the matter density (plausibly between 10^{-29} and 10^{-31} g/cm^3) could be a small perturbation of the total energy density (possibly up to 10^{-20} g/cm^3), so that the total galaxy number density would not be directly related to the total energy density; rather it would be determined by the galaxy formation processes taking place in the background space–time, whose curvature is determined by the energy density of gravitational waves and neutrinos. What would also be demanded, in this situation, is a mechanism that would result in a matter having, on average, the same 4-velocity as the radiation. This would have to be the result of the forces which control the suggested 'drift' process; there would have to be sufficient coupling between the matter and whatever radiation is present, that the radiation pressure could act on the matter strongly enough to hold it at the same average 4-velocity. This would lead to some difficulties in astrophysical terms, but they may not be insuperable. (And see the further possibilities raised in Section 7.)

6 The General Relativity field equations

6.1 THE FIELD EQUATIONS

The field equations for metric (1) with source tensor given by (10), (2), are:

$$\frac{f''}{f} + \frac{g''}{g} + \frac{f'g'}{fg} = p - \Lambda, \tag{16a}$$

$$2\frac{f''}{f} + \frac{f'^2}{f^2} - \frac{1}{f^2} = -\mu - \Lambda, \tag{16b}$$

$$2\frac{f'g'}{fg} + \frac{f'^2}{f^2} - \frac{1}{f^2} = p - \Lambda, \tag{16c}$$

where Λ is the cosmological constant; the integrability conditions are the conservation equations (11), and the equations are completed by an equation of state $p = p(\mu)$. The equations (16), (11) and the regularity conditions (3) are invariant under the 2-parameter group of transformations

$$g \to \lambda g, \tag{17a}$$

$$r \to \kappa^{-1} r, \quad f \to \kappa^{-1} f, \quad \mu \to \kappa^2 \mu, \quad \Lambda \to \kappa^2 \Lambda, \quad p \to \kappa^2 p, \tag{17b}$$

where λ, κ are non-zero real numbers. In general the solution $\{f(r), g(r), \mu(r)\}$ of (16) depends essentially on four of the six constants $\{f(r_0), g(r_0), f'(r_0), g'(r_0), \mu(r_0), \Lambda\}$ at some value r_0 of r, because of the first integral (16c) and the invariance (17a); however, the existence of the regular centre reduces the arbitrariness to two constants. To see this, consider the situation on the central world line C; there the regularity conditions (3) restrict the constants, so that the solution depends only on two values specified on this line. We will take these as $\{\mu_0 \equiv \mu(0), \Lambda\}$, defining an initial-value plane at C.

Thus in the SSS case, there is even greater 'lack of options' [9] than in the FRW case: given an equation of state, the solution is determined by one arbitrary constant if $\Lambda = 0$, and two if $\Lambda \neq 0$.

6.2 METHOD AND CONCLUSION

Even with the simple equation of state

$$p = \mu/3 > 0 \tag{18}$$

which we assume in view of the argument in the preceding section, it does not seem possible to obtain analytic solutions to these equations in general, or even to obtain the qualitative behaviour of the solutions by use of phase planes (except in the cases $\Lambda = 0; \mu = 0$ or $\Lambda = \mu$). Accordingly we have examined the solutions of equations (16), (18) numerically. A useful simplification results because the equation of state (18), as well as the field equations (16), are invariant under the transformation group (17b); so properties of the solution need only be investigated on a one-dimensional subspace of the initial-value plane, in order to find the behaviour over the whole plane.

As well as integrating the equations numerically, we have investigated the (m, z) relations predicted for these space–times by equation (7), and attempted to fit the predicted curves to the observed data points. The outcome of the investigation, which we briefly report below, is that one cannot fit the observations well. The essential reason for this is the 'regularity' condition (3) at the central world line C, which restricts the solutions so much that we cannot use the available freedom to obtain a good fit (the redshift, which starts off with zero derivative at the origin, initially builds up too slowly as r increases).

6.3 THE NATURE OF THE SOLUTIONS

The general nature of the solutions is shown in Fig. 3. Fig. 3(a) shows the initial-value plane $\{\mu_0, \Lambda; \mu_0 > 0\}$. Numerical integrations of the field equations were carried out for a sample of points in this plane; the nature of the solutions is shown in Fig. 3(b). One can usefully remember here that by the invariance (17b), the nature of the solution is the same on every straight line through the origin in the initial-value plane.

In three specific cases, one can obtain the exact analytic solution. Case (1): $\Lambda = 0 = \mu_0$, is Minkowski space–time. We obtain this as the empty solution with zero redshift and $f(=r)$ increasing linearly without limit. Case (2): $\Lambda \neq 0$, $\mu_0 = 0$, is de Sitter space–time obtained as the zero-mass 'Schwarzschild solution' with non-zero Λ term and with a regular centre (see, e.g. Rindler [17], pp. 208–212). Note that we are not using the congruence of world lines used in de Sitter space–time to obtain the usual Steady State world model, for we are considering as world lines integral curves of a time-like Killing vector field. In this empty solution, when $\Lambda > 0$, f has a pure 'sine' behaviour, increasing to a maximum and then decreasing to zero as $r \to \pi(3/\Lambda)^{1/2}$; and z increases monotonically with r. Case (3): $\Lambda = \mu_0 > 0$, is the Einstein static universe, with zero redshift and with f again having a pure 'sine' behaviour, going to zero again as $r \to \pi(3/2\Lambda)^{1/2}$.

In Case (1), the space sections are complete and unbounded; in Cases (2) and (3), there are regular second centres, so the space sections are complete but bounded (and compact). However, in Case (2), this is an unstable situation; as the redshift diverges, an infinitesimal non-zero value of μ_0 would lead to energy divergences and incomplete space-sections.

In the other cases, we have to rely on numerical integration. Case (4): is when $\Lambda > \mu_0 > 0$. Our integrations indicate that now both z and f increase to maximum values and then decrease, with f going to zero, and z through zero to -1; a second singularity occurs, but it is one where the density goes to zero. The maximum value of f occurs before the maximum value of z. Case (5): is when $\mu_0 > \Lambda > 0$. In this case, f again increases to a maximum and then decreases to zero, but z initially decreases to negative values (corresponding to blue shifts), reaches a minimum, and then increases through zero to $+\infty$. A second singularity

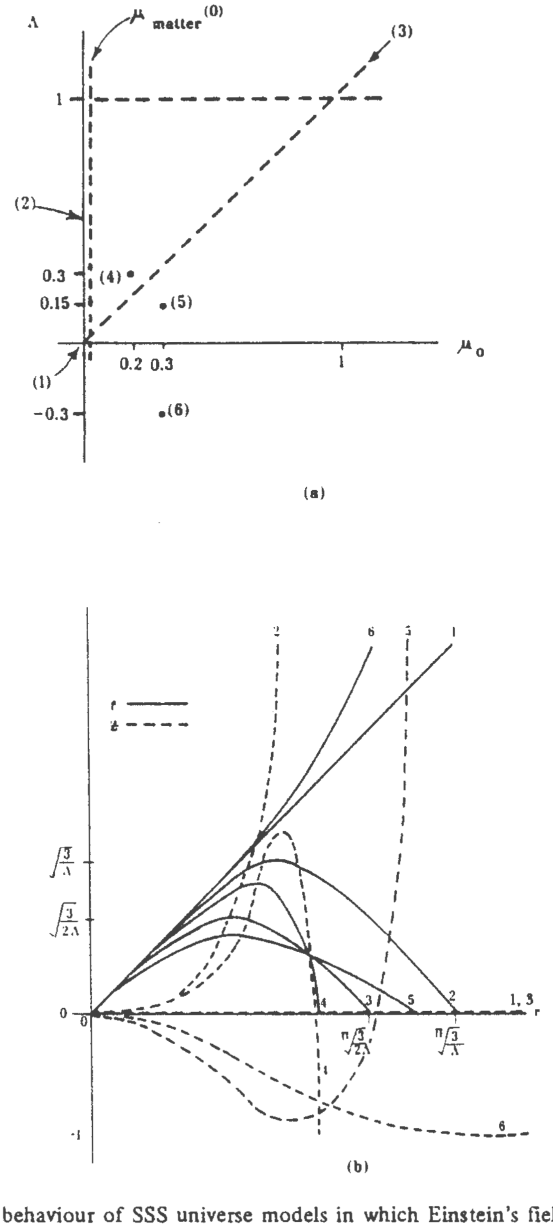

Figure 3. The radial behaviour of SSS universe models in which Einstein's field equations hold with an equation of state $p = \mu/3$. (a) The initial-value plane for these universe models, parametrized by values of the cosmological constant Λ and the central density μ_0. Limits on Λ and μ_0, and the requirement of positive redshifts near the central world line, confine plausible cosmological models to the triangle delineated by broken lines. (Units are $10^{-54}\ \mathrm{cm}^{-2}$; $8\pi G = c = 1$.) (b) Sketches of the radial variation of the observer area distance f and the redshift z, for solutions with initial values indicated by the numbers 1–6 in (a).

occurs here, where the density diverges. Case (6): is when $\Lambda < 0$, $\mu_0 \neq 0$. One can show analytically from the field equations that f increases monotonically, and the space sections are complete. The redshift is negative, decreasing monotonically from zero to -1, and the density decreases monotonically to zero, but does not reach zero for any finite radius (see [8] for analytic proofs when $\Lambda = 0$).

The redshift behaviour just discussed shows that if such universe models are to be realistic, $z > 0$ near the origin will imply that we do *not* get an infinite redshift build up near the singularity. This means that in these models we would have to postulate that the black-body radiation was emitted as such by the singularity. This represents a partial, but not fatal, setback to the picture discussed in the last section.

The condition $\mu > 0$ is an energy condition which implies a focusing tendency on radial null geodesics through the central world line C (one can explicitly check that $d^2f/dv^2 < 0$ where v is an affine parameter). However, one can also check that an actual turnover cannot occur for finite values of r if $\Lambda < 0$, because (16c) prohibits the occurrence of a point where the spatial radial derivative $f'(r)$ is zero. On the other hand, our integrations show that f does turn over when $\Lambda > 0$; in Fig. 4(a), the radial coordinate value r_1 and the value $f_T \equiv f(r_1)$ of f at the turnover, are given as functions of (μ_0/Λ). Thus in all these universes, the past and future light cones of each point on the central world line C refocus either to a singularity or to a second centre. (There need not necessarily be a singularity there, despite the singularity theorems of Hawking & Ellis [4], because the Λ term is sufficiently large that the time-like convergence condition need not be satisfied on all time-like geodesics.) Correspondingly there will be minimum angular diameters for rigid objects in these universes, and the function f (the 'area distance' r_0, see (5)) cannot be used as a global coordinate, for it does not characterize distance uniquely.

The initial behaviour of the redshift near the centre may be seen from the power-series expansion

$$z = \frac{1}{6}(\Lambda - \mu_0)\, r^2 + \frac{1}{360}(15\Lambda - 16\mu_0)(\Lambda - \mu_0)\, r^4 + O(r^6), \tag{19}$$

obtained from equations (11), (16) and (3). This does not necessarily give a good approximation to the actual redshift (the central line may be a bad set of points about which to make such an expansion), but it confirms the initial values of the redshift found in the numerical integration (positive if $\Lambda - \mu_0 > 0$, zero if $\Lambda - \mu_0 = 0$, and negative if $\Lambda - \mu_0 < 0$).

It is the centrality condition which results in (19) being a power series in even terms only. As z starts off with zero derivative, we need to build up redshift rapidly through the quadratic term, in order to achieve something like the observed values. This can be done; the maximum build up is achieved by maximizing $(\Lambda - \mu_0)$. Now local observations in the solar system allow very large values of Λ (see [17]). One can argue [9] that application of the Newtonian limit to small clusters of galaxies limits $|\Lambda|$ to less than $10^{-57}\,\text{cm}^{-2}$, but this is not a very firm limit in view of the uncertainties in the application; if we take the limit as $10^{-54}\,\text{cm}^{-2}$ (see [17, 18]) and take $\mu_0 = 0$, one finds $r \simeq (6z/\Lambda)^{1/2}$ for small z. One could then achieve a redshift of 0.004 at a proper radial distance of 10^{26} cm (i.e. 10^8 light year); and for the required large Λ values, the distance scale would be little changed even if we took μ_0 to be the 'critical density' of $10^{-29}\,\text{g/cm}^3$ (or $10^{-56}\,\text{cm}^{-2}$). Thus one could achieve approximately the same redshift build up as in the FRW universes. The trouble is that the build up is now too fast at larger values of z; and one cannot use the higher order terms in order to get an almost linear behaviour, for they turn the curve right over. Consequently in these universes the attempt to use redshift to characterize distance fails (one value of z corresponds to two values of r). In addition there exist maximum or minimum redshifts in these universes; Fig. 4(a) shows the maximum value z_T of the redshift in those universes which start with a positive redshift (i.e. where $\Lambda > \mu_0 > 0$), and the r value r_2 at the turnover point. For smaller values of Λ, one would get different scales (one could reach $z \simeq 0.004$ at much larger values of r; this would necessarily be the case if one adopted the Λ limit in [9]).

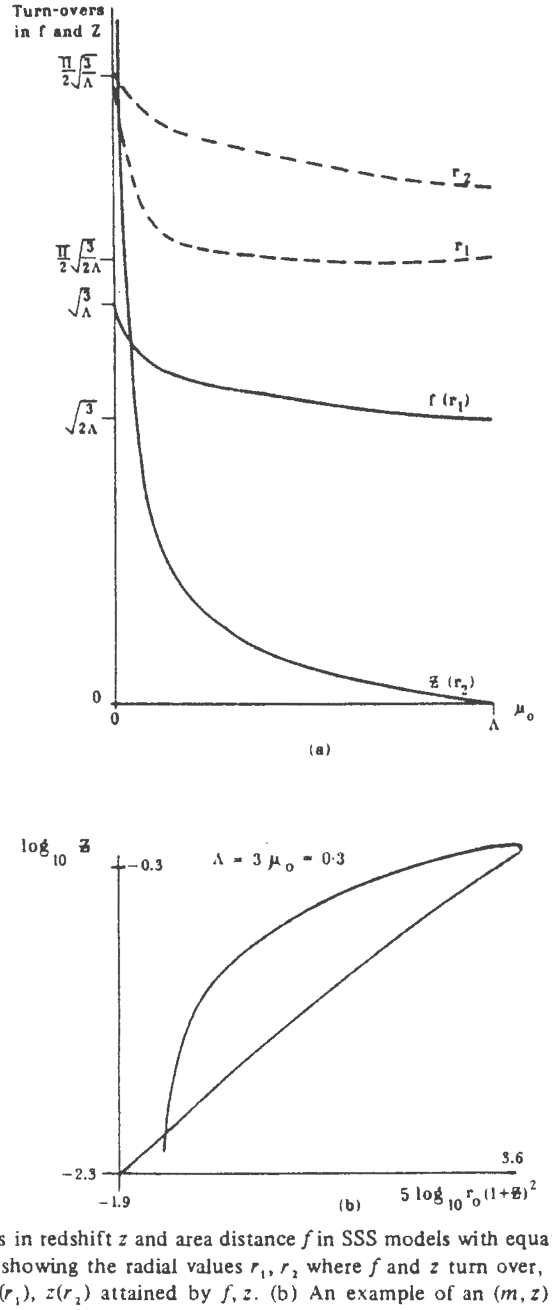

Figure 4. The turnovers in redshift z and area distance f in SSS models with equation of state $p = \mu/3$ and $\Lambda > \mu_0 > 0$. (a) Curves showing the radial values r_1, r_2 where f and z turn over, as a function of μ_0; and the maximum values $f(r_1)$, $z(r_2)$ attained by f, z. (b) An example of an (m, z) curve for these models, showing how the turnovers in z and f lead to a turnover in the (m, z) curve.

6.4 BEST FIT

In order to fit the predicted to the actual (m, z) curve, we of course demand that $z > 0$ for nearby galaxies; so we must have

$$10^{-54}\,\mathrm{cm}^{-2} \gtrsim \Lambda > \mu_0 \gtrsim 10^{-58}\,\mathrm{cm}^{-2}, \tag{20}$$

where the outer inequalities correspond to upper possible limits on Λ and lower limits on the observed matter in the Universe, and the middle inequality leads to positive redshifts near C. This region in the initial-value plane is the area where we may hope to find reasonable universe models (see Fig. 3(a)). The general behaviour of the models is that described above as Case (4).

In these models, the turnover of f and then z results in the (m, z) curve turning over at $z = z_T$; the kind of (m, z) curve that results is shown in Fig. 4(b). The turnover divides the (m, z) relation into two branches, that for $0 < r < r_2$, i.e. for small distances, and that for $r > r_2$, i.e. for larger distances (but the same z values). Our integrations indicate that the slope of the branch $r > r_2$ is steeper than that of the branch $r < r_2$, for small z (*cf.* Fig. 4(b)). They also indicate that the latter slope exceeds the slope defined by the observed (m, z) points, so we attempted to fit the observations for the branch $r < r_2$. Since under (17),

$$\left.\begin{aligned} &z \to z, \\ &5 \log_{10} r_0(1+z)^2 \to 5 \log_{10} r_0(1+z)^2 - 5 \log_{10} \kappa \end{aligned}\right\}, \tag{21}$$

we are free to shift the (m, z) curve along the m axis, with the corresponding shift $\mu_0 \to \kappa^2\mu_0$, $\Lambda \to \kappa^2\Lambda$ in the initial-value plane (which may move the model out of the region (20)). Observational (m, z) values for distant galaxies were taken from Sandage, Kristian & Westphal [19]. For fixed Λ and a range of μ_0 values such that $0 < \mu_0 < \frac{1}{5}\Lambda$ (note that $\mu_0 > \frac{1}{5}\Lambda \Rightarrow z_T < 1$), the numerical (m, z) curve was interpolated to give m values at the same z values as the observed points; and then shifted, by use of (21), so that the sum of squares of differences between observed and numerical values was a minimum. This minimum value is plotted against μ_0 in Fig. 5(a). The best fit is obtained for the de Sitter solution where $\Lambda > 0$, $\mu_0 = 0$ (and for which analytic expressions may be obtained for the (m, z) curve); and this is not an acceptable fit to the observations (see Fig. 5(b)). Thus we are unable to obtain a good fit to the observations, if $p = \mu/3$.

7 Alternatives

The argument presented has shown that current (m, z) observations cannot be fitted by the SSS cosmological models discussed, as the slope of the observed points is too shallow relative to the curve predicted by these models. One can, however, query what alternatives there might be to some of the assumptions made in the argument.

7.1 EQUATION OF STATE

Substantially different results might be obtained if the *equation of state* of the cosmological fluid had been incorrectly estimated (the SSS universe models are much more sensitive to the equation of state, than the FRW universes). In the situation we have in mind, it is quite plausible that the equation of state would vary with radius. If either neutrinos or electromagnetic fields are significant on cosmological scales, one might have anisotropic fluid pressures; and considerable differences from the standard fluid model are conceivable in a gas where a long-range force, such as gravity, is dominant (see the 'gravithermodynamics' discussion by Saslaw [20]). The consequences might simply be that evaluation of the constants in the equation of state was incorrect, or it might be that a substantially different model of matter was required (an 'imperfect fluid' or a plasma description, for example). (If neutrinos dominate, one can have what is essentially a long-range interaction resembling elasticity.) It would, of course, be possible to take a cosmographic model which fitted the observed data and run the field equations in reverse in order to deduce the matter behaviour required to give the observed (m, z) curve. However, such 'matter' would probably have

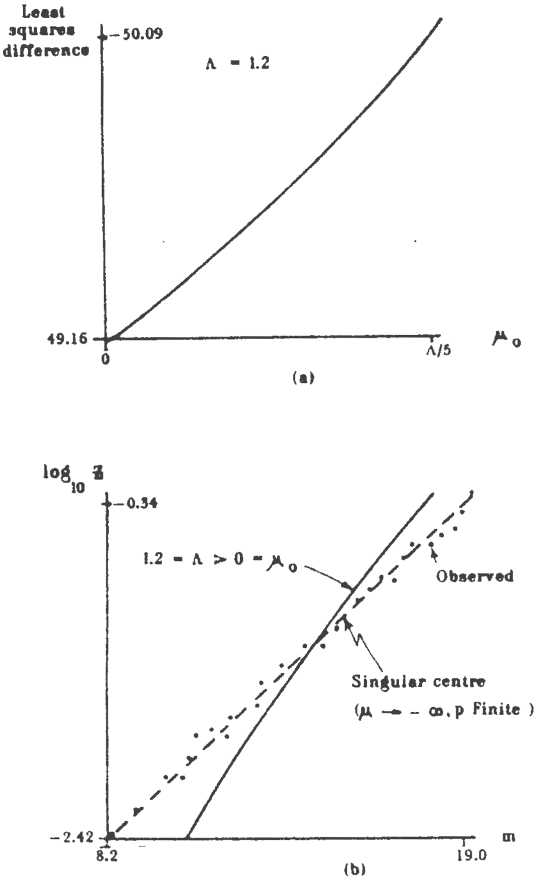

Figure 5. The observational fit to the (m, z) curve, for models with $p = \mu/3$ and $\Lambda > \mu_0 > 0$. The observations are from Sandage *et al.* [19]. (a) Curve showing variation with μ_0 of the least-squares error of the predicted curve (for $r < r_2$), relative to the observations. (b) The best-fit theoretical (m, z) curve, for $\Lambda = 1.2$, $\mu_0 = 0$; and the observed points. The curve has been shifted along the m axis by an amount corresponding to $\kappa = 1.26$ ($\rightarrow \Lambda \rightarrow 2.57\, \Lambda$; see Section 6). We have taken an absolute magnitude $M \equiv -2.5 \log_{10} L = -23.16$. Also shown is the fit that can be obtained if one drops the 'regular centre' condition and lets $\mu \rightarrow -\infty, p \rightarrow$ finite value at the centre (see Section 7).

unrealistic physical properties. It is not clear that any alterations in the equation of state representing realistic matter properties, would enable the redshift to build up fast enough.

7.2 CALIBRATION AND RADIAL EVOLUTION

One should, however, be cautious here. The whole *calibration* on which the interpretation of the usual (m, z) curves is based, is made on the understanding that the FRW universe model is basically correct. In the situation we have in mind, the redshift values might well correspond to much larger distances than in the FRW universe models; so the whole calibration system for magnitudes might work out substantially differently in the SSS universes. A related property is that *radial changes* in galactic properties (corresponding to the 'evolutionary corrections' in the FRW universes) might be significant at smaller redshifts than in the FRW universe models. A careful investigation would be needed to follow all the ramifications of the change of scale that might result if the SSS interpretation were adopted.

7.3 OTHER FIELD EQUATIONS

More radically, one could investigate the consequences of adopting *different field equations.* One could even seriously consider the possibility that matter and radiation momentum-conservation would no longer hold, with such altered equations; for some sort of cosmological force field could enter modified momentum-conservation equations in an appropriate way (this would be the spatial analogue of the Steady State universe situation, where the 'c-field' results in energy conservation being abandoned).

It is not clear how good a fit such changes would enable one to achieve. The problem remains that of obtaining a reasonably fast build up of redshift, relative to the change in area distance as one moves away from the centre, without then having a serious 'overshoot'.

7.4 SINGULAR CENTRES

One way which one might achieve a good fit is by weakening the *centrality conditions.* Specifically, if one were to allow the function $g(r)$ to have discontinuous first derivatives at $r = 0$ (by dropping the condition $g'(r) \to 0$ as $r \to 0$), and perhaps weakened the conditions on $f(r)$, one could hope to achieve, for a galaxy very near the centre, a much better fit for the (m, z) curves. For example, if we take

$$g(r) = 1 - H_0 r + ar^2 + 0(r^3), \quad f(r) = r + \{(\alpha - 1) H_0 - aH_0^{-1}\} r^2 + 0(r^3),$$

where $\alpha \equiv \frac{1}{2}(1 - q_0)$ and $a \equiv -\frac{1}{2}(1 + 2q_0) H_0^2, H_0$ and q_0 constants, one obtains

$$5 \log_{10} r_0(1 + z)^2 = -\log_{10} H_0 + \log_{10} z + (\tfrac{1}{2} \log_{10} e)\{1 - q_0\} z + 0(z^2),$$

which is precisely the FRW form. These functional forms imply (by (16)) an 'equation of state' such that $p \to$ finite value but $\mu \to -\infty$ as $r \to 0$. Fig. 5(b) shows a plot of this (m, z) relation for $\alpha = 0, M = -23.16, H_0 = 55$ km/(s Mpc) $= 60.5 \times 10^{-30}$/cm. While the (m, z) fit is good, obviously the physical situation implied leaves much to be desired. Other singular behaviours at the centre we might consider do not necessarily involve an infinite density singularity there; but the real problem would be that any such supposition would make our space–time position too special. One would have to be extremely close to the central line in order to see almost isotropic redshifts. Having chosen such a space, it would be difficult to justifiy such a special position for our Galaxy in terms of the argument applied previously.

7.5 DOPPLER OR LOCAL GRAVITATIONAL REDSHIFT CONTRIBUTIONS

Perhaps the most interesting possibilities involve new contributions to the redshift. First, one could have a static universe (whose curvature was governed by a static radiation field) in which the *matter was non-static*, i.e. in which the matter moved as a test field in the static background space–time. Then part of the observed redshift would be a Doppler shift. This corresponds exactly to the situation in both the Milne and Steady State universes (in each case the space–time is locally static but the matter non-static). This proposal – that the average 4-velocity of matter should not exactly align with that of the background radiation field – is perhaps more plausible than the strict SSS proposal that these 4-velocities should be the same. To make such a suggestion plausible would require a detailed investigation of the matter 'drift' mechanism, and associated galaxy-formation processes; it is quite possible that problems would be revealed by such an investigation. The second possibility would be to pursue those lines of argument which suggest that at least a substantial part of each red-shift observed, is a *locally generated redshift* (e.g. due to gravitational redshift at the source). It might well be that a SSS universe would provide at least as plausible a background geometry for such a situation, as do the FRW universes; the situation could then be different from that which we have considered up to now, in that in this case one might be able to have plausible matter dominated SSS models.

7.6 CONCLUSION

None of these alternatives are immediately compelling. However, they do show that there are a variety of possibility which need consideration before the SSS models are categorically ruled out.

8 Implications

We have seen that although it is cosmographically possible to construct a SSS universe model, it is difficult (though not proved totally impossible) to construct plausible cosmological models of this kind. Although we have looked only at exactly isotropic universe models, it is very plausible that deviations from isotropy would not change the situation; i.e. if our observations eventually show there are deviations from isotropy on a cosmological scale, there would again be problems in constructing a static model for this situation. (One should note here the unproven, but plausible, assumption that isotropic (m, z) and (N, z) relations will only occur in an isotropic space–time.)

What, then, has been gained by this investigation?

8.1 EXPANSION OF THE UNIVERSE

First, we have given *an indication as to how one can deduce the Universe is expanding*, on the basis of present observational evidence. It may be claimed that better or briefer arguments could be given. We would welcome such arguments. What we do claim, is that such arguments are necessary if one is to make clear statements about the observational basis of cosmology. Until they are made explicit, the interpretation of the current evidence as showing that the Universe is expanding is based on the assumption of spatial homogeneity, which is made on philosophical rather than observational grounds.

8.2 INHOMOGENEOUS EXPANDING UNIVERSES

Secondly, our argument has thrown up certain other features of interest in the study of observational cosmology. Although the FRW models have won wide acceptance, there are various problems associated with them, partly related to their spatial homogeneity (e.g. the problem of galaxy formation from exactly homogeneous initial conditions), and partly to their isotropy (evidence has been led that there might be anisotropy in the Hubble diagram [21], or some kind of hierarchical structure or other inhomogeneity [22]). We have pointed out that many of the features of a FRW universe could occur in an inhomogeneous static universe model with our Galaxy somewhere near the centre, but had difficulty fitting a detailed static model.

However, this is a very extreme case. Many new features occurred in the static models, which suggest a new spectrum of possibilities that could be considered even if one does not claim the Universe is static. What may well be worth pursuing is the possibility that there are *inhomogeneous spherically-symmetric universe models which are expanding, but retain some of the interesting features of the SSS universe models*. Their properties could include features of both the FRW and the SSS models. We consider briefly some of the features that could arise in this case.

8.2.1

In such expanding models, the major part of the redshift as observed from near the centre could be due to the expansion, as in the FRW models. Accordingly, there would be no difficulty in fitting the (m, z) curve predicted for an observer near the centre, to current observations. Further one could again advance plausible arguments as to why we should observe such a universe from near its centre (as in Section 4 above). There would be no more difficulty in fitting the number counts of galaxies, than in the FRW models (where one has to hypothesize rapid galactic evolution in order to get a plausible model).

8.2.2

The best current limits on the energy densities of various matter and radiation components are based on deductions from the field equations – essentially, on evaluation of Raychaudhuri's equation at the present time – which are made on the assumption that the Universe is spatially homogeneous. This applies particularly to coagulated matter (rocks), neutrinos and gravitational waves. In expanding inhomogeneous universe models there are weaker limits on the possible densities of these components; one can again attempt to evaluate Raychaudhuri's equation at the present time, but it becomes difficult to bound the acceleration term in this case. One might conceivably have many times the critical density present in these forms of matter and radiation. For the same reason one could contemplate a much larger Λ term than in the FRW universes.

8.2.3

The redshift might correspond to rather different distance scales than in the FRW case. Moreover, it might fail as a distance indicator, as one might have the same redshift for two very different proper distances. Associated with this, if it occurred, would be the existence of maximum redshifts for observed objects, and possibly a turnover in the (m, z) curve. However, if this were to occur, the 'natural' explanation for the microwave background radiation would be in danger, unless the redshift at some later radius again began increasing to large values. This would not be a total disaster, as the singularity would presumably be quite capable of producing blackbody radiation by itself; but this would not be as attractive an explanation as before.

8.2.4

An observer at the centre of a SSS universe would measure exactly isotropic microwave background radiation (see Section 4 above). What is rather unexpected is that *all* observers outside the hot fireball in a SSS universe would measure exactly isotropic background radiation; that is, isotropy of the background radiation would in no way imply that the observer was at the centre of the Universe. In fact the redshift measured by an observer at coordinate value r for radiation emitted at coordinate value r_e, would be given by (*cf.* equation (4))

$$1 + z = g(r)/g(r_e),$$

so the temperature of blackbody background radiation (emitted at temperature T_e at coordinate value r_d) measured by an observer at r, would be (*cf.* equation (9))

$$T(r) = T_e g(r_d)/g(r).$$

The important thing to notice is that this holds independent of the direction of observation; the observed temperature depends only on the gravitational potential difference between the point of emission and the point of reception. Accordingly one could not, in the SSS universe models, use isotropy of the microwave background radiation as evidence that the Universe was spherically symmetric about one's position (since the radiation would be isotropic no matter where one was situated). (This would not be true for the galactic redshifts, i.e. exactly isotropic (m, z) observations would imply one was exactly at the centre.)

This opens up the possibility that the same could be true for some expanding spherically-symmetric universes; it might be that an off-centre observer – who would in general measure anisotropic galactic redshifts – would observe isotropic blackbody background radiation.

8.2.5

Perhaps the most interesting feature of the SSS models which might carry over, are some of the essential features of the singularity structure. We have briefly discussed how in the SSS models one could have a singularity situated (for all time) 'over there' around the Universe, rather than in the past; and how then this singularity could be actively interacting with the Universe, absorbing and emitting radiation and matter, controlling the boundary conditions for differential equations (and so controlling stability properties), and so on. One would then also have the 'thermal history' taking place in a spatial rather than a time direction, with element formation taking place continuously in the hot fireball, pair production taking place continuously, and so on. One can conceive of much the same qualitative features occurring in expanding spherically-symmetric models; i.e. one might still have the essential features of Fig. 2 remaining true, but now with radii a function of time (because the distances between the galaxies vary).

We have not been able to find a detailed model of this kind because of the difficulty of integrating the field equations in the case of a perfect fluid with a realistic equation of state. However, we have been able to find a model (an exact solution of Einstein's equations) which has some of these features, in which the matter component is pressure-free matter ('dust'). Because the pressure vanishes it is obviously not realistic at high densities, and this prevents us from demonstrating all the features we would like; nevertheless it shows that the situation we contemplate might be possible.

Details of the solution – adapted from Bondi's discussion [23] – are given in Appendix B. It is an exact spherically-symmetric inhomogeneous solution of the field equations. The central world line is given by $r = 0$; on this world line, the expansion is just like that of a FRW universe with a large Λ term, which collapses from infinity to a minimum finite radius and then re-expands, without a singularity ever occurring on this world line. The same behaviour will occur on a comoving sphere of particles surrounding the origin; so there is a sphere of matter around the central particle which never runs into an initial (or final) singularity. However, for each proper time t, the space sections orthogonal to the fluid flow end at a singularity where the matter density diverges, within a finite proper distance. That is, the matter in the Universe is, for all time, surrounded by a singularity 'over there', which raises all the possibilities mentioned previously.

Because of the lack of pressure in this model, it cannot include some other features we might look for (e.g. the model cannot be made almost static near the singularity). Nevertheless it does indicate there are a range of possibilities similar to those suggested by the SSS models, which could be looked at by investigating expanding, spherically-symmetric, inhomogeneous models.

9 Conclusion

We have given an indication of how to show from present observations, without making unverifiable *a priori* assumptions about space–time geometry, that the Universe is expanding. There may be better methods of proving this result, than those we have given; our contention is that *some* such proof must be given, for otherwise the statement that the Universe is expanding cannot be regarded as an observationally-based statement.

In addition, our investigation throws up various other points of interest. In particular, we given an example which shows that even if the microwave background radiation in a cosmological model is measured to be exactly isotropic about an observer, this does not imply that the space–time is isotropic about him. Further we have shown there exist singularity structures in expanding but inhomogeneous universe models, which are completely different in nature from those in the FRW universes, but which give the same observational predictions.

While inhomogeneous, expanding cosmological models have been suggested before (see, e.g. [24]), the singularity structures in such spaces (see, e.g. [31]) have not been explored in the way suggested here.

Acknowledgments

We should like to thank the Research Committee of the University of Cape Town for a research grant. In addition, G. Ellis would like to thank the Max Planck Institute for Physics and Astrophysics, Munich, for hospitality and support; S. D. Nel, to acknowledge support of an AECI Research Fellowship; and R. Maartens, of a UCT scholarship.

We wish to record our thanks to the members of the relativity groups at the University of Cape Town and the Max Planck Institute, particularly J. Perry, for many useful discussions.

References

[1] Sciama, D. W., 1971. *Modern cosmology*, Cambridge University Press.
[2] Peebles, P. J. E., 1971. *Physical cosmology*, Princeton University Press.
[3] Weinberg, S., 1972. *Gravitation and cosmology*, Wiley and Sons, New York.
[4] Hawking, S. W. & Ellis, G. F. R., 1973. *The large scale structure of space–time*, Cambridge University Press.
[5] Ehlers, J., 1961. Beiträge zur relativistischen Mechanik kontinuierlicher Medien, *Abh. Akad. Wiss. u. Lit. Mainz, Mat. Nat. Kl.*, 11.
[6] Ellis, G. F. R., 1975. *Q. Jl R. astr. Soc.*, 16, 245.
[7] Ellis, G. F. R., 1971. Relativistic cosmology, in *General relativity and cosmology*, ed. Sachs, R. K., Proceedings of the XLVII Enrico Fermi School of Physics at Varenna, Academic Press, New York.
[8] Schmidt, B., 1966. *Kugelsymmetrische statische Materielösungen der Einsteinschen Feldgleichungen*, Diplomarbeit, Hamburg University.
[9] Misner, C. W., Thorne, K. S. & Wheeler, J. A., 1973. *Gravitation*, Freeman and Co., San Francisco.
[10] Synge, J. L., 1960. *Relativity: the general theory*, North-Holland, Amsterdam.
[11] Ellis, G. F. R. & Sciama, D. W., 1972. Global and non-global problems in cosmology, in *General relativity*, ed. O'Raifeartaigh, L., Clarendon Press, Oxford.
[12] Bondi, H., 1960. *Cosmology*, Cambridge University Press.
[13] Ellis, G. F. R. & Harrison, E. R., 1974. *Comm. Astr. Space Sci.*, VI, 23.
[14] Carter, B., 1974. Large number coincidences and the anthropic principle in cosmology, in *Confrontation of cosmological theories with observational data*, ed. Longair, M., IAU Symp. 63, D. Reidel, Holland.
[15] Rees, M. J., 1971. *Mon. Not. R. astr. Soc.*, 154, 187.
Dautcourt, G., 1974. Low frequency gravitational waves in cosmology, in *Confrontation of cosmological theories with observational data*, ed. Longair, M., IAU Symp. 63, D. Reidel, Holland.

Dautcourt, G., 1974. *Astr. Nachr.*, 295, 6.
Burke, W., 1975. *Astrophys. J.*, 196, 329.
[16] de Graaf, T., 1970. *Astr. Astrophys.*, 5, 335.
[17] Rindler, W., 1969. *Essential relativity*, van Nostrand Reinhold, New York.
[18] Rindler, W., 1969. *Astrophys. J.*, 157, L147.
[19] Sandage, A., Kristian, J. & Westphal, J. A., 1976. *Astrophys. J.*, 205, 688.
[20] Saslaw, W., 1968. *Mon. Not. R. astr. Soc.*, 141, 1.
Saslaw, W., 1969. *Mon. Not. R. astr. Soc.*, 143, 437.
Saslaw, W., 1970. *Mon. Not. R. astr. Soc.*, 147, 253.
[21] de Vaucouleurs, 1976. *Astrophys. J.*, 205, 13.
[22] Rees, M. J. & Sciama, D. W., 1967. *Nature*, 213, 347.
Rees, M. J. & Sciama, D. W., 1968. *Nature*, 217, 511.
de Vaucouleurs, G., 1970. *Science*, 167, 1203.
[23] Bondi, H., 1947. *Mon. Not. R. astr. Soc.*, 107, 410.
[24] Omer, G. C., 1949. *Astrophys. J.*, 109, 164.
Kantowski, R., 1969. *Astrophys. J.*, 155, 1023.
Bonnor, W. B., 1972. *Mon. Not. R. astr. Soc.*, 159, 261.
Varshni, Y. P., 1976. *Astrophys. Space Sci.*, 43, 3.
Mavrides, S., 1976. *Mon. Not. R. astr. Soc.*, 177, 709.
[25] Marx, G. & Szalay, S. A., 1973. *Int. Centre Theor. Phys.*, Trieste, IC/73/71.
Cowsik, R. & McClelland, J., 1973. *Astrophys. J.*, 180, 7.
[26] Bahcall, J. N. & Davis, R., 1971. *Phys. Rev. Lett.*, 26, 662.
[27] Papini, G., 1974. *Can. J. Phys.*, 52, 880.
[28] Press, W. H. & Thorne, K. S., 1972. Gravitational wave astronomy, *A. Rev. Astr. Astrophys.*, 10, 335.
[29] Weber, J., 1967. *Phys. Rev. Lett.*, 18, 498.
[30] Gibbons, G. W. & Hawking, S. W., 1971. *Phys. Rev. D.*, 4, 2191.
[31] Eardley, D., Liang, E. & Sachs, R. K., 1972. *J. Math. Phys.*, 13, 99.
[32] Ellis, G. F. R., 1967. *J. Math. Phys.*, 8, 1171.
Ellis, G. F. R. & Stewart, J. M., 1968. *J. Math. Phys.*, 9, 1072.
[33] Stabell, R. & Refsdal, S., 1966. *Mon. Not. R. astr. Soc.*, 132, 379.

This paper constitutes an expansion and further development of the essay 'Is the Universe Expanding?', awarded the 2nd prize by the Gravity Research Foundation in 1977 (*J. Gen. Rel. Grav.*, 1978).

Appendix A: density limits on neutrino and gravitational-wave background radiation

A.1 NEUTRINO BACKGROUND

First, we consider the limits on the energy density of a massless neutrino (ν_e) background (the situation might be substantially different if one considered neutrinos with non-zero rest mass, *cf.* [25]).

Within a FRW background, one can use the best estimates of the deceleration parameter q_0 to put limits on the neutrino background, but in the SSS framework, we are concerned with direct experimental limits. We will consider here specifically the energy densities that might reside in two energy ranges, (i) very-low energy ranges, $E_\nu \lesssim$ few eV; (ii) intermediate energy, $1\,\text{MeV} \lesssim E_\nu \lesssim 100\,\text{MeV}$.

(i) In the low-energy band, a completely degenerate ν_e Fermi gas could exist without having been detected [3, 16]. Taking the conservative upper limit $\epsilon_F \lesssim 1\,\text{eV}$ on the Fermi energy, one could have present energy densities as high as $10^{-20}\,\text{g cm}^{-3}$ [16].

(ii) In the intermediate-energy band, limits are placed on the neutrino energy density by the reaction $\nu + {}^{37}\text{Cl} \rightarrow {}^{37}\text{Ar} + e^-$, used in the solar-neutrino detection experiment [26]. These experiments, together with a power-law approximation to the E_ν dependence of the absorption cross-section, indicate that there could be energy densities as high as $10^{-27.6}\,\text{g cm}^{-3}$ [26].

Obviously there could be large densities residing in other energy bands; however, the limits given are useful, as the experiments can be taken as giving definite upper limits for these two bands.

A.2 GRAVITATIONAL-WAVE BACKGROUND

Secondly, we consider limits on the energy density of gravitational waves in the Universe, which result from direct experiment (see [27, 28] for reviews). Table 1 gives limits on the energy density of gravitational waves that might reside in frequency bands centred on various frequencies, without having been detected; and the nature of the detector.

Table 1.

Detector	Frequency (s^{-1})	Upper limit ($g\,cm^{-3}$)	Reference
Earth (normal mode)	$\sim 10^{-4}$	$\sim 8\times 10^{-31}$	[29]
Earth–Moon	~ 1	$\sim 10^{-13}$	[9]
Weber's cylinder	$\sim 10^{3}$	$\sim 10^{-27}$	[30]

Appendix B: spherically-symmetric dust solutions

B.1 BONDI'S SOLUTION

We follow Bondi's method [23] of integrating a spherically-symmetric dust solution of Einstein's equations. The metric will be taken to be

$$ds^2 = -dt^2 + X^2(r,t)\,dr^2 + Y^2(r,t)(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{B1}$$

with the 'dust' stress tensor

$$T_{ab} = \rho u_a u_b, \quad u^a = \delta^a_0, \quad \rho(r,t) > 0. \tag{B2}$$

Suppose that at some time t_0, one has $\rho(r,t_0) = \rho_0(r)$, $Y(r,t_0) = Y_0(r)$. The solution to the field equations is given by the functions $X(r,t)$, $Y(r,t)$ that obey

$$\dot{Y}^2(r,t) = \Lambda Y^2/3 + W^2(r) - 1 + 2Z(r)/Y, \tag{B3}$$

$$X(r,t) = \frac{1}{W(r)}\,\partial Y(r,t)/\partial r, \tag{B4}$$

with the initial value of Y given by Y_0, where $W(r)$ is an arbitrary non-zero function of r, and

$$Z(r) = \int_0^r \tfrac{1}{6}\rho_0(r)[Y_0^3(r)]'\,dr. \tag{B5}$$

(Here, $\dot{}$ means '$\partial/\partial t$' and $'$ means '$\partial/\partial r$'). The matter density ρ is then given by

$$\rho(r,t) = \rho_0(r)[Y_0^3(r)]'/[Y^3(r,t)]'. \tag{B6}$$

The geometric and physical meaning of the space–time are discussed in [23] (and see also [32]).

The 'centrality' condition for the metric (B1) is

$$Y(r,t) \to 0, \quad \frac{1}{X(r,t)}\,\partial Y(r,t)/\partial r \to 1 \quad \text{as} \quad r \to 0; \tag{B7}$$

in view of (B4), this condition that $r = 0$ is a regular centre C for the space–time can be written

$$Y(r,t) \to 0, \quad W(r) \to 1 \quad \text{as} \quad r \to 0. \tag{B8}$$

These conditions are preserved under the propagation equations (B3). Accordingly, one can obtain a solution to the field equations for metric (B1) with matter source (B2), a regular centre at $r = 0$, and the desired initial values of the function Y and the density ρ, by (A) choosing functions $\{Y_0(r), W(r), \rho_0(r)\}$ which are all positive and such that $dY_0(r)/dr \neq 0$ for all r, and $Y_0(r) \to 0$, $W(r) \to 1$, $\rho_0(r)' \to 0$ as $r \to 0$; (B) finding $Z(r)$ from (B5); (C) solving (B3) for $Y(r,t)$ with $Y_0(r)$ as the initial value; and, (D) finding X and ρ from (B4), (B6). Note that in the final solution, the coordinate r may be changed by a transformation $r \to \bar{r} = \bar{r}(r)$ which preserves the metric form (B1) but will change the specific forms of the functions X, Y, ρ.

B.2 ROBERTSON–WALKER INITIAL CONDITIONS

If one chooses, for r in some interval I,

$$\left.\begin{aligned} W^2(r) &= 1 - KY_0^2(r), \\ \rho_0(r) = \rho_0 &\Leftrightarrow Z(r) = MY_0^3(r) \end{aligned}\right\} \tag{B9}$$

where K, ρ_0 and M are constants ($M = \rho_0/6$), then for these values of r the solution is an exact FRW solution with $Y(r,t) = R(t)\,Y_0(r)$ and

$$R^{\cdot 2}(t)/R^2(t) = \Lambda/3 - K/R^2(t) + 2M/R^3(t), \tag{B10}$$

which is the Friedmann equation. (It is important to note that, as in Newtonian theory, if C is a centre of spherical symmetry 'the spherical shells of matter further away from C than a particle P do not affect the motion of P at all' (Bondi [23], p. 416). Accordingly if the central (comoving) sphere of matter for $|r| < R$ starts out as a FRW solution, it remains a FRW solution for *all* time.) If the initial values do not obey (B9), the solution does not develop as a FRW space–time (although along each world line, the evolution equation for Y is still the Friedmann equation (B3)).

B.3 THE CHOSEN SOLUTION

Our procedure now is to choose $Y_0(r) = r$, which amounts to choosing r as an area coordinate, and to define $W(r) = 1 - K(r)\,r^2$. Choose positive finite numbers, R_1, R_2 ($0 < R_1 < R_2$). Now choose $\{K(r), \rho_0(r)\}$ as C^∞ functions for $0 < r < R_2$, such that the following conditions hold: $\rho_0(r) > 0$, $W(r) > 0$; both $K(r)$ and $\rho_0(r)$ are constant for $0 < r < R_1$; $\rho_0(r)$ is non-constant for $R_1 < r < R_2$, and $\rho_0(r)$ diverges as $r \to R_2$. We adjust the constants K, ρ_0 and Λ on $0 < r < R_1$ such that the model is an oscillating FRW universe without singularity (an M_2 model in the notation of Stabell & Refsdal [33], or a C model in the notation of [17]).

This is now a model of the required type. The sphere of matter for (comoving) coordinate values $0 < r < R_1$ is locally exactly a FRW universe model which is singularity free (the matter comes from infinity, reaches a minimum volume and associated maximum density, and then expands again to infinity); while for all time, there is a singularity as $r \to R_2$ (or perhaps, at late or early times, for smaller values of r in the interval (R_1, R_2)). Thus the Universe is at all times surrounded by a matter singularity where the density diverges. However, the particles in the central sphere have no singular past or future history; and they live in a spatially homogeneous region of space–time where the redshifts are given the conventional 'Doppler' interpretation. At the time t the proper distance from the origin to the singularity is at most

$$\int_0^{R_2} X(r, t)\, dr = Y(R_2, t),$$

and so is finite as long as Y is finite. Equation (B3) will show that Y remains finite for all finite values of t. Accordingly, this space–time has the features mentioned at the end of Section 8.

B.3 STABILITY

The situation here is probably unstable; it is quite likely that if one were to introduce non-spherically symmetric perturbations, a singularity would occur on each world line in the past (or future); that is, one would have both the features of the FRW initial singularity in the past, and the SSS interactive 'surrounding' singularity.

Commun. math. Phys. 59, 273—283 (1978)

Communications in Mathematical Physics
© by Springer-Verlag 1978

Decomposable Differential Operators in a Cosmological Context

R. Maartens and S. D. Nel
Department of Applied Mathematics, University of Cape Town, Rondebosch, Cape Town, South Africa

Abstract. The integrability conditions for a certain second order ordinary differential equation in two variables are studied via the concept of decomposability of the associated differential operator. The results are applied to regain in a unified manner the known exact solutions for locally rotationally symmetric, spatially homogeneous cosmological models. In addition, new solutions are obtained.

1. Introduction

Einstein's equations

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = T_{ab} \tag{1.1}$$

for a spacetime metric tensor g_{ab}, where R_{ab} is the Ricci tensor, $R = R^a_a$, and Λ is the cosmological constant, can be solved exactly only in cases of rather high spacetime symmetry, and for relatively simple forms of the energy-momentum tensor T_{ab}.

In this paper, we consider exact solutions of (1.1) for spacetimes in which local coordinates $(x^a) = (t, x^\alpha)(a = 0, \ldots, 3; \alpha = 1, \ldots, 3)$ may be chosen so that one or more of the field equations, or combinations thereof, take the generic form

$$A_1 \frac{\ddot{X}}{X} + A_2 \frac{\ddot{Y}}{Y} + A_3 \frac{\dot{X}^2}{X^2} + A_4 \frac{\dot{Y}^2}{Y^2} + A_5 \frac{\dot{X}}{X} + A_6 \frac{\dot{Y}}{Y} + A_7 \frac{\dot{X}\dot{Y}}{XY} + H(X, Y, t) = 0 \tag{1.2}$$

where $A_i \in \mathbf{R}(i = 1, \ldots, 7)$, a dot denotes differentiation with respect to t, and $X(t)$, $Y(t)$ are metric component functions. This is the case, for example, when the spacetime is locally rotationally symmetric and admits a 1-parameter family of homogeneous hypersurfaces (see, e.g., Refs. [1–3]). If a first integral of (1.2) can be found, then together with the remaining field equations and the conservation equations $T^{ab}_{;b} = 0$, this often allows us to obtain a reduction of the system of equations to quadratures, or to a single ordinary differential equation in one variable, plus quadratures.

0010-3616/78/0059/0273/$02.20

To study the integrability conditions for (1.2), we define a differential operator on functions of differentiability class $C^k(k \geqq 2)$

$$L_{\mathbf{a}} : C^k(\mathbf{R}) \times C^k(\mathbf{R}) \to C^{k-2}(\mathbf{R})$$

by

$$L_{\mathbf{a}}[u, v](t) = a_1\ddot{u}(t) + a_2\ddot{v}(t) + a_3\dot{u}^2(t) + a_4\dot{v}^2(t) + a_5\dot{u}(t) + a_6\dot{v}(t) + a_7\dot{u}(t)\dot{v}(t) \tag{1.3}$$

where $a_1^2 + a_2^2 \neq 0$.

$L_{\mathbf{a}}$ will be said to be *decomposable* if there exist a function $F = F(u, v, t)$, of class C^k in its arguments, and a function $G = G(u, \dot{u}, v, \dot{v}, t)$, of class C^{k+1} in its arguments, such that

$$L_{\mathbf{a}}[u, v](t) = F(u, v, t)\dot{G}(u, \dot{u}, v, \dot{v}, t). \tag{1.4}$$

If $L_{\mathbf{a}}$ is decomposable, then the differential equation

$$L_{\mathbf{a}}[u, v](t) = 0 \tag{1.5}$$

is integrable, with integrating factor F^{-1}. In addition, (1.2), which is of form

$$L_{\mathbf{a}}[u, v](t) + H(u, v, t) = 0$$

where $u = \log|X|, v = \log|Y|$, and $\mathbf{a} = (A_1, A_2, A_3 + A_1, A_4 + A_2, A_5, A_6, A_7)$, then has a first integral

$$G\left(u, \frac{du}{dt'}, v, \frac{dv}{dt'}, t(t')\right) = t'_0 - t' \qquad (t'_0 \in \mathbf{R})$$

where

$$t'(t) = \int \frac{H(u(t), v(t), t)}{F(u(t), v(t), t)}\, dt.$$

In §2, the complete characterisation of decomposability of operators of form (1.3), is found. Although this does not lead to the full integrability conditions for (1.5) (since in general an integrating factor for (1.5) will depend also on $\dot{u}$ and $\dot{v}$), it is sufficient to lead to the regaining of all known exact solutions for perfect fluid spacetimes which are locally rotationally symmetric and spatially homogeneous (as listed by MacCallum [4]). Together with these known solutions, new exact solutions are given in §3. §4 contains some remarks on possible extensions of the method.

2. Integrability Conditions

Theorem 2.1. *$L_{\mathbf{a}}$ is decomposable iff*

$$a_1^2 a_4 + a_2^2 a_3 - a_1 a_2 a_7 = 0 \tag{D1}$$

$$a_5^2 a_4 + a_6^2 a_3 - a_5 a_6 a_7 = 0. \tag{D2}$$

Proof. Necessity. Suppose (1.4) holds. Then, by (1.3)

$$F\frac{\partial G}{\partial \dot{u}} = a_1 \tag{2.1a}$$

$$F\frac{\partial G}{\partial \dot{v}} = a_2 \tag{2.1b}$$

$$F\left[\frac{\partial G}{\partial u}\dot{u} + \frac{\partial G}{\partial v}\dot{v} + \frac{\partial G}{\partial t}\right] = a_3\dot{u}^2 + a_4\dot{v}^2 + a_5\dot{u} + a_6\dot{v} + a_7\dot{u}\dot{v}. \tag{2.1c}$$

By differentiating (2.1) with respect to $\dot{u}$ and $\dot{v}$ we obtain

$$F\frac{\partial^2 G}{\partial \dot{u}\partial u} = a_3 \tag{2.2a}$$

$$F\frac{\partial^2 G}{\partial \dot{v}\partial v} = a_4 \tag{2.2b}$$

$$F\left[\frac{\partial^2 G}{\partial \dot{u}\partial v} + \frac{\partial^2 G}{\partial \dot{v}\partial u}\right] = a_7 \tag{2.2c}$$

$$\frac{\partial}{\partial t}\left[G - \frac{\partial G}{\partial \dot{u}}\dot{u} - \frac{\partial G}{\partial \dot{v}}\dot{v}\right] = 0. \tag{2.2d}$$

Case 1: $a_1 a_2 \neq 0$
By differentiating (2.1a), (2.1b) with respect to u, v (respectively), and using (2.2a), (2.2b), we obtain expressions, which, with (2.2c), yield (D1).

Also, these expressions, together with (2.2) and (D1), give expressions for a_5, a_6, which, when differentiated with respect to v, u (respectively), and subtracted, yield

$$a_2 a_3 a_6 - a_1 a_4 a_5 = [a_1^2 a_4 - a_2^2 a_3]\frac{\partial}{\partial t}\log F$$

while differentiation with respect to t gives, using (2.1a), (2.1b) and (2.2d), and subtracting

$$(a_2 a_5 - a_1 a_6)\frac{\partial F}{\partial t} = 0.$$

Hence

$$(a_2 a_5 - a_1 a_6)(a_2 a_3 a_6 - a_1 a_4 a_5) = 0 \tag{2.3}$$

(D2) follows from (D1) and (2.3).

Case 2: $a_2 = 0$
By (2.1b) and (2.2b)

$$a_4 = 0$$

which is just (D1) for $a_2 = 0$.

(D2) is obtained by steps similar to those in Case 1.

Sufficiency

Case 1: $a_1 a_2 \neq 0$

Case 1(a): $a_3 = 0 = a_4$

(D1) implies $a_7 = 0$, and hence we can set

$$F = 1$$

$$G = a_1\dot{u} + a_2\dot{v} + a_5 u + a_6 v.$$

Case 1(b): $a_3^2 + a_4^2 \neq 0$

(D1), (D2) imply (2.3), and hence we can set

$$F = \exp\left[\alpha t - \frac{a_3}{a_1}u - \frac{a_4}{a_2}v\right]$$

$$G = F^{-1}(a_1\dot{u} + a_2\dot{v} + \beta)$$

where α and β are defined as follows:

if $a_2 a_5 - a_1 a_6 = 0$ then $\alpha = -\dfrac{a_5}{a_1} = -\dfrac{a_6}{a_2}$ and $\beta = 0$

if $a_2 a_5 - a_1 a_6 \neq 0$ then $\alpha = 0$ and $\beta = \begin{cases} \dfrac{a_1 a_5}{a_3} & \text{if } a_3 \neq 0 \\ \dfrac{a_2 a_6}{a_4} & \text{if } a_4 \neq 0. \end{cases}$

The remaining case (*Case 2:* $a_2 = 0$; *Case 2(a):* $a_3 = 0 = a_7$, *Case 2(b):* $a_3^2 + a_7^2 \neq 0$) is treated similarly.

In all cases we obtain $F(u, v, t)\dot{G}(u, \dot{u}, v, \dot{v}, t) = L_a[u, v](t)$. □

We now investigate the behaviour of the decomposability conditions (D1), (D2) under transformation of variable. Define differentiable functions

$$D_\sigma : \mathbf{R}^7 \to \mathbf{R} \qquad (\sigma = 1, 2)$$

by

$$D_1(x^i) = (x^1)^2 x^4 + (x^2)^2 x^3 - x^1 x^2 x^7$$

$$D_2(x^i) = (x^1)^2 x^4 + (x^6)^2 x^3 - x^5 x^6 x^7$$

where

$$(x^1)^2 + (x^2)^2 \neq 0.$$

Then

$$\mathscr{D} = D_1^{-1}(0) \cap D_2^{-1}(0)$$

is a 6-dimensional differentiable variety of $\mathbf{R}^7$, and $\{L_{\mathbf{a}} | \mathbf{a} \in \mathscr{D}\}$ is, by Theorem 2.1, the set of all decomposable operators $L_{\mathbf{a}}$.

A transformation of variable $(u, v, t) \mapsto (u', v', t')$, given by

$$\begin{aligned} u &= \phi^1(u', v', t') \\ v &= \phi^2(u', v', t') \\ t &= \phi^3(u', v', t') \end{aligned} \tag{2.4a}$$

results in a transformation

$$L_{\mathbf{a}}[u, v](t) \mapsto M\left(u', \frac{du'}{dt'}, \frac{d^2u'}{dt'^2}, v', \frac{dv'}{dt'}, \frac{d^2v'}{dt'^2}, t'\right).$$

If

$$M(u', \dots, t') = L_{\mathbf{a}'}[u', v'](t') \tag{2.4b}$$

where $\mathbf{a}' \in \mathbf{R}^7$, the transformation will be said to be *form-preserving*. Let $\mathscr{G}$ denote the group of form-preserving transformations, and denote by $\mathscr{G}^*$ the group of parameter transformations $\mathbf{a} \mapsto \mathbf{a}'$ of $\mathbf{R}^7$ induced by $\mathscr{G}$. (2.4b) implies that $\phi^3 = \phi^3(t')$ and hence it is shown readily that $\mathscr{G}$ preserves the property of decomposability. $\mathscr{D}$ is therefore an invariant variety of $\mathscr{G}^*$: if $L_{\mathbf{a}}$ is not decomposable, no form-preserving transformation (2.4a) of the variables can render it decomposable.

(Explicitly, $\mathscr{G}$ is given locally by the transformation equations

$$\begin{aligned} u &= g^1u' + g^2v' + g^6t' + g^8 \\ v &= g^3u' + g^4v' + g^7t' + g^9 \\ t &= g^5t' + g^{10} \end{aligned}$$

where $g^A \in \mathbf{R} (A = 1, \dots, 10)$, $(g^1g^4 - g^2g^3)g^5 \neq 0$, and

$$L_{\mathbf{a}}\left[\left(\frac{g^6}{g^5}\right)t, \left(\frac{g^7}{g^5}\right)t\right] = 0.$$

Then

$$D_1(a_i') = (g^1g^4 - g^2g^3)^2(g^5)^5 D_1(a_i)$$

$$D_2(a_i') = (g^1g^4 - g^2g^3)^2(g^5)^4 D_2(a_i) + (4a_3a_4 - a_7^2)L_{\mathbf{a}}\left[\left(\frac{g^6}{g^5}\right)t, \left(\frac{g^7}{g^5}\right)t\right]$$

verifying that $\mathscr{D}$ is an invariant variety of $\mathscr{G}^*$).

3. Examples

In this section we show, by means of illustrative examples, how in a cosmological context the foregoing results sometimes allow a reduction of Einstein's field equations to one or more ordinary differential equations involving only one dependent variable. We shall assume throughout that the large scale matter and radiation distribution in the universe is represented by a perfect fluid obeying the equation of state $p = (\gamma - 1)\mu$, where $1 \leqq \gamma \leqq 2$.

Example 1: Orthogonal, Locally Rotationally Symmetric, Spatially Homogeneous Models [1]

For these models, there exist co-moving co-ordinates (t, x, y, z) such that

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)[dy^2 + f^2(y)dz^2] - X^2(t)h(y)[2dx - h(y)dz]dz$$

where

$$f(y) = \begin{bmatrix} \sin y \\ y \\ \sinh y \end{bmatrix}, \ h(y) = \begin{bmatrix} 2c\cos y \\ -c^2y^2 \\ -2c\cosh y \end{bmatrix} \text{ for } k = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

and $c, k \in \mathbf{R}$ are parameters related to the symmetry group of the space-time.

The fluid 4-velocity is $\mathbf{u} = \frac{\partial}{\partial t}$, and the field equations are

$$\frac{\ddot{X}}{X} + \frac{\dot{X}\dot{Y}}{XY} + \frac{\ddot{Y}}{Y} + c^2\frac{X^2}{Y^4} = (1-\gamma)\mu + \Lambda \tag{3.1a}$$

$$2\frac{\ddot{Y}}{Y} + \frac{\dot{Y}^2 + k}{Y^2} - 3c^2\frac{X^2}{Y^4} = (1-\gamma)\mu + \Lambda \tag{3.1b}$$

$$2\frac{\dot{X}\dot{Y}}{XY} + \frac{\dot{Y}^2 + k}{Y^2} - c^2\frac{X^2}{Y^4} = \mu + \Lambda \tag{3.1c}$$

while the conservation equations reduce to

$$\frac{1}{\gamma}\frac{\dot{\mu}}{\mu} + \frac{\dot{X}}{X} + 2\frac{\dot{Y}}{Y} = 0. \tag{3.1d}$$

Using the reduction technique described in this paper, we have been able to regain all the hitherto known exact solutions of (3.1) (of which we are aware [4]). These are:

$c = 0$: *Bianchi* I($k = 0$): General solutions for $1 \leqq \gamma \leqq 2$; Vacuum.
Kantowski-Sachs I($k = +1$), II($k = -1$): General solutions for $\gamma = 1, \frac{4}{3}(\Lambda = 0), 2(\Lambda = 0)$; Vacuum.

$c \neq 0$: *Bianchi* II($k = 0$): Special solutions for $1 \leqq \gamma \leqq 2$; General solution for $\gamma = 2(\Lambda = 0)$; Vacuum.
Bianchi VIII($k = -1$), IX($k = +1$): Vacuum.

Detailed references may be found in [4].

In addition, we have obtained the (apparently new) general solution for the case $ck \neq 0, \gamma = 2$ (Bianchi types VIII and IX) and a general reduction to one second order o.d.e. and two quadratures for the case $c \neq 0, k = 0$ (Bianchi II) (in both cases with $\Lambda = 0$). These are as follows.

Case (*i*): $ck \neq 0, \gamma = 2$. (*Bianchi types* VIII *and* IX)

(3.1a) and (3.1c) decompose to

$$[Y(XY)^{\cdot}]^{\cdot} = -kX. \tag{3.2}$$

If we define a new time parameter τ by

$$\tau(t) = \int X(t)dt \tag{3.3}$$

(3.2) may be integrated twice to give

$$(XY)^2 = -k\tau^2 + a\tau + b. \tag{3.4}$$

Substituting for $\mu = d^2X^{-2}Y^{-4}(d \in \mathbf{R})$ from (3.1d) and for Y from (3.4) into (3.1c), yields an equation for X:

$$[(X^{-2})']^2 = [(a^2 + 4bk - d^2)X^{-2} - 4c^2](-k\tau^2 + a\tau + b)^{-2} \tag{3.5}$$

where $' \equiv \dfrac{d}{d\tau}$.

Since (3.5) is separable it integrates immediately to give $X = X(\tau)$, and then $Y = Y(\tau)$ is found from (3.4).

Thus the full solution is obtained in parametric form, where the proper time t is related to the parameter τ by (3.3).

Note that the vacuum solution ($\mu = 0$) is given by setting $d = 0$ in (3.5).

Case (*ii*): $c \neq 0, k = 0$. (*Bianchi* II)

Using (3.1a) and (3.1c) as before, we find after substituting for μ from (3.1d) and decomposing

$$[Y(XY)^{\cdot}]^{\cdot} = (2-\gamma)d^2(XY^2)^{1-\gamma}. \tag{3.6}$$

If $\gamma = 2$, we proceed as in (i) and obtain (3.5) with $k = 0$.

If $1 \leqq \gamma < 2$, we define a new time parameter η by

$$\eta = [(2-\gamma)d^{2(3-\gamma)}]^{1/2}\int [X(t)Y^2(t)]^{1-\gamma}dt. \tag{3.7}$$

Then (3.6) may be integrated to give

$$\frac{X'}{X} + \frac{Y'}{Y} = \mu^{(2-\gamma)/\gamma}(\eta - \eta_0) \tag{3.8}$$

where $' \equiv \dfrac{d}{d\eta}$ and $\eta_0 \in \mathbf{R}$.

Now (3.8) and (3.1d) allow us to express $\dfrac{X'}{X}$ and $\dfrac{Y'}{Y}$ purely in terms of μ. Substituting this into the linear combination (3.1) [(b)–3(c)], we obtain a second order equation for μ in terms of η:

$$\mu\mu'' - (4/\gamma)\mu'^2 - 8(\eta - \eta_0)\mu^{2/\gamma}\mu' - 6\gamma(\eta - \eta_0)^2\mu^{4/\gamma} + [\gamma - \gamma(\gamma+2)/2(2-\gamma) \cdot d^{2(3-\gamma)}]\mu^{(2+\gamma)/\gamma} = 0. \tag{3.9}$$

After solving (3.9) to obtain $\mu = \mu(\eta)$, the solution is completed by the quadratures

$$X = a\mu^{1/\gamma} \exp\left[2\int \mu^{(2-\gamma)/\gamma}(\eta - \eta_0)d\eta\right]$$
$$Y = b\mu^{-1/\gamma} \exp\left[-\int \mu^{(2-\gamma)/\gamma}(\eta - \eta_0)d\eta\right]$$

where $a, b \in \mathbf{R}$.

Example 2: Tilted, Locally Rotationally Symmetric, Spatially Homogeneous Models [2]

For these models, (invariant under a group of Bianchi type V), there exist coordinates (t, x, y, z) such that

$$ds^2 = -dt^2 + X^2(t)dx^2 + Y^2(t)\exp(-2A_0 x)(dy^2 + dz^2)$$

where $A_0 \in \mathbf{R}$, $A_0 \neq 0$, and the fluid 4-velocity is

$$\mathbf{u} = (\cosh\psi)\frac{\partial}{\partial t} + (X^{-1}\sinh\psi)\frac{\partial}{\partial x},$$

where ψ is the hyperbolic tilt angle.

The field equations are (with $\Lambda = 0$)

$$\frac{\ddot{X}}{X} + 2\frac{\dot{X}\dot{Y}}{XY} - 2\frac{A_0^2}{X^2} = \tfrac{1}{2}(2-\gamma)\mu + \gamma\mu\sinh^2\psi \tag{3.10a}$$

$$\frac{\ddot{Y}}{Y} + \frac{\dot{Y}^2}{Y^2} + \frac{\dot{X}\dot{Y}}{XY} - 2\frac{A_0^2}{X^2} = \tfrac{1}{2}(2-\gamma)\mu \tag{3.10b}$$

$$\frac{2A_0}{X}\left(\frac{\dot{X}}{X} - \frac{\dot{Y}}{Y}\right) = \gamma\mu\sinh\psi\cosh\psi \tag{3.10c}$$

$$\frac{\dot{Y}^2}{Y^2} + 2\frac{\dot{X}\dot{Y}}{XY} - 3\frac{A_0^2}{X^2} = \mu\cosh^2\psi + (\gamma-1)\mu\sinh^2\psi \tag{3.10d}$$

while the conservation equations can be written as

$$[\log(\mu^\gamma X Y^2 \cosh\psi)]^{\cdot} = \frac{2A_0}{X}\tanh\psi \tag{3.10e}$$

$$[\log(\mu^{(\gamma-1)/\gamma} X \sinh\psi)]^{\cdot} = 0. \tag{3.10f}$$

We shall solve these equations for the case $\gamma = 2$. This appears to be the first exact solution for this class of model with non-vanishing pressure.

By (3.10e),

$$\mu = A_0^2(\alpha\sinh\psi X)^{-2} \tag{3.11}$$

where $\alpha \in \mathbf{R}$.

Substituting (3.11) into (3.10a) and (3.10b), we find

$$\alpha^2\frac{\ddot{X}}{X} - (\alpha^2+1)\frac{\ddot{Y}}{Y} - (\alpha^2+1)\frac{\dot{Y}^2}{Y^2} + (\alpha^2-1)\frac{\dot{X}\dot{Y}}{XY} = 0. \tag{3.12}$$

By rescaling the x-coordinate, we may set $\alpha^2 = 1$, in which case (3.12) satisfies the conditions of Theorem 2.1, and decomposes to

$$[X^2(X^{-1}Y^2)^{\cdot}]^{\cdot} = 0. \tag{3.13}$$

Furthermore, (3.10b) decomposes to

$$[X(Y^2)^{\cdot}]^{\cdot} = 2A_0^2 Y^2 X^{-1} \tag{3.14}$$

which, after the introduction of a new time parameter ξ, defined by

$$\xi(t) = \int X^{-1}(t)dt \tag{3.15}$$

becomes

$$(Y^2)'' - 2A_0^2 Y^2 = 0$$

where $' \equiv \dfrac{d}{d\xi}$.

Hence

$$Y^2 = a \sinh(2A_0\xi) + b\cosh(2A_0\xi)$$

$a, b \in \mathbf{R}$, and by (3.13)

$$X = cY^2 \exp(-d\int Y^{-2}(\xi)d\xi)$$

where $c, d \in \mathbf{R}$.

The hyperbolic tilt angle ψ can be found from (3.10c), (3.14) and (3.16) to be

$$\psi = \text{arc}\coth[Y^{-2}(YY' + d)]$$

and now μ is determined by (3.11).

As in Example 1, the full solution is thereby obtained in parametric form.

The Farnsworth dust solution ($\gamma = 1$) [5] can be regained by our method after changing to comoving coordinates (τ, x, y, z) in which

$$ds^2 = -d\tau^2 + 2F(\tau)d\tau dx + [W^2(\tau) - F^2(\tau)]dx^2 + Y^2(\tau)\exp(-2A_0x)[dy^2 + dz^2]$$

and $\mathbf{u} = \dfrac{\partial}{\partial\tau}$.

Example 3: Tilted, Non-Locally Rotationally Symmetric, Spatially Homogeneous Models [2]

The method we have developed can be extended easily to the case where the generic Equation (1.2) contains more than 2 dependent variables. Even without this extension, however, it is sometimes possible to obtain the necessary decomposition by defining new variables in such a way that the problem involves essentially only 2 dependent variables.

For example, a certain class (the case Bianchi II, $\Sigma_{23} = 0$, in the notation of [2]) of tilted, spatially homogeneous models may be given by

$$ds^2 = -dt^2 + X^2(t)[dx - f(t, z)dy]^2 + Y^2(t)dy^2 + Z^2(t)dz^2$$

where

$$f(t,z) = -\sqrt{2a}Z - \sqrt{2b}\int Y(t)[X^3(t)Z(t)]^{-1}dt$$

$a, b > 0$, and the fluid 4-velocity is

$$\mathbf{u} = \cosh\psi\frac{\partial}{\partial t} + Z^{-1}\sinh\psi\frac{\partial}{\partial z}.$$

The field equations are (with $\Lambda = 0$)

$$\frac{\ddot{X}}{X} + \frac{\dot{X}\dot{Y}}{XY} + \frac{\dot{X}\dot{Z}}{XZ} + a\left(\frac{X}{YZ}\right)^2 - \frac{b}{X^4Z^2} = \tfrac{1}{2}(2-\gamma)\mu \tag{3.16a}$$

$$\frac{\ddot{Y}}{Y} + \frac{\dot{Y}\dot{Z}}{YZ} + \frac{\dot{X}\dot{Y}}{XY} - a\left(\frac{X}{YZ}\right)^2 + \frac{b}{X^4Z^2} = \tfrac{1}{2}(2-\gamma)\mu \tag{3.16b}$$

$$\frac{\ddot{Z}}{Z} + \frac{\dot{X}\dot{Z}}{XZ} + \frac{\dot{Y}\dot{Z}}{YZ} - a\left(\frac{X}{YZ}\right)^2 = \tfrac{1}{2}(2-\gamma)\mu + \gamma\mu\sinh^2\psi \tag{3.16c}$$

$$(\gamma\mu XYZ^2\sinh\psi\cosh\psi)^{\cdot} = 0 \tag{3.16d}$$

$$\frac{\dot{X}\dot{Y}}{XY} + \frac{\dot{Y}\dot{Z}}{YZ} + \frac{\dot{X}\dot{Z}}{XZ} - \tfrac{1}{2}a\left(\frac{X}{YZ}\right)^2 - \tfrac{1}{2}\frac{b}{X^4Z^2} = \mu\cosh^2\psi + (\gamma-1)\mu\sinh^2\psi \tag{3.16e}$$

and the conservation equations both integrate to give

$$\mu^{\gamma}XYZ\cosh\psi = d \tag{3.16f}$$

$$\mu^{(\gamma-1)/\gamma}Z\sinh\psi = e \tag{3.16g}$$

$d, e \in \mathbf{R}$.

We obtain a reduction of these equations for the case $\gamma = 2$. By (3.16a) and (3.16b),

$$[Z(XY)^{\cdot}]^{\cdot} = 0 \tag{3.17}$$

which suggests introducing a new time parameter ξ

$$\xi(t) = \int Z^{-1}(t)dt \tag{3.18}$$

so that (3.17) may be integrated twice to give

$$XY = g(\xi - \xi_0) \tag{3.19}$$

$g, \xi_0 \in \mathbf{R}$.

Using (3.16a), (3.16b) and (3.19), we can now derive an equation for X, which may be written as

$$WW'' - W'^2 + (\xi - \xi_0)^{-1}WW' + g(\xi - \xi_0)^{-1}W(aW^3 - b) = 0 \tag{3.20}$$

where $' \equiv \dfrac{d}{d\xi}$ and $W = X^4[g(\xi - \xi_0)]^{-1}$.

Equation (3.20) is a special case of the defining equation

$$SS'' - S'^2 - \xi^{-1}SS' - \xi^{-1}(a_1S^3 + a_2S) + a_3S^4 + a_4 = 0$$

of the 3rd Painleve transcendent [6].

Having solved (3.20), we obtain Y from (3.19), and ψ from (3.16f) and (3.16g)

$$\psi = \text{arc}\tanh\left[(eg/2d)(\xi - \xi_0)\right].$$

Now (3.16e) constitutes a linear, first order equation for Z^2, which may readily be solved to complete the solution.

4. Extensions

The method developed in this paper has enabled us not only to obtain in a unified manner the known perfect fluid, Locally Rotationally Symmetric exact solutions, but also to obtain some new solutions. In addition it should be possible to regain the known exact solutions for those L.R.S. spatially homogeneous spacetimes which admit a non-vanishing magnetic field by use of this method (see [4] for references to these solutions).

A natural extension of the results of this paper would be to obtain the decomposability conditions for the case where there are n dependent variables. For $n = 3$, this could be applied to the non-L.R.S. spatially homogeneous spacetimes containing a perfect fluid or magnetic field. In this way, it should be possible to regain the known solutions [4], and perhaps to obtain new solutions (one new solution for $n = 3$ was given in §3, where a new choice of dependent variable reduced the problem to $n = 2$).

It may be that the differential Equation (1.2) occurs in contexts other than the cosmological. In general, the concept of decomposability (corresponding to a restriction on the functional dependence of the integrating factor) could be applied to other types of differential equation (corresponding to operators other than $L_\bullet$), to obtain useful integrability conditions.

Acknowledgements. We would like to thank Prof. G. F. R. Ellis for helpful discussions. S. D. N. acknowledges the support of an AECI Research Fellowship, R. M. of U. C. T. and C. S. I. R. scholarships.

References

1. Ellis, G. F. R., MacCallum, M. A. H. : A class of homogeneous cosmological models. Commun. math. Phys. **12**, 108–141 (1969)
2. King, A. R., Ellis, G. F. R. : Tilted homogeneous cosmological models. Commun. math. Phys. **31**, 209–242 (1973)
3. Synge, J. L. : Relativity : The general theory, chap. 7. Amsterdam: North-Holland 1960
4. MacCallum, M. A. H. : Cosmological models from a geometric point of view. In: Cargèse lectures in physics, Vol. 6 (ed. E. Schatzman). New York-London-Paris : Gordon and Breach 1973
5. Farnsworth, D. L. : Some new general relativistic dust metrics possessing isometries. J. Math. Phys. **8**, 2315–7 (1967)
6. Ince, E. L. : Ordinary differential equations. New York: Dover Publications 1956

Communicated by R. Geroch

Received October 13, 1977

REFERENCES

[1] Ellis, G.F.R. (1979): "Limits to verification in cosmology." To appear in Ann. New York Acad. Sci.

[2] Ellis, G.F.R., Maartens, R. & Nel, S.D.:"Observational cosmology I: Ideal cosmographic observations." In preparation.

[3] Nel, S.D., Ellis, G.F.R. & Maartens, R.: "Observational cosmology II: Ideal cosmological observations." In preparation.

[4] Maartens, R., Nel, S.D. & Ellis, G.F.R.: "Observational cosmology III: Realistic cosmological observations." In preparation.

[5] Maartens, R. (1980): "Idealised observations in relativistic cosmology." Ph.D thesis, University of Cape Town.

[6] Ellis, G.F.R. (1975): "Cosmology and verifiability." Q. Jl. R. astr. Soc., *16*, 245.

[7] Harwitt, M. (1975): "The number of class A phenomena characterising the universe." Q.Jl. R. astr. Soc., *16*, 378.

[8] Dicke, R.H., Peebles, P.J.E., Roll, P.G. & Wilkinson, D.T. (1965): "Cosmic-black-body radiation." Astrophys. J., *142*, 414.

[9] Rees, M.J. (1978): "Origin of the pregalactic microwave background radiation." Nature, *275*, 35.

[10] Burbidge, G. & Brecher, K. (1971): "Extragalactic cosmic rays - a reappraisal." Comments on Astrophys. and Space Sci., *3*, 141.

[11] Schramm, D.N. & Wagoner, R.V. (1977): "Element production in the early universe." Annu. Rev. Nucl. Sci., *27*.

[12] Barrow, J. (1977): " The synthesis of light elements in turbulent cosmologies." Mon. Not. R. astr. Soc., *178*, 625.

[13] Jones, B.J.T. (1976): "The origin of galaxies: a review of theoretical developments and their comparison with observation." Rev. Mod. Phys., *48*, 107.

[14] Ellis, G.F.R. & Sciama, D.W. (1972): "Global and non-global problems in cosmology", in *General Relativity*, ed. L. O'Raifearty. Clarendon Press, Oxford.

[15] Ellis, G.F.R. (1971): "Relativistic cosmology", in *General Relativity and Cosmology*, ed. R.K. Sachs. Academic press, London.

[16] Weinberg, S.W. (1972): *Gravitation and Cosmology*. Wiley & Sons, New York.

[17] Ellis, G.F.R. (1979): "The homogeneity of the universe." J. Gen. Rel. Grav., *11*, 281.

[18] Penrose, R. (1979): "Singularities and time-asymmetry", in *General Relativity*, ed. S.W. Hawking & W. Isreal. Cambridge University Press.

[19] Carter, B. (1974): Large number coincidences and the anthropic principle in cosmology", in *Confrontation of Cosmological Theories with Observational Data*, ed. M. Longair. A. Reidel, Holland.

[20] Misner, C.W. (1968): " The isotropy of the universe." Astrophys. J., *151*, 431.

[21] Ellis, G.F.R., Maartens, R. & Nel, S.D. (1978): "The expansion of the universe." Mon. Not. R. astr. Soc., *184*, 439.

[22] Dyson, F.J. (1972): "The fundamental constants and their time variation", in *Aspects of Quantum Theory*, ed. A. Salam & E.P. Wigner. Cambridge University Press.

[23] Barrow, J.D. (1978): "A cosmological limit on the possible variation of G." Mon. Not. R. astr. Soc., *184*, 677.

[24] Will, C.M. (1979): "The confrontation between gravitational theory and experiment", in *General Relativity*, ed. S.W. Hawking & W. Isreal. Cambridge University Press.

[25] Hawking, S.W. & Ellis, G.F.R. (1973): *The large scale structure of space-time*. Cambridge University Press.

[26] Hughston, L.P. & Ward, R.S. (Editors) (1979): *Advances in Twistor Theory*. Pitman, London.

[27] Collins, C.B. & Ellis, G.F.R. (1979): "Singularities in Bianchi cosmologies." Phys. Rep., *56*, 2.

[28] de Vaucouleurs, G. (1976): "The case for a hierarchical cosmology." Science, *167*, 1203.

[29] de Vaucouleurs, G. (1976): "Supergalactic studies V." Astrophys. J., *205*, 13.

[30] Stewart, J.M. (1971): "Non-equilibrium Relativistic Kinetic Theory". Lecture notes in physics, *10*, Springer-Verlag, Berlin.

[31] Ellis, G.F.R. (1973): "Relativistic cosmology", in *Cargese Lectures, 6*, ed. E. Schatzman. Gordon & Breach, New York.

[32] Stewart, J.M. (1968): "Neutrino viscosity in cosmological models." Astrophys. Lett., *2*, 133.

[33] Misner, C.W. (1967): "Neutrino viscosity and the isotropy of primordial blackbody radiation." Phys. Rev. Lett., *19*,533.

[34] MacCallum, M.A.H. (1979): "Anisotropic and inhomogeneous relativistic cosmologies", in *General Relativity*, ed. S.W. Hawking & W. Isreal. Cambridge University Press.

[35] Smoot, G.F., Gorenstein, M.V. & Muller, R.A. (1977): "Detection of anisotropy in the cosmic blackbody radiation." Phys. Rev. Lett., *39*, 898.

[36] Rowan-Robinson, M. (1977): "Aether drift detected at last." Nature, *270*, 9.

[37] Rubin, V.C., Thonnard, N., Ford, W.K. & Roberts, M.S. (1976): "Motion of the galaxy and the local supercluster determined from the velocity anisotropy of distant Cs I galaxies. II." Astronom. J., *81*, 719.

[38] Ellis, G.F.R. & Perry, J. (1979): "Towards a correctionless cosmology." Mon. Not. R. astr. Soc., *187*, 357.

[39] Perry, J. & Ellis, G.F.R.: "Cosmological observations." In preparation.

[40] McCrea, W.H. (1960): "The interpretation of cosmology." Nature, *186*, 1035.

[41] Ellis, G.F.R. (1975): "Singularities and cosmology." Ann. New York Acad. Sci., *262*, 231.

[42] Collins, C.B. (1974): "Tilting at cosmological singularities." Commun. math. Phys., *39*, 131.

[43] King, A.R. & Ellis, G.F.R. (1973): "Tilted homogeneous cosmological models." Commun. math. Phys., *31*, 209.

[44] Ellis, G.F.R. (1971): "Topology and cosmology." J. Gen. Rel. Grav., *2*, 7.

[45] Barrow, J. (1976): "Light element abundances and the isotropy of the universe." Mon. Not. R. astr. Soc., *175*, 359.

[46] Barrow, J.D. & Matzner, R.A. (1977): "The homogeneity and isotropy of the universe." Mon. Not. R. astr. Soc., *181*, 719.

[47] Wasserman, I. & Brecher, K. (1978): "Time variation of Planck's constant." Phys. Rev. Lett., *41*, 920.

[48] Tayler, R.J. (1978): "Neutrino stability and cosmological helium production." Nature, *274*, 232.

[49] Longair, M.S. (1974): "The counts of radio sources", in *Confrontation of Cosmological Theories with Observatiol Data*, ed. M.S. Longair. A. Reidel, Holland.

[50] Kantowski, R. & Sachs, R.K. (1966): "Some spatially homogeneous anisotropic relativistic cosmological models." J. Math. Phys., *7*, 443.

[51] Tomita, K. (1968): "Theoretical relations among observable quantities in an anisotropic and homogeneous universe." Prog. Theor. Phys., *40*, 264.

[52] Collins, C.B. & Hawking, S.W. (1973): "The rotation and distortion of the universe." Mon. Not. R. astr. Soc., *162*, 307.

[53] MacCallum, M.A.H. & Ellis, G.F.R. (1970): "A class of homogeneous cosmological models II. Observations." Commun. math. Phys., *19*, 31.

[54] Kristian, J. & Sachs, R.K. (1966): "Observations in cosmology." Astrophys. J., *143*, 379.

[55] Newman, E.T. & Penrose, R. (1962): "An approach to gravitational radiation by a method of spin coefficients." J. Math. Phys., *3*, 566; *4*, 998.

[56] Penrose, R. (1963): "Null hypersurface initial data for classical fields of arbitrary spin and for general relativity", in *ARL Tech. Documentary Report 63-56*, P.G. Bergmann.

[57] Müller zum Hagen, H. & Seifert, H.J. (1976): "On characteristic initial-value and mixed problems." J. Gen. Rel. Grav., *8*, 259.

[58] Budic R.A. & Sachs, R.K. (1976): "Deterministic space-times." J. Gen. Rel. Grav., *7*, 21.

[59] Bondi, H., van der Burg, M.G.J. & Metzner, A.W.K. (1962): "Gravitational waves in general relativity VII. Waves from axi-symmetric isolated systems." Proc. R. Soc. Lond., A*269*, 21.

[60] Sachs, R.K. (1962): "Gravitational waves in general relativity VIII. Waves in asymptotically flat space-time." Proc. R. Soc. Lond., A*270*, 103.

[61] Newman, E.T. & Unti, T.W.J. (1962): "Behaviour of asymptotically flat empty spaces." J. Math. Phys., *3*, 891.

[62] Tamburino, L.A. & Winicour, J.H. (1960): "Gravitational fields in finite and conformal Bondi frames." Phys. Rev., *150*, 1039.

[63] Rorty, R. (1967): *The Linguistic Turn*. University of Chicago Press.

[64] Ellis, G.F.R. (1967): "Dynamics of pressure-free matter in general relativity." J. Math. Phys., *7*, 1171.

[65] Wainwright, J., Ince, W.C.W. & Marshman, B.J. (1979): "Spatially homogeneous and inhomogeneous cosmologies with equation of state $p = \mu$." Commun. math. Phys., *10*, 259.

[66] Maartens, R. & Nel, S.D. (1978): "Decomposable differential operators in a cosmological context." Commun. math. Phys., *59*, 273.

[67] Posadas, R. (1970): "The motion and structure of singularities in general relativity." Ph.D thesis, University of Pittsburgh.

[68] Newman, E.T. & Posadas, R. (1971): "Motion and structure of singularities in general relativity." J. Math. Phys., *12*, 2319.

[69] Pirani, F.A.E. (1964): "Introduction to gravitational radiation theory", in *Lectures in General Relativity*, ed. S. Deser & K.W. Ford. Prentice-Hall, New Jersey.

[70] Press, W.H. & Gunn, J.E. (1973): "Method for detecting a cosmological density for condensed objects." Astrophys. J., *185*, 397.

[71] Penrose, R. Private communication.

[72] Dixon, W.G. (1970): "Analysis of the Newman-Unti integration procedure for asymptotically flat space-times." J. Math. Phys., *11*, 1238.

[73] Penrose, R. (1968): "Structure of space-time", in *Battelle Rencontres*, ed. C.M. DeWitt & J.A. Wheeler. W.A. Bejamin, New York.

[74] Goldberg, J.N., MacFarlane, A.J., Newman, E.T., Rohrlich, F. & Sudarshan, E.A. (1967): " Spin-s spherical harmonics and ð." J. Math. Phys., *8*, 2155.

[75] Newman, E.T. & Penrose, R. (1966): "Note on the Bondi-Metzner-Sachs group." J. Math. Phys., *7*, 863.

[76] Geroch, R., Held, A. & Penrose, R. (1973): "A space-time calculus based on pairs of null directions." J. Math. Phys., *14*, 874.

[77] Harrison, E.R. (1974): "Cosmological principles. II: Physical principles." Comments on Astrophys. and Space Sci., *6*, 29.

[78] Penrose, R. (1966): "General-relativistic energy flux and elementary optics", in *Perspectives in Geometry and Relativity*, ed. B. Hoffman. Indiana University Press.

[79] Liebert, J., Angel, J.R.P., Hege, E.K., Martin, P.C. & Blair, W.P. (1979): "The moving emission features in SS433 require dynamical interpretation." Nature, *279*, 384.

[80] de Vaucouleurs, G. & Bollinger, J. (1979): "The extragalactic distance scale. VII. The velocity-distance relations in different directions and the Hubble ratio within and without the local super-cluster." Astrophys. J., *233*, 433.

[81] Bondi, H. (1947): "Spherically symmetrical models in general relativity." Mon. Not. R. astr. Soc., *107*, 410.

[82] Friedrich, H. (1979): "Eine untersuchung der Einsteinschen vakuumfeldgleichungen in der umgebung regulärer und singulärer nullhyperflächen." Ph.D thesis, University of Hamburg.

[83] Chellone, D.S. & Williams, D.P. (1972): " A characteristic initial value problem in general relativity in the case of a perfect fluid with axial symmetry." Proc. R. Soc. Lond., *A332*, 549.

[84] Bachman, G. & Narici, L. (1966): *Functional Analysis*. Academic Press, New York.

[85] Wainwright, J. (1971): "Characterisation of Szekeres inhomogeneous cosmologies as algebraically special solutions." J. Math. Phys., *18*, 672.

[86] Carmeli, M. (1977): *Group Theory and General Relativity*. McGraw-Hill, New York.

2 1 JAN 1981