The space of cosmological space-times

George F R Ellis

January 1, 2004

Abstract

I first focus on how to best describe the space of cosmological space-times and what its essential properties are and some comments on the dynamical behaviour revealed by studies that will be described in more detail by John Wainwright. I will then relate this both to observations and to anthropic issues[i.e. the possible existence of observers]. This space includes some viable singularity free solutions which will be briefly described, thus posing the issue of the tension between very special initial conditions and the existence of initial singularities. I will conclude with remarks on the issue of realised infinities in this context and the concept of multiverses

1 Miami talks: Introduction

The overall attempt in these talks will be to consider the description and uses of universe models and of the space of universe models in relation to (a) physical understanding of dynamics, (b) astronomical observations, and (c) irreversible local processes and the arrow of time. We are interested in how special our universe is - how 'fine tuned' initial conditions have to be to lead to this particular universe in which we live, so our framework should relate to that issue also.

To describe universes, we must characterise the space of possibilities for universes we choose to consider, and its controlling parameters. Then the aim is to determine the evolution of models in this space, in particular determining what is special and what is generic behaviour. Thus we may usefully consider ensembles of possible universes. There are some claims that such ensembles are realised, indeed that "all that can happen, happens" in a really existing multiverse; however this intriguing proposal is hard to relate to observational or other empirical tests. This issue is discussed by Ellis Kirchner and Stoeger (2003).

In order that cosmology be a proper science, the observational relations implied by cosmological models must be compared with astronomical observations. This determines those solutions that can usefully be considered as viable cosmological models of the real universe. A major aim of the present lectures is to point out that this class is wider than just the standard Friedmann–Lemaître/ Robertson–Walker ('FLRW') cosmologies, even though those models are very successful as models fo the universe; indeed those models cannot be realistic on all scales of description, but they may also be inaccurate on large scales, or at very early and very late times. To examine this, we need to consider the subspace of the space of all cosmological solutions that contains models with observational properties like those of the real universe at some stage of their histories. Thus we are interested in the *full* state space of solutions, allowing us to see how realistic models are related to each other and to higher symmetry models, including particularly the FLRW models.

Finally if we are interested in observational relations, as required to test if a specific model is a good model of the real universe, then it is taken for granted that observers can exist in that universe. But this is a highly non-trivial requirement. Hence a key question is which of the universes contemplated do indeed allow complex structures to exist and life to evolve.

2 The Set of Possible Universes

We first consider the overall project of describing the space of space-times, i.e. the set of possible cosmologies, relating this both to observations and to anthropic issues (i.e. the possible existence of observers).

The basis for describing universes is contained in the structure and the dynamics of a space $\mathcal{M}$ of all possible universes m, each of which can be described in terms of a set of states s in a state space $\mathcal{S}$. Each universe in $\mathcal{M}$ will be characterised by a set $\mathcal{P}$ of distinguishing parameters p, which are coordinates on $\mathcal{S}$. Some will be logical parameters, some will be numerical constants, and some will be functions or tensor fields defined in local coordinate neighbourhoods for s. Each universe m will evolve from its initial state to some final state according to the dynamics operative, with some or all of its parameters varying as it does so. The course of this evolution of states will be represented by a path in the state space $\mathcal{S}$, depending on the parametrisation of $\mathcal{S}$. Thus, each such path (in degenerate cases a point) is a representation of one of the universes m in $\mathcal{M}$. The coordinates in $\mathcal{S}$ will be directly related to the parameters specifying members of $\mathcal{M}$. The parameter space $\mathcal{P}$ has dimension N which is the dimension of the space of models $\mathcal{M}$; the space of states $\mathcal{S}$ has $N+1$ dimensions, the extra dimension indicating the change of each model's states with time, characterised by an extra parameter, e.g., the Hubble parameter H which does not distinguish between models but rather determines what is the state of dynamical evolution of each model. Note that N may be infinite, and indeed will be so unless we consider only geometrically highly restricted sets of universes.

It is possible that with some parameter choices the same physical universe m will be multiply represented by this description; thus a significant issue is the *equivalence problem* – identifying which different representations might in fact represent the same universe model. In *self-similar cases* we get a single point in $\mathcal{S}$ described in terms of the chosen parameters $\mathcal{P}$: the state remains unchanged in terms of the chosen variables. But we can always get such variables for any evolution, as they are just comoving variables, not necessarily indicating anything interesting is happening dynamically. The interesting issue is if this invariance is true in physically defined variables, e.g., expansion normalised variables; then physical self-similarity is occurring.

The very description of this space $\mathcal{M}$ of possibilities is based on an assumed set of laws of behaviour, either laws of physics or meta-laws that determine the laws of physics, which all universes m have in common; without this, we have no basis for setting up its description. The detailed characterisation of this space, and its relationship to $\mathcal{S}$, will depend on the matter description used and its behaviour. The overall characterisation of $\mathcal{M}$ therefore must incorporate a description both of the geometry of the allowed universes and of the physics of matter. Thus the set of parameters $\mathcal{P}$ will include both geometric and physical parameters.

The space $\mathcal{M}$ has a number of important subsets, for example:

1. $\mathcal{M}_{\mathrm{FLRW}}$ – the subset of all possible exactly Friedmann-Lemaître-Robertson-Walker (FLRW) universes, described by the state space $\mathcal{S}_{\mathrm{FLRW}}$ (in the case of dust plus non-interacting radiation a careful description of this phase space has been given by Ehlers and Rindler 1989).

2. $\mathcal{M}_{\mathrm{almost-FLRW}}$ – the subset of all perturbed FLRW model universes. These need to be characterised in a gauge-invariant way (see e.g. Ellis and Bruni 1989) so that we can clearly identify those universes that are almost-FLRW and those that are not.

3. $\mathcal{M}_{\mathrm{anthropic}}$ – the subset of all possible universes in which life emerges at some stage in their evolution. This subset intersects $\mathcal{M}_{\mathrm{almost-FLRW}}$, and may even be a subset of $\mathcal{M}_{\mathrm{almost-FLRW}}$, but does not intersect $\mathcal{M}_{\mathrm{FLRW}}$ (realistic models of a life-bearing universe like ours cannot be exactly FLRW, for then there is no structure).

4. $\mathcal{M}_{\mathrm{Observational}}$ – the subset of models compatible with current astronomical observations. Precisely because we need observers to make observations, this is a subset of $\mathcal{M}_{\mathrm{anthropic}}$.

If $\mathcal{M}$ truly represents all possibilities for universes, as is required in order to see how special or general a universe model is, one must have a description that is wide enough to encompass *all* possibilities. It is here that major issues arise: how do we decide what all the possibilities are? What are the limits of possibility? What classifications of possibility are to be included? Proponents of the multiverse idea suggest "All that can happen happens" - all possibilities, as characterised by our description in terms of families of parameters must occur, and they must occur in all possible combinations. The full space $\mathcal{M}$ must be large enough to represent all of these possibilities. The larger the possibility space considered, the more fine-tuned the actual universe appears to be - for with each extra possibility that is included in the possibility space, unless it can be shown to relate to already existing parameters, the actual universe and its close neighbours will live in a smaller fraction of the possibility space. For example if we assume General Relativity then there is only the parameter G to measure; but if we consider scalar-tensor theories, then we have to explain why we are so close to General Relativity now. Hence there is a tension between including all possibilities in what we consider, and giving an explanation for fine tuning.

2.1 Adequately Specifying Possible Anthropic Universes

When defining any ensemble of universes, possible or realised, we must specify all the parameters which differentiate members of the ensemble from one another at any time in their evolution. The values of these parameters may not be known or determinable initially in many cases – some of them may only be set by transitions that occur via processes like symmetry breaking within given members of the ensemble. In particular, some of the parameters whose values are important for the origination and support of life may only be fixed later in the evolution of universes in the multiverse.

We can separate our set of parameters $\mathcal{P}$ for the space of all possible universes $\mathcal{M}$ into different categories, beginning with the most basic or fundamental, and progressing to more contingent and more complex categories. Ideally they should all be independent of one another, but we will not be able to establish that independence for each parameter, except for the most fundamental cosmological ones. In order to categorise our parameters, we can doubly index each parameter p in $\mathcal{P}$ as $p_j(i)$ such that those for $j = 1 - 2$ describe basic physics, for $j = 3 - 5$ describe the cosmology (given that basic physics), and $j = 6 - 7$ pertain specifically to emergence and life (we must include the latter if we seriously intend to address anthropic issues). Our characterisation is as follows:

1. $p_1(i)$ are the basic physics parameters within each universe, excluding gravity - parameters characterising the basic non-gravitational laws of physics in action, related constants such as the fine-structure constant α, and including parameters describing basic particle properties (masses, charges, spins, etc.) These should be logical parameters or dimensionless parameters, otherwise one may be describing the same physics in other units.

2. $p_2(i)$ are basic parameters describing the nature of the cosmological dynamics, e.g., $p_2(1) = 1$ indicates Einstein gravity dominates, $p_2(1) = 2$ indicates Brans-Dicke theory dominates, $p_2(1) = 3$ indicates electromagnetism dominates, etc. Associated with each choice are the relevant parameter values, e.g., $p_2(2) = G$, $p_2(3) = \Lambda$, and in the Brans-Dicke case $p_2(4) = \omega$. If gravity can be derived from more fundamental physics in some unified fundamental theory, these will be related to the parameters $p_1(i)$; for example the cosmological constant may be determined from quantum field theory and basic matter parameters.

3. $p_3(i)$ are cosmological parameters characterising the nature of the matter content of a universe. These parameters encode whether radiation, bary-

onic matter, dark matter, neutrinos, scalar fields, etc. occur, in each case specifying the relevant equations of state and auxiliary functions needed to determine the physical behaviour of matter (e.g. a barotropic equation of state for a fluid and the potential function for a scalar field). These are characterisations of physical possibilities for the macro-states of matter arising out of fundamental physics, so the possibilities here will be related to the parameters in $p_1(i)$. Realistic representations of the Universe will include all the above, but simplified ensembles considered for exploratory purposes may exclude some or many of them.

4. $p_4(i)$ are physical parameters determining the relative amounts of each kind of matter present in the specific cosmological solutions envisaged, for example the density parameters Ω_i of various components at some specific stage of its evolution (which then for example determine the matter to anti-matter ratio and the entropy to baryon ratio). The matter components present will be those characterised by $p_3(i)$.

5. $p_5(i)$ are geometrical parameters characterising the spacetime geometry of the cosmological solutions envisaged- for example the scale factor $a(t)$, Hubble parameter $H(t)$, and spatial curvature parameter k in FLRW models. These will be related to $p_4(i)$ by the gravitational equations set in $p_2(i)$, for example the Einstein Field Equations.

6. $p_6(i)$ are parameters related to the functional emergence of complexity in the hierarchy of structure, for example allowing the existence of chemically complex molecules. Thus $p_6(1)$ might be the number of different types of atoms allowed (as characterised in the periodic table), $p_6(2)$ the number of different states of matter possible (crystal, glass, liquid, gas, plasma for example), and $p_6(3)$ the number of different types of molecules characterised in a suitable way. These are emergent properties arising out of the fundamental physics in operation, and so are related to the parameters set in $p_1(i)$.

7. $p_7(i)$ are biologically relevant parameters related specifically to the functional emergence of life and of self-consciousness, for example $p_7(1)$ might characterise the possibility of supra-molecular chemistry and $p_7(2)$ that of living cells. This builds on the complexity allowed by $p_6(i)$ and relates again to the parameter set $p_1(i)$.

It is important to note that these parameters will describe the set of possibilities we are able to characterise on the basis of our accumulated scientific experience. The limits of our understanding are relevant here, in the relation between what we conceive of as this space of possibilities, and what it really is. There may be universes which we believe are possible on the basis of what we know of physics, that may in fact not be possible. There may also be universes which we conceive of as being impossible for one reason or another, that turn out to be possible. And it is very likely that we simply may not be able to imagine or envisage all the possibilities. However this is by no means a statement that "all that can occur" is arbitrary. On the contrary, specifying the set of possible parameters determines a uniform high-level structure that is obeyed by all universes in $\mathcal{M}$.

We see, then, that a possibility space $\mathcal{M}$ is the set of universes (one-parameter sets of states $\mathcal{S}$) obeying the dynamics characterised by a parameter space $\mathcal{P}$, which may be considered to be the union of all allowed parameters $p_j(i)$ for all i, j as briefly discussed above:

$$\mathcal{M} = \{\mathcal{S}, \mathcal{P}\}, \quad \mathcal{P} = \cup_{i,j}\, p_j(i).$$

Because the parameters $\mathcal{P}$ determine the dynamics, the set of paths in $\mathcal{S}$ characterising individual universes m are determined by this prescription. In some particular envisaged ensemble, some of these parameters ('class parameters') may be fixed across the ensemble, thus defining a class of universes considered, while others ('member parameters') will vary across the ensemble, defining the individual members of that class. Thus

$$\mathcal{P} = \mathcal{P}_{\text{class}} \cup \mathcal{P}_{\text{member}}.$$

As we consider more generic ensembles, class parameters will be allowed to vary and so will become member parameters. In an ensemble in which all that is possible happens, all parameters will be member parameters; however that is so hard to handle that we usually analyse sub-spaces characterised by particular class parameters.

2.2 Cosmological Models

Having made these remarks, we now proceed making standard assumptions about the physics, but remembering that there are wider possibilities as indicated in the present discussion.

A cosmological model represents the universe at a particular scale. We will usually assume that on large scales, space-time geometry is described by General Relativity (but note that there are other possibilities in this regard). Then a cosmological model is defined by specifying:

* the *space-time geometry* represented on some specific averaging scale and determined by the metric $g_{ij}(x^k)$, which — because of the requirement of compatibility with observations — must either have some expanding Robertson–Walker ('RW') geometries as a regular limit, or else be demonstrated to have observational properties compatible with the major features of current astronomical observations of the universe;

* the *matter present in the universe*, represented on the same averaging scale, and its *physical behaviour* (the energy-momentum tensor of each matter component, the equations governing the behaviour of each such component, and the interaction terms between them), which must represent physically plausible matter (ranging from early enough times to the present day, this will include most of the interactions described by present-day physics); and

* the *interaction of the geometry and matter* — how matter determines the geometry, which in turn determines the motion of the matter. We assume this is through the *Einstein gravitational field equations* ('EFE') given by[1]

$$G_{ab} \equiv R_{ab} - \frac{1}{2} R\, g_{ab} = T_{ab} - \Lambda\, g_{ab} \ , \tag{1}$$

which, because of the *twice-contracted Bianchi identities*, guarantee the conservation of total energy-momentum

$$\nabla_b G^{ab} = 0 \ \Rightarrow \ \nabla_b T^{ab} = 0 \ , \tag{2}$$

provided the *cosmological constant* Λ satisfies the relation $\nabla_a \Lambda = 0$, i.e., it is constant in time and space.

Together, these determine the combined dynamical evolution of the model and the matter in it. The description must be sufficiently complete to determine

* the *observational relations* predicted by the model for both discrete sources and background radiation, implying a well-developed theory of *structure growth* for very small and for very large physical scales (i.e. for light atomic nuclei and for galaxies and clusters of galaxies), and of *radiation absorbtion and emission.*

To be useful in an explanatory role, a cosmological model must be easy to describe — that means they have symmetries or special properties of some kind or other. The usual choices for the matter description will be some combination of a fluid with a physically well-motivated equation of state, for example a perfect fluid with specified equation of state (beware of imperfect fluids, unless they have well-defined and motivated physical properties); a mixture of fluids, usually with different 4-velocities; a set of particles represented by a kinetic theory description; a scalar field ϕ, with a given potential $V(\phi)$ (at early times); an electromagnetic field described by the Maxwell field equations.

2.3 Describing the Geometry of Possible Universes

2.3.1 The preferred vector field

Cosmological models are characterised by a preferred timelike vector field $u : \ u^a u_a = -1$, usually the fluid flow vector (Ellis 1971), but sometimes chosen for other reasons, e.g., to fit local symmetries.n a cosmological space-time $(\mathcal{M}, \mathbf{g})$, at late times there will be a family of preferred worldlines representing the average motion of matter at each point[2] (notionally, these represent the histories of clusters of galaxies, with associated 'fundamental observers'); at early times there will

[1]Throughout this review we employ geometrised units characterised by $c = 1 = 8\pi G/c^2$. Consequently, all geometrical variables occurring have physical dimensions that are integer powers of the dimension [length]. The index convention is such that space-time and spatial indices with respect to a general basis are denoted by $a, b, \ldots = 0, 1, 2, 3$ and $\alpha, \beta, \ldots = 1, 2, 3$, respectively, while space-time indices in a coordinate basis are $i, j, \ldots = 0, 1, 2, 3$.

[2]We are here assuming a fluid description can be used on a large enough scale [3, 6]. The alternative is that the matter distribution is hierarchically structured at all levels or fractal, so that a fluid description does not apply. The success of the FLRW models encourages us to use the approach taken here.

be uniquely defined notions of the average velocity of matter (at that time, interacting gas and radiation), and corresponding preferred worldlines. In each case their *4-velocity* is

$$u^a = \frac{dx^a}{d\tau} \ , \ \ u_a u^a = -\,1 \ , \tag{3}$$

where τ is proper time measured along the fundamental worldlines. We assume this 4-velocity is unique: that is, there is a well-defined preferred motion of matter at each space-time event. At recent times this is taken to be the 4-velocity defined by the vanishing of the dipole of the cosmic microwave background radiation ('CBR'): for there is precisely one 4-velocity which will set this dipole to zero. It is usually assumed that this is the same as the average 4-velocity of matter in a suitably sized volume; indeed this assumption is what underlies studies of large scale motions and the 'Great Attractor'.

Given u^a, there are defined unique *projection tensors*

$$U^a{}_b = -\,u^a\,u_b \quad \Rightarrow \quad U^a{}_c\,U^c{}_b = U^a{}_b \ , \ U^a{}_a = 1 \ , \ U_{ab}\,u^b = u_a \ , \tag{4}$$

$$h_{ab} = g_{ab} + u_a\,u_b \quad \Rightarrow \quad h^a{}_c\,h^c{}_b = h^a{}_b \ , \ h^a{}_a = 3 \ , \ h_{ab}\,u^b = 0 \ . \tag{5}$$

The first projects parallel to the 4-velocity vector u^a, and the second determines the (orthogonal) metric properties of the instantaneous rest-spaces of observers moving with 4-velocity u^a. There is also defined a *volume element* for the rest-spaces:

$$\eta_{abc} = u^d\,\eta_{dabc} \ \Rightarrow \ \eta_{abc} = \eta_{[abc]} \ , \ \eta_{abc}\,u^c = 0 \ , \tag{6}$$

where η_{abcd} is the 4-dimensional volume element ($\eta_{abcd} = \eta_{[abcd]}$, $\eta_{0123} = \sqrt{|\,g_{ab}\,|}$).

Moreover, *two derivatives* are defined: the covariant time derivative $\dot{}$ along the fundamental worldlines, where for any tensor $T^{ab}{}_{cd}$

$$\dot{T}^{ab}{}_{cd} = u^e T^{ab}{}_{cd;e} \ , \tag{7}$$

and the fully orthogonally projected covariant derivative ${}^3\tilde{\nabla}$, where for any tensor $T^{ab}{}_{cd}$

$${}^3\tilde{\nabla}_e T^{ab}{}_{cd} = h^a{}_f\,h^b{}_g\,h^p{}_c\,h^q{}_d \ \ h^r{}_e \nabla_r\,T^{fg}{}_{pq} \ , \tag{8}$$

with total projection on all free indices. The tilde serves as a reminder that if u^a has *non-zero* vorticity, ${}^3\tilde{\nabla}_e$ is *not* a proper 3-dimensional covariant derivative . Finally, we use angle brackets to denote orthogonal projections of vectors and the orthogonally projected symmetric trace-free part of tensors: $v^{<a>} = h^a{}_b\,v^b$, $T^{<ab>} = [\ h^{(a}{}_c\,h^{b)}{}_d - \frac{1}{3}\,h^{ab}\,h_{cd}\ ]\,T^{cd}$;for convenience the angle brackets are also used to denote othogonal projections of covariant time derivatives along u^a ('*Fermi derivatives*'): $\dot{v}^{<a>} = h^a{}_b\,\dot{v}^b$, $\dot{T}^{<ab>} = [\ h^{(a}{}_c\,h^{b)}{}_d - \frac{1}{3}\,h^{ab}\,h_{cd}\ ]\,\dot{T}^{cd}$. Note that the projected time and space derivatives of U_{ab}, h_{ab} and η_{abc} all vanish.

The existence of the preferred vector field leads to a 1+3 decomposition of all interesting physical and geometrical quantities into components relevant for fundnamental observers. For example the *Maxwell field strength tensor* F_{ab} of an electromagnetic field is split relative to u^a into *electric* and *magnetic field* parts by the relations

$$E_a = F_{ab}\, u^b \Rightarrow E_a u^a = 0 \;, H_a = \frac{1}{2}\, \eta_{abc}\, F^{bc} \Rightarrow H_a u^a = 0 \;.$$

In analogy to F_{ab}, the *Weyl conformal curvature tensor* C_{abcd} is split relative to u^a into *'electric'* and *'magnetic' Weyl curvature* parts according to

$$\begin{aligned} E_{ab} = C_{acbd}\, u^c\, u^d \quad &\Rightarrow \quad E^a{}_a = 0 \;,\; E_{ab} = E_{(ab)} \;,\; E_{ab}\, u^b = 0 \;, \\ H_{ab} = \frac{1}{2}\, \eta_{ade}\, C^{de}{}_{bc}\, u^c \quad &\Rightarrow \quad H^a{}_a = 0 \;,\; H_{ab} = H_{(ab)} \;,\; H_{ab}\, u^b = 0 \;. \end{aligned}$$

These represent the 'free gravitational field', enabling gravitational action at a distance (tidal forces, gravitational waves), and influence the motion of matter and radiation through the *geodesic deviation equation* for timelike and null vectors, respectively. Together with the Ricci tensor R_{ab} (determined locally at each point by the matter tensor through the EFE (1)), these quantities completely represent the space-time *Riemann curvature tensor* R_{abcd}, which in fully $1+3$-decomposed form becomes

$$\begin{aligned} R^{ab}{}_{cd} &= R_P^{ab}{}_{cd} + R_I^{ab}{}_{cd} + R_E^{ab}{}_{cd} + R_H^{ab}{}_{cd} \;, \\ R_P^{ab}{}_{cd} &= \frac{2}{3}\,(\mu + 3p - 2\Lambda)\, u^{[a}\, u_{[c}\, h^{b]}{}_{d]} + \frac{2}{3}\,(\mu + \Lambda)\, h^a{}_{[c}\, h^b{}_{d]} \;, \\ R_I^{ab}{}_{cd} &= -\,2\, u^{[a}\, h^{b]}{}_{[c}\, q_{d]} - 2\, u_{[c}\, h^{[a}{}_{d]}\, q^{b]} - 2\, u^{[a}\, u_{[c}\, \pi^{b]}{}_{d]} + 2\, h^{[a}{}_{[c}\, \pi^{b]}{}_{d]} \;, \\ R_E^{ab}{}_{cd} &= 4\, u^{[a}\, u_{[c}\, E^{b]}{}_{d]} + 4\, h^{[a}{}_{[c}\, E^{b]}{}_{d]} \;, \\ R_H^{ab}{}_{cd} &= 2\, \eta^{abe}\, u_{[c}\, H_{d]e} + 2\, \eta_{cde}\, u^{[a}\, H^{b]e} \;. \end{aligned}$$

2.3.2 The space-time geometry

To describe a cosmological spacetime locally we must give a description of its (generally inhomogeneous and anisotropic) geometry via suitable parameters $p_5(i)$. This description may be usefully given in terms of a tetrad basis as follows (see Ellis and van Elst 1999, Wainwright and Ellis 1996, Uggla, *et al.* 2003):

Feature 1: a set of local coordinates $\mathcal{X} = \{x^i\}$ must be chosen in each chart of a global atlas. This will in particular have a time coordinate t which will be used to characterise evolution of the universe; this should be chosen in as uniform as possible a way across all the universes considered, for example it may be based on surfaces of constant Hubble parameter H for the preferred vector field u.

Feature 2: in each chart, to determine the geometry we must be given the components $\mathcal{E} = [e^i{}_a(x^j)]$ of an orthonormal tetrad with the fluid flow vector chosen as the timelike tetrad vector ($a, b, c\ldots$ are tetrad indices; four of these components can be set to zero by suitable choice of coordinates). Together the coordinates and the tetrad form the reference frame

$$\mathcal{P}_{\rm frame} \equiv \{\mathcal{X}, \mathcal{E}\}. \tag{9}$$

The metric tensor is then

$$ds^2 = g_{ij}\,(x^k)dx^i dx^j = \eta_{ab}\, e^a{}_i(x^k)\, e^b{}_j(x^l)dx^i dx^j$$

where η_{ab} is the Minkowski metric:

$$\eta_{ab} = e_a.e_b = diag(-1,+1,+1,+1)$$

(because the tetrad is orthonormal) and $e^b{}_j(x^j)$ are the inverse of $e^i{}_a(x^j)$:

$$e^i{}_a(x^j)e_i{}^b(x^j) = \delta^b_a.$$

Thus the metric is given by

$$ds^2 = -\left(e^0{}_i dx^i\right)^2 + \left(e^1{}_i dx^i\right)^2 + \left(e^2{}_i dx^i\right)^2 + \left(e^3{}_i dx^i\right)^2. \tag{10}$$

The basic geometric quantities used to determine the spacetime geometry are the rotation coefficients $\Gamma^a{}_{bc}$ of this tetrad, defined by

$$\Gamma^a{}_{bc} = e_j{}^a e^j{}_{c;k} e^k{}_b.$$

They may conveniently be given in terms of geometric quantities

$$\mathcal{P}_{\rm geometry} \equiv \{\dot{u}_\alpha, \theta, \sigma_{\alpha\beta}, \omega_{\alpha\beta}, \Omega_\gamma, a^\alpha, n_{\alpha\beta}\}. \tag{11}$$

characterised as follows:

$$\begin{aligned}
\Gamma_{\alpha 00} &= \dot{u}_\alpha,\\
\Gamma_{\alpha 0\beta} &= \frac{1}{3}\theta + \sigma_{\alpha\beta} - \omega_{\alpha\beta},\\
\Gamma_{\alpha\beta 0} &= \epsilon_{\alpha\beta\gamma}\Omega^\gamma,\\
\Gamma_{\alpha\beta\gamma} &= a_{[\alpha}\delta_{\beta]\gamma} + \epsilon_{\gamma\delta[\alpha}n^\delta{}_{\beta]} + \frac{1}{2}\epsilon_{\alpha\beta\delta}n^\gamma{}_\delta,
\end{aligned}$$

where $\dot{u}_\alpha$ is the acceleration of the fluid flow congruence, θ is its expansion, $\sigma_{\alpha\beta} = \sigma_{(\alpha\beta)}$ is its shear ($\sigma^b{}_b = 0$), and $\omega_{\alpha\beta} = \omega_{[\alpha\beta]}$ its vorticity, while $n_{\alpha\beta} = n_{(\alpha\beta)}$ and a_α determine the spatial rotation coefficients (see Wainwright and Ellis 1996, Ellis and van Elst 1999). Greek indices (with range $1-3$) indicate that all these quantities are orthogonal to u^a. They are related by

$\nabla_a u_b = -\,\dot{u}_a u_b + {}^3\,\tilde{\nabla}_a u_b = -\,\dot{u}_a u_b + \,\frac{1}{3}\,\Theta\, h_{ab} + \sigma_{ab} + \omega_{ab}$.

The stated meaning for these quantities follows from the evolution equation for a relative position vector $\eta^a_\perp = h^a{}_b\eta^b$, where η^a is a deviation vector for the family of fundamental worldlines, i.e. $u^b\nabla_b\eta^a = \eta^b\nabla_b u^a$. Writing $\eta^a_\perp = \delta\ell\, e^a$, $e_a e^a = 1$, we find the relative distance $\delta\ell$ obeys the propagation equation

$$\frac{(\delta\ell)^{\cdot}}{\delta\ell} = \frac{1}{3}\,\Theta + (\sigma_{ab}e^a e^b) \ , \tag{12}$$

(the generalised Hubble law), and the relative direction vector e^a the propagation equation

$$\dot{e}^{<a>} = (\sigma^a{}_b - (\sigma_{cd}e^c e^d)\, h^a{}_b - \omega^a{}_b)\, e^b \ , \tag{13}$$

giving the observed rate of change of position in the sky of distant galaxies. The *average length scale* S determined by

$$\frac{\dot{S}}{S} = \frac{1}{3}\,\Theta \ , \tag{14}$$

so the volume of a fluid element varies as S^3.Finally $\dot{u}^a = u^b\nabla_b u^a$ is the *relativistic acceleration* vector, representing the degree to which the matter moves under forces other than gravity plus inertia (which cannot be covariantly separated from each other in General Relativity: they are different aspects of the same effect). The acceleration vanishes for matter in free fall (i.e. moving under gravity plus inertia alone).

The Jacobi identities, Bianchi identities, and Einstein field equations can all be written out in terms of these quantities, as can the components $E_{\alpha\beta}$, $H_{\alpha\beta}$ of the Weyl tensor (see Ellis and van Elst 1999 and the Appendix). Except in the special cases of isotropic spacetimes and locally rotationally symmetric spacetimes, the basis tetrad can be chosen in an invariant way so that three of these quantities vanish and all the rest are scalar invariants.

Thus the geometry is determined by the 36 spacetime functions in the combined set $(\mathcal{E}, \mathcal{P}_{\rm geometry})$ with some chosen specification of coordinates $\mathcal{X}$, with the metric then determined by (10). For detailed dynamical studies it is often useful to rescale the variables in terms of the expansion (see Wainwright and Ellis 1996, Uggla et al 2003 for details). Note that the same universe may occur several times over in this space; the *equivalence problem* is determining when such multiple representations occur. We do not recommend going to a quotient space where each universe occurs only once, as for example in the dynamical studies of Fischer and Marsden (1979), for the cost of doing so is to destroy the manifold structure of the space of spacetimes. It is far better to allow multiple representations of the same universe (for example several representations of the same Bianchi I universe occur in the Kasner ring in the space of Bianchi models, see Wainwright and Ellis 1996) both to keep the manifold structure intact and because then the dynamical structure becomes clearer.

Feature 3: To determine the global structure, we need a set of composition functions relating different charts in the atlas where they overlap, thus determining the global topology of the universe.

Together these are the parameters $p_5(i)$ needed to distinguish model states. A particular model will be represented as a path through those states. The nature of that evolution will be determined by the matter present.

2.4 Describing the Physics of Possible Universes

The matter *energy-momentum tensor* T_{ab} can be decomposed relative to u^a in the form

$$
\begin{aligned}
T_{ab} &= \mu\, u_a\, u_b + q_a\, u_b + u_a\, q_b + p\, h_{ab} + \pi_{ab} \ , \\
q_a\, u^a &= 0 \ , \ \pi^a{}_a = 0 \ , \ \pi_{ab} = \pi_{(ab)} \ , \ \pi_{ab}\, u^b = 0 \ ,
\end{aligned}
$$

where $\mu = (T_{ab} u^a u^b)$ is the *relativistic energy density* relative to u^a, $q^a = -\,T_{bc}\, u^b\, h^{ca}$ is the *relativistic momentum density*, which is also the energy flux relative to u^a, $p = \frac{1}{3}\,(T_{ab} h^{ab})$ is the *isotropic pressure* , and $\pi_{ab} = T_{cd}\, h^c{}_{<a}\, h^d{}_{b>}$ is the trace-free *anisotropic pressure* (stress).

Feature 4: To determine the matter stress-energy tensor we must specify the quantities

$$\mathcal{P}_{\rm matter} \equiv \{\mu, q_\alpha, p, \pi_{ab}, \Phi_A\} \tag{15}$$

for all matter components present, where μ is the energy density, q_α is the momentum flux density, p is the pressure, $\pi_{ab} = \pi_{(ab)}$ the anisotropic pressure ($\pi^b_{\ b} = 0$), and Φ_A ($A = 1..A_{\rm max}$) is some set of internal variables sufficient to make the matter dynamics deterministic when suitable equations of state are added (for example these might include the temperature, the entropy, the velocity v^i of matter relative to the reference frame, some scalar fields and their time derivatives, or a particle distribution function). These are parameters $p_4(i)$ for each kind of matter characterised by $p_3(i)$. Some of these dynamical quantities may vanish (for example, in the case of a 'perfect fluid', $q_\alpha = 0$, $\pi_{ab} = 0$) and some of those that do not vanish will be related to others by the equations of state (for example, in the case of a barotropic fluid, $p = p(\mu)$) and dynamic equations (for example the Klein Gordon equation for a scalar field). The *physics* of the situation is in the *equations of state* relating these quantities; for example, the commonly imposed restrictions

$$q^a = \pi_{ab} = 0 \ \Rightarrow \ T_{ab} = \mu\, u_a\, u_b + p\, h_{ab} \tag{16}$$

characterise a 'perfect fluid' with, in general, equation of state $p = p(\mu, s)$. If in addition we assume that $p = 0$, we have the simplest case: pressure-free matter ('dust' or 'Cold Dark Matter'). Otherwise we must specify an equation of state determining p from μ and possibly other thermodynamical variables. Whatever these relations may be, we usually require that various *energy conditions* hold: one or all of

$$\mu > 0 \ , \ (\mu + p) > 0 \ , \ (\mu + 3p) > 0 \ , \tag{17}$$

(the latter, however, being violated by scalar fields in inflationary universe models), and additionally demand the *isentropic speed of sound* $c_s^2 = (\partial p/\partial \mu)_{s=\rm const}$ obeys

$$0 \le c_s^2 \le 1 \Rightarrow 0 \le \left(\frac{\partial p}{\partial \mu}\right)_{s=\rm const} \le 1 \ , \tag{18}$$

as required for local stability of matter (lower bound) and causality (upper bound), respectively.

These equations of state can be used to reduce the number of variables in $\mathcal{P}_{\text{ matter}}$; when they are not used in this way, they must be explicitly stated in a separate parameter space $\mathcal{P}_{\text{eos}}$ in $p_3(i)$. In broad terms

$$\begin{aligned}\mathcal{P}_{\text{eos}} \quad \equiv \quad & \{q_\alpha = q_\alpha(\mu, \Phi_A), \quad p = p(\mu, \Phi_A), \\ & \pi_{ab} = \pi_{ab}(\mu, \Phi_A),\ \dot{\Phi}_A = \dot{\Phi}_A(\Phi_A)\}. \end{aligned} \tag{19}$$

Given this information the equations become determinate and we can obtain the dynamical evolution of the models in the state space; see for example Wainwright and Ellis (1996), Hewitt et al (2003), Horwood et al (2003) for the case of Bianchi models (characterised by all the variables defined above depending on the time only) and Uggla et al (2003), Lim et al (2003) for the generic case.

Feature 5: However more general features may vary: the gravitational constant, the cosmological constant, and so on; and even the dimensions of space-time or the kinds of forces in operation. These are the parameters $\mathcal{P}_{\text{physics}}$ comprising $p_1(i)$ and $p_2(i)$. What complicates this issue is that some or many of these features may be emergent properties, resulting for example from broken symmetries occurring as the universe evolves. Thus they may come into being rather than being given as initial conditions that then hold for all time.

Initially one might think that considering all possible physics simply involves choices of coupling constants and perhaps letting some fundamental constant vary. But the issue is more fundamental than that. Taking seriously the concept of including *all* possibilities in the ensembles, the space of physical parameters $\mathcal{P}_{\text{physics}}$ used to describe $\mathcal{M}$, the parameters $p_2(i)$ might for example include a parameter $p_{\text{grav}}(i)$ such that: for $i = 1$ there is no gravity; for $i = 2$ there is Newtonian gravity; for $i = 3$ general relativity is the correct theory at all energies – there is no quantum gravity regime; for $i = 4$ loop quantum gravity is the correct quantum gravity theory; for $i = 5$ a particular version of superstring theory or M-theory is the correct theory.

Choices such as these will arise for all the laws and parameters of physics. In some universes there will be a fundamental unification of physics expressible in a basic "theory of everything", in others this will not be so. Some universes will be realised as branes in a higher dimensional spacetime, others will not.

2.4.1 Energy equation

It is worth commenting here that, because of the equivalence principle, there is no agreed energy conservation equation for the gravitational field itself, nor is there a definition of its entropy. Thus the above set of equations does not contain expressions for gravitational energy or entropy, and the concept of energy conservation does not play the major role for gravitation that it does in the rest of physics, neither is there any agreed view on the growth of entropy of the gravitational field. However, energy conservation of the matter content of space-time, expressed by the divergence equation $\nabla_b T^{ab} = 0$, is of course of major importance.

If we assume a perfect fluid the energy conservation equation takes the form

$$\dot{\mu} + (\mu + p)3\frac{\dot{S}}{S} = 0. \quad (20)$$

with a (linear) γ-law equation of state, this shows that

$$p = (\gamma - 1)\,\mu\ ,\ \dot{\gamma} = 0\ \Rightarrow\ \mu = M/S^{3\gamma}\ ,\ \dot{M} = 0\ . \quad (21)$$

One can approximate ordinary fluids in this way with $1 \le \gamma \le 2$ in order that the causality and energy conditions are valid, with 'dust' and Cold Dark Matter ('CDM') corresponding to $\gamma = 1 \Rightarrow \mu = M/S^3$, and radiation to $\gamma = \frac{4}{3} \Rightarrow \mu = M/S^4$.

In the case of a mixture of non-interacting matter, radiation and CDM having the *same* 4-velocity, represented as a single perfect fluid, the total energy density is simply the sum of these components: $\mu = \mu_{\rm dust} + \mu_{\rm CDM} + \mu_{\rm radn}$. (NB: This is only possible in universes with spatially homogeneous radiation energy density, because the matter will move on geodesics which by the momentum conservation equation implies ${}^3\tilde{\nabla}_a p_{\rm radn} = 0 \Leftrightarrow^3 \tilde{\nabla}_a \mu_{\rm radn} = 0$. This will not be true for a general inhomogeneous or perturbed FLRW model, but will be true in exact FLRW and orthogonal Bianchi models.)

The pressure can still be related to the energy density by a γ-law as in (21) in this case of non-interacting matter and radiation, but γ will no longer be constant. A scalar field has a perfect fluid energy-momentum tensor if the surfaces $\{\phi = \text{const}\}$ are spacelike and we choose u^a normal to these surfaces. Then it approximates the equation of state (21) in the 'slow-rolling' regime, with $\gamma \approx 0$, and in the velocity-dominated regime, with $\gamma \approx 2$. In the former case the energy conditions are no longer valid, so 'inflationary' behaviour is possible, which changes the nature of the attractors in the space of space-times in an important way.

2.5 The Anthropic subset

We are interested in the subset of universes that allow intelligent life to exist. That means we need a function on the set of possible universes that describes the probability that life may evolve. An adaptation of the Drake equation gives for the expected number of planets with intelligent life in any particular universe m in an ensemble,

$$N_{\rm life}(m) = N_g * N_S * \Pi * F, \quad (22)$$

where N_g is the number of galaxies in the model and N_S the average number of stars per galaxy, the probability that a star provides a habitat for life is expressed by the product

$$\Pi = f_S * f_p * n_e \quad (23)$$

and the probability of coming into existence of life, given such a habitat, is expressed by the product

$$F = f_l * f_i. \quad (24)$$

Here f_S is the fraction of stars that can provide a suitable environment for life (they are 'Sun-like'), f_p is the fraction of such stars that are surrounded by planetary systems, n_e is the mean number of planets in each such system that are suitable habitats for life (they are 'Earth-like'), f_l is the fraction of such planets on which life actually originates, and f_i represents the fraction of those planets on which there is life where intelligent beings develop. The anthropic subset of a possibility space is that set of universes for which $N_{\text{life}}(m) > 0$.

The quantities $\{N_g, N_S, f_S, f_p, n_e, f_l, f_i\}$ are functions of the physical and cosmological parameters characterised above, so there will be many different representations of this parameter set depending on the degree to which we try to represent such interrelations.

The astrophysical issues expressed in the product Π are the easier ones to investigate. We can in principle make a cut between those consistent with the eventual emergence of life and those incompatible with it by considering each of the factors in N_g, N_S, and Π in turn, taking into account their dependence on the parameters $p_1(i)$ to $p_5(i)$, and only considering the next factor if all the previous ones are non-zero (an approach that fits in naturally with Bayesian statistics and the successive allocation of relevant priors). In this way we can assign " bio-friendly intervals" to the possibility space $\mathcal{M}$. If $N_g * N_S * \Pi$ is non-zero we can move on to considering similarly whether F is non-zero, based on the parameters $p_6(i)$ to $p_7(i)$ determining if true complexity is possible, which in turn depend on the physics parameters $p_1(i)$ in a crucial way that is not fully understood. It will be impossible at any stage to characterise that set of the multiverse in which *all* the conditions *necessary* for the emergence of self-conscious life and its maintenance have been met, for we do not know what those conditions are (for example, we do not know if there are forms of life possible that are not based on carbon and organic chemistry). Nevertheless it is clear that life demands unique combinations of many different parameter values that must be realised simultaneously. When we look at these combinations, they will span a very small subset of the whole parameter space.

2.6 Parameter space revisited

It is now clear that some of the parameters discussed above are dependent on other ones, so that while we can write down a more or less complete set at varying levels of detail they will in general not be an independent set. There is a considerable challenge here: to find an independent set. *Inter alia* this involves solving both the initial value problem for general relativity and the way that galactic and planetary formation depend on fundamental physics constants (which for example determine radiation transfer properties in stars and in proto-planetary gas clouds), as well as relations there may be between the fundamental constants and the way the emergent complexity of life depends on them. We are a long way from understanding all these issues; allowing the cosmological parameters to vary simultaneously can change the predictions enormously. This means we can provide necessary sets of parameter values but cannot guarantee completeness or independence.

3 Exact solutions: FLRW models

A particularly important involutive subspace is that of the Friedmann–Lema ître ('FL') universes, the standard models of cosmology, based on the everywhere-isotropic Robertson–Walker ('RW') geometry. It is characterised by a perfect fluid matter tensor and the condition that *local isotropy* holds everywhere:

$$0 = \dot{u}^a = \sigma_{ab} = \omega^a \;\Leftrightarrow\; 0 = E_{ab} = H_{ab} \;\Rightarrow\; 0 =^3 \tilde{\nabla}_a \mu =^3 \tilde{\nabla}_a p =^3 \tilde{\nabla}_a \Theta \;, \quad (25)$$

the first conditions stating the kinematical quantities are locally isotropic, the second that these universes are conformally flat, and the third that they are spatially homogeneous, thus showing that isotropy everywhere implies spatial homogeneity in this case. We develop it here to show the dynamics and observational relations in this model.

3.1 Coordinates and metric

It follows then that:

1. Comoving coordinates can be found so that the metric takes the form

$$ds^2 = -\,dt^2 + S^2(t)\,(\,dr^2 + f^2(r)\,d\Omega^2\,)\;,\; u^a = \delta^a{}_0\;, \quad (26)$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta\,d\phi^2$, $u_a = -\,\nabla_a t$, and $\dot{S}/S = \frac{1}{3}\,\Theta$, characterising $S(t)$ as the scale factor for distances between any pair of fundamental observers. The expansion of matter depends only on one scale length, so it is isotropic (there is no distortion or rotation).

2. The Ricci tensor ${}^3R_{ab}$ is isotropic, so the 3-spaces $\{t = \text{const}\}$ are 3-spaces of constant curvature k/S^2 where k can be normalised to ± 1, if it is non-zero. Using the geodesic deviation equation in these 3-spaces, one finds that

$$f(r) = \sin r\;,\; r\;,\; \sinh r \text{ if } k = +1\;,\; 0\;,\; -\,1\;. \quad (27)$$

Thus when $k = +1$ the surface area $4\pi\,S^2(t)\,f^2(r)$ of a geodesic 2-sphere in these spaces, centred on the (arbitrary) point $r = 0$, increases to a maximum at $r = \pi/2$ and then decreases to zero again at the antipodal point $r = \pi$; hence the point at $r = 2\pi$ has to be the same point as $r = 0$, and these 3-spaces are necessarily closed, with finite total volume. In the other cases the 3-spaces are usually unbounded and the surface areas of these 2-spaces increase without limit; however, unusual topologies still allow the spatial sections to be closed.

3.2 Dynamical equations

The remaining non-trivial equations are the energy equation (20), the Raychaudhuri equation

$$3\,\frac{\ddot{S}}{S} + \frac{1}{2}\,(\mu + 3p) = 0\;, \quad (28)$$

and the Friedmann equation

$$\mu - \frac{1}{3}\,\Theta^2 = \frac{3\,k}{S^2}\ , \tag{29}$$

where k is a constant. Any two of these equations imply the third if $\dot{S} \neq 0$ (the latter equation being a first integral of the other two). All one has to do then to determine the dynamics is to solve the Friedmann equation. The solution depends on what form is assumed for the matter: Usually it is taken to be a perfect fluid with equation of state $p = p(\mu)$, or as a sum of such fluids, or as a scalar field with given potential $V(\phi)$. For the γ-law discussed above, the energy equation integrates to give (21), which can then be used to represent μ in the Friedmann equation.

On using the obvious tetrad above, all the other $1 + 3$ covariant and tetrad equations are identically true when these equations are satisfied.

3.2.1 Basic parameters

As well as the parameters $H_0 = (\dot{S}/S)_0$, $\Omega_0 = (\kappa\mu_0/3H_0^2)$, $\Omega_\Lambda = (\Lambda/3H_0^2)$ and $q_0 = (3\ddot{S}/SH_0^2)$, the FLRW models are characterised by the spatial curvature parameter $K_0 = k/S_0^2 = {}^3R_0/6$. These parameters are related by the equations (28) and (29).

3.2.2 Singularity and ages

The existence of the big bang, and age limits on the universe, follow directly from the Raychaudhuri equation, together with the energy assumption $(\mu + 3p) > 0$ (true at least when quantum fields do not dominate), because the universe is expanding today ($\Theta_0 > 0$). That is, the singularity theorem above applies in particular to FLRW models. Furthermore, from the Raychaudhuri equation, in any FLRW model, the fundamental age relation holds:

> **Age Theorem**: In an expanding FLRW universe with vanishing cosmological constant and satisfying the active gravitational mass density energy condition, ages are strictly constrained by the Hubble expansion rate: namely, at every instant, the age t_0 of the universe (the time since the big bang) is less than the inverse Hubble constant at that time:
>
> $$(\mu + 3p) > 0\ ,\ \Lambda = 0\ \Rightarrow\ t_0 < 1/H_0\ . \tag{30}$$

More precise ages $t_0(H_0, \Omega_0)$ can be determined for any specific cosmological model from the Friedmann equation (29); in particular, in a matter-dominated early universe the same result will hold with a factor 2/3 on the right-hand side, while in a radiation dominated universe the factor will be 1/2. Note that this relation applies in the early universe when the expansion rate was much higher, and, hence, shows that the hot early epoch ended shortly after the initial singularity .

3.3 Exact and approximate solutions

If $\Lambda = 0$ and the energy conditions are satisfied, FLRW models expand forever from a big bang if $k = -1$ or $k = 0$, and recollapse in the future if $k = +1$. A positive value of Λ gives a much wider choice for behaviours

3.3.1 Simplest models

a) *Einstein static*: $S(t)$ =const, $k = +1$, $\Lambda = \frac{1}{2}(\mu + 3p) > 0$, where everything is constant in space and time, and there is no redshift. This model is unstable (see above).
b) *de Sitter*: $S(t) = S_{\rm unit} \exp(H\,t)$, H = const, $k = 0$, a steady state solution in a constant curvature space-time: it is empty, because $(\mu + p) = 0$, i.e. it does not contain ordinary matter, but rather a cosmological constant,[3] or a scalar field in the strict 'no-rolling' case. It has ambiguous redshift because the choice of families of worldlines and space sections is not unique in this case.
c) *Milne*: $S(t) = t$, $k = -1$. This is flat, empty space-time in expanding coordinates (again $(\mu + p) = 0$).
d) *Einstein–de Sitter*: the simplest non-empty expanding model, with

$$k = 0 = \Lambda, \;\; p = 0, \;\; S(t) = a\,t^{2/3} \;, \;\; a = const \quad .$$

$\Omega = 1$ is always identically true in this case (this is the critical density case that just manages to expand forever). The age of such a universe is $t_0 = 2/(3H_0)$; if the cosmological constant vanishes, higher density universes ($\Omega_0 > 1$) will have ages less than this, and lower density universes ($0 < \Omega_0 < 1$) ages between this value and (30).

3.3.2 Early-time solutions

At early times, when matter is relativistic or negligible compared with radiation, the equation of state is $p = \frac{1}{3}\mu$ and the curvature term can be ignored. The solution is

$$S(t) = c\,t^{1/2} \;, \; c = \text{const} \;, \; \mu = \frac{3}{4}\,t^{-2} \;, \; T = \left(\frac{3}{4a}\right)^{1/4} \frac{1}{t^{1/2}} \;, \qquad (31)$$

which determines the expansion time scale during nucleosynthesis and so the way the temperature T varies with time (and hence determines the element fractions produced), and has no adjustable parameters. Consequently the degree of agreement attained between nucleosynthesis theory based on this time scale and element abundance observations may be taken as supporting both a FLRW geometry and the validity of the EFE at that epoch.

The standard thermal history of the hot early universe follows; going back in time, the temperature rises indefinitely (at least until an inflationary or quantum-dominated epoch occurs), so that the very early universe is an opaque

[3] A fluid with $(\mu + p) = 0$ is equivalent to a cosmological constant.

near-equilibrium mixture of elementary particles that combine to form nuclei, atoms, and then molecules after pair production ends and the mix cools down as the universe expands, while various forms of radiation (gravitational radiation, neutrinos, electromagnetic radiation) successively decouple and travel freely through the universe that has become transparent to them. This picture is very well supported by the detection of the extremely accurate black body spectrum of the CBR, together with the good agreement of nucleosynthesis observations with predictions based on the FLRW time scales (31) for the early universe.

3.3.3 Scalar field

The inflationary universe models use many approximations to model a FLRW universe with a scalar field ϕ as the dominant contribution to the dynamics, so allowing accelerating models that expand quasi-exponentially through many efoldings at a very early time, possibly leading to a very inhomogeneous structure on very large (super-particle-horizon) scales where decay of the scalar field to radiation happens at different times. This then leads to important links between particle physics and cosmology, and there is a very large literature on this subject. If an inflationary period occurs in the very early universe, the matter and radiation densities drop very close to zero while the inflaton field dominates, but is restored during 'reheating' at the end of inflation when the scalar field energy converts to radiation.

This will not be pursued further here, except to make one point: because the potential $V(\phi)$ is unspecified (the nature of the inflaton is not known) and the initial value of the 'rolling rate' $u\phi$ can be chosen at will, it is possible to specify a precise procedure whereby any desired evolutionary history $S(t)$ is attained by appropriate choice of the potential $V(\phi)$ and the initial 'rolling rate' (see [11] for details). Thus, inflationary models may be adjusted to give essentially any desired results in terms of expansion history.

3.3.4 Kinetic theory

While a fluid description is used most often, it is also of interest to use a kinetic theory description of the matter in the universe. The details of collisionless isotropic kinetic models in a FLRW geometry are given by Ehlers, Geren and Sachs [4].

3.4 Phase planes

¿From these equations, as well as finding simple exact solutions, one can determine evolutionary phase planes for this family of models; see Refsdal and Stabell [25] for (Ω_m, q_0), Ehlers and Rindler [5] for $(\Omega_m, \Omega_r, q_0)$, Wainwright and Ellis [30] for (Ω_0, H_0), and Madsen and Ellis [23] for (Ω, S). The latter are based on the phase-plane equation

$$\frac{d\Omega}{dS} = -\,(3\gamma - 2)\,\frac{\Omega}{S}\,(1 - \Omega)\ . \tag{32}$$

This equation is valid for any γ, i.e. for arbitrary relations between μ and p, but gives a (Ω, S) phase plane flow if $\gamma = \gamma(\Omega, S)$, and in particular if $\gamma = \gamma(S)$ or γ =const. Non-static solutions can be followed through turnaround points where $uS = 0$ (and so Ω is infinite). This enables one to attain complete (time-symmetric) phase planes for models with and without inflation; see [23] for details.

3.5 Observations

Astronomical observations are based on radiation travelling to us on the *geodesic null rays* that generate our *past light cone*. In the case of a FLRW universe, we may consider only radial null rays as these are generic (because of spatial homogeneity, we can choose the origin of coordinates on any light ray of interest; because of isotropy, light rays travelling in any direction are equivalent to those travelling in any other direction). Thus we may consider geodesic null rays travelling in the FLRW metric (26) such that $ds^2 = 0 = d\theta = d\phi$; then it follows that $0 = -\,dt^2 + S^2(t)\,dr^2$ on these geodesics. Hence, radiation emitted at E and received at O obeys the basic relations

$$r = \int_E^O dr = \int_{t_E}^{t_0} \frac{dt}{S(t)} = \int_{S_E}^{S_0} \frac{dS}{S(t)\,\dot{S}(t)}\ , \tag{33}$$

where the term $\dot{S}$ may be found from the Friedmann equation (29), once a suitable matter description has been chosen.

3.5.1 Redshift

The first fundamental observational quantity is *redshift*. Considering two successive pulses sent from E to O, each remaining at the same comoving coordinate position, it follows from (33) that the cosmological redshift in a FLRW model is given by

$$(1 + z_c) = \frac{\lambda_0}{\lambda_E} = \frac{\Delta T_0}{\Delta T_E} = \frac{S(t_0)}{S(t_E)}\ , \tag{34}$$

and so directly measures the expansion of the universe between when light was emitted and when it is received. Two comments are in order. First, redshift is essentially a time-dilation effect, and will be apparent in all observations of a source, not just in its spectra; this characterisation has the important consequences that (i) redshift is achromatic — the fractional shift in wavelength is independent of wavelength, (ii) the width of any emitted frequency band $d\nu_E$ is altered proportional to the redshift when it reaches the observer, i.e. the observed width of the band is $d\nu_0 = (1 + z)\,d\nu_E$, and (iii) the observed rate of emission of radiation and the rate of any time variation in its intensity will both also be proportional to $(1+z)$. Second, there can be local gravitational and Doppler contributions z_0 at the observer, and z_E at the emitter; observations of spectra tell us the overall redshift z, given by

$$(1+z) = (1+z_0)\,(1+z_c)\,(1+z_E)\ , \tag{35}$$

but cannot tell us what part is cosmological and what part is due to local effects at the source and the observer. The latter can be determined from the CBR anisotropy, but the former can only be estimated by identifying cluster members and subtracting off the mean cluster motion. The essential problem is in identifying which sources should be considered members of the same cluster.

3.5.2 Areas

The second fundamental issue is apparent size. Considering light rays converging to the observer at time t_0 in a solid angle $d\Omega = \sin\theta\, d\theta\, d\phi$, from the metric form (26) the corresponding null rays[4] will be described by constant values of θ and ϕ and at the time t_E will encompass an area $dA = S^2(t_E) f^2(r) d\Omega$ orthogonal to the light rays, where r is given by (33). Thus, on defining the *observer area distance* $r_0(z)$ by the standard area relation, we find

$$dA = r_0^2\, d\Omega\ \Rightarrow\ r_0^2 = S^2(t_E)\, f^2(r)\ . \tag{36}$$

Because these models are isotropic about each point, the *same* distance will relate the observed angle α corresponding to a linear length scale ℓ orthogonal to the light rays:

$$\ell = r_0\, \alpha\ . \tag{37}$$

One can now calculate r_0 from this formula together with (33) and the Friedmann equation, or from the geodesic deviation equation (see [14]), to obtain for a non-interacting mixture of matter and radiation,

$$r_0(z) = \frac{1}{H_0 q_0 (q_0 + \beta - 1)}\, \frac{\left[\ (q_0 - 1)\left\{1 + 2q_0 z + q_0 z^2 (1-\beta)\right\}^{1/2} - (q_0 - q_0\beta z - 1)\ \right]}{(1+z)^2}\ , \tag{38}$$

where β represents the matter to radiation ratio: $(1-\beta)\,\rho_{m0} = 2\,\beta\,\rho_{r0}$. The standard Mattig relation for pressure-free matter is obtained for $\beta = 1$ and the corresponding radiation result for $\beta = 0$.

An important consequence of this relation is refocusing of the past light cone: the universe as a whole acts as a gravitational lens, so that there is a redshift z_* such that the area distance reaches a maximum there and then decreases for larger z; correspondingly, the apparent size of an object of fixed size would reach a minimum there and then increase as the object was moved further away. As a specific example, in the simplest (Einstein–de Sitter) case with $p = \Lambda = k = 0$, we find

$$\beta = 1\ ,\ q_0 = \frac{1}{2}\ \Rightarrow\ r_0(z) = \frac{2}{H_0}\, \frac{1}{(1+z)^{3/2}}\, (\,\sqrt{1+z} - 1\,)\ , \tag{39}$$

[4]Bounded by geodesics located at (ϕ_0, θ_0), $(\phi_0 + d\phi, \theta_0)$, $(\phi_0, \theta_0 + d\theta)$, $(\phi_0 + d\phi, \theta_0 + d\theta)$.

which refocuses at $z_* = 5/4$; objects further away will look the same size as much closer objects. For example, an object at a redshift $z_1 = 1023$ (i.e. at about last scattering) will appear the same angular size as an object of identical size at redshift $z_2 = 0.0019$ (which is very close — it corresponds to a speed of recession of about 570 km/ sec). In a low density universe, refocusing takes place further out, at redshifts up to $z \approx 4$, depending on the density, and with apparent sizes depending on possible source size evolution.

The predicted (*angular size, distance*)–relations are difficult to test observationally because objects of more or less fixed size (such as spherical galaxies) do not have sharp edges that can be used for measuring angular size and so one has rather to measure isophotal diameters, while objects with well-defined linear dimensions, such as double radio sources, are usually rapidly evolving and so one does not know their intrinsic size. Thus, these tests, while in principle clean, are in fact difficult to use in practice.

3.5.3 Luminosity and reciprocity theorem

There is a remarkable relation between upgoing and downgoing bundles of null geodesics connecting the source at t_E and the observer at t_0. Define *galaxy area distance* r_G as above for observer area distance but for the upgoing rather than downgoing bundle of null geodesics. The expression for this distance will be exactly the same as (36) except that the times t_E and t_0 will be interchanged. Consequently, on using the redshift relation (34),

> **Reciprocity Theorem**: The observer area distance and galaxy area distance are identical up to redshift factors:
>
> $$\frac{r_0^2}{r_G^2} = \frac{1}{(1+z)^2} \ . \qquad (40)$$

This is true in any space-time as a consequence of the standard first integral of the geodesic deviation equation [6].

Now from photon conservation, the *flux of light* received from a source of *luminosity* L at time t_E will be measured to be

$$F = \frac{L(t_E)}{4\pi} \frac{1}{(1+z)^2} \frac{1}{S\,{}^2_0, f^2(r)} \frac{1}{r_G^2} \ ,$$

with r given by (33), and the two factors $(1+z)$ coming from photon redshift and time dilation of the emission rate, respectively. On using the reciprocity result this becomes

$$F = \frac{L(t_E)}{4\pi} \frac{1}{(1+z)^4} \frac{1}{r_0^2} \ , \qquad (41)$$

where $r_0(q_0, z)$ is given by (38). On taking logarithms, this gives the standard (*luminosity, redshift*)–relation of observational cosmology. Observations of this Hubble relation basically agree with these predictions, but are not accurate enough to distinguish between the various FLRW models. The hopes that this relation would determine q_0 from galaxy observations have faded away because of the major problem of *source evolution*: we do not know what the source luminosity would have been at the time of evolution. We lack standard candles of known luminosity (or equivalently, rigid objects of known linear size, from which apparent size measurements would give the answer). Various other distance estimators such as the Tully–Fisher relation have helped considerably, but not enough to give a definitive answer. Happily it now seems that Type Ia supernovae can provide the answer, because their luminosity can be determined from their light curves, which should depend only on local physics rather than their evolutionary history.

3.5.4 Specific intensity

In practice, we measure (a) in a limited waveband rather than over all wavelengths as the 'bolometric' calculation above suggests; and (b) real detectors measure specific intensity (radiation received per unit solid angle) at each point of an image, rather than total source luminosity. Putting these together, we see that if the *source spectrum* is $\mathcal{I}(\nu_E)$, i.e. a fraction $\mathcal{I}(\nu_E)\, d\nu_E$ of the source radiation is emitted in the frequency range $d\nu_E$, then the observed *specific intensity* at each image point is given by [5]

$$I_\nu \, d_\nu = \frac{B_E}{(1+z)^3}\, \mathcal{I}(\nu(1+z))\, d\nu \ , \tag{42}$$

where B_E is the *surface brightness* of the emitting object, and the area distance $r_0(z)$ has canceled out (because of the reciprocity theorem). This tells us the apparent intensity of radiation detected in each direction — which is independent of (area) distance, and dependent only on the source redshift, spectrum, and surface brightness. Together with the angular diameter relation (37), this determines what is actually measured by a detector.

An immediate application is black body radiation: if any radiation is emitted as black body radiation at temperature T_E, it follows from the black body expression $\mathcal{I}_\nu = \nu^3\, b(\nu/T_E)$ that the received radiation will also be black body (i.e. have this same black body form) but with a measured temperature of

$$T_0 = \frac{T_E}{(1+z)} \ . \tag{43}$$

Note this is true in all cosmologies: the result does not depend on the FLRW symmetries. The importance of this, of course, is that it applies to the observed CBR.

3.5.5 Number counts

If we observe sources in a given solid angle $d\Omega$ in the distance range $(r, r+dr)$, the corresponding volume is $dV = S^3(t_E)\, r_0^2\, dr\, d\Omega$, so if the source density is

[5] Absorption effects will modify this if there is sufficient absorbing matter present; see [6] for relevant formulae.

$n(t_E)$ and the probability of detection is p, the number of sources observed will be

$$dN = p\, n(t_E)\, dV = p \left[\frac{n(t_E)}{(1+z)^3} \right] S^3(t_0)\, f^2(r)\, dr\, d\Omega \ , \tag{44}$$

with r given by (33). This is the *basic number count relation*, where dr can be expressed in terms of observable quantities such as dz; the quantity in brackets is constant if source numbers are conserved in a FLRW model, that is

$$n(t_E) = n(t_0)(1+z)^3 \, . \tag{45}$$

The FLRW predictions agree with observations only if we allow for source number and/or luminosity evolution (cf. the discussion of spherically symmetric models in the next section); but we have no good theory for source evolution.

The additional problem is that there are many undetectable objects in the sky, including entire galaxies, because they lie below the detection threshold; thus we face the problem of *dark matter*, which is very difficult to detect by cosmological observations except by its lensing effects (if it is clustered) and its effects on the age of the universe (if it is smoothly distributed). The current view is that there is indeed such dark matter, detected particularly through its dynamical effects in galaxies and clusters of galaxies, with the present day total matter density most probably in the range $0.1 \leq \Omega_0 \leq 0.3$, while the baryon density is of the order of $0.01 \leq \Omega_0^{\mathrm{baryons}} \leq 0.03$ (from nucleosynthesis arguments). Thus, most of the dark matter is probably non-baryonic.

3.5.6 Other observations and the concordance model

Good cosmological models must fit (a) all available astronomical observations of the type characterised above: number counts, magnitude-redshift and angular diameter-redshift measurements, etc, in particular the recent estimiates of the deceleration parameter from observations of Type Ia supernovae in distant galaxies. They must also (b) provide agreement between nucleosynthesis predictions and element abundance measurments, stellar age estimates and the age of the universe, and (c) give agreement between the cosmic background radiation anisotropy spectrum and observations of the power spectrum of inhomogeneities in the matter distribution, based on a theoretical model of structure formation in the early universe.

We do indeed get good agreement of all these factors for a ‘concordance model’ - a perturbed FLRW universe that at the present time has a non-zero effective cosmological constant and a large preponderance of dark matter, with nearly flat space sections: $\Omega_m \simeq 0.3$, $\Omega_{baryons} \simeq 0.02$, $\Omega_\Lambda \simeq 0.7$, $\Omega_0 \simeq 1$ but with marginal evidence that $\Omega_0 > 1$.The structure formation models are premised on an inflationarly early era, but the nature of the field causaing inflation - the inflaton - is completely unknown, also the nature of the non-baryonic dark matter present is unknown as is the nature of the present-day effective cosmological constant (is it in fact constant or - as more usually supposed - in fact an effective scalar field?)

3.6 Observational limits

The first basic observational limit is that we cannot observe anything outside our past light cone, given by (33). Combined with the finite age of the universe, this leads to a maximum comoving coordinate distance from the origin for matter with which we can have had any causal connection: namely

$$r_{ph}(t_0) = \int_0^{t_0} \frac{dt}{S(t)} \ , \tag{46}$$

which converges for any ordinary matter. Matter outside is not visible to us, indeed we cannot have had any causal contact with it. Consequently (see Rindler [26]), the particles at this comoving coordinate value define the *particle horizon*: they separate that matter which can have had any causal contact with us since the origin of the universe from that which cannot. This is most clearly seen by using Penrose's conformal diagrams, obtained on using as coordinates the comoving radius and conformal time; see Penrose [24] and Tipler, Clarke and Ellis [27]. The present day distance to the horizon is

$$D_{ph}(t_0) = S(t_0)\, r_{ph} = S(t_0) \int_0^{t_0} \frac{dt}{S(t)} \ . \tag{47}$$

From (33), this is a sphere corresponding to infinite measured redshift (because $S(t) \to 0$ as $t \to 0$).

Once comoving matter has entered the particle horizon, it cannot leave it (i.e. once causal contact has been established in a FLRW universe, it cannot cease).

Actually we cannot even see as far as the particle horizon: on our past light cone information rapidly fades with redshift (because of (42)); and because the early universe is opaque, we can only see (by means of any kind of electromagnetic radiation) to the *visual horizon* (Ellis and Stoeger [13]), which is the sphere at comoving coordinate distance

$$r_{vh}(t_0) = \int_{t_d}^{t_0} \frac{dt}{S(t)} \ , \tag{48}$$

where t_d is the time of *decoupling of matter and radiation*, when the universe became transparent (at about a redshift of $z = 1100$). The matter we see at that time is the matter which emitted the CBR we measure today with a present temperature of $2.73\,^{\circ}$K; its present distance from us is

$$D_{vh}(t_0) = S(t_0)\, r_{vh} \ . \tag{49}$$

If we evaluate these quantities in an Einstein–de Sitter universe, we find an interesting result: the present day distance to the particle horizon is $D_{ph}(t_0) = 3\, ct_0$.

Finally it should be noted that an early inflationary era will move the particle horizon out to very large distances, thus solving the causal problem presented by the isotropy of CBR arriving here from causally disconnected regions, but it will have no effect on the visual horizon. Thus, it changes the causal limitations, but does not affect the visual limits on the part of the universe we can see.

3.6.1 Small universes

The existence of visual horizons represent absolute limits on what we can ever know; because of them, we can only hope to investigate a small fraction of all the matter in the universe. Furthermore, they imply we do not in fact have the data needed to predict to the future, for at any time gravitational radiation from as yet unseen objects (e.g. domain walls in a chaotic inflationary universe) may cross the visual horizon and undermine any predictions we may have made. However, there is one exceptional situation: it is possible we live in a small universe, with a spatially closed topology on such a length scale (say, 300 to 800 Mpc) that we have already seen around the universe many times, thus already having seen all the matter there is in the universe. The effect is like being in a room with mirrors on the floor, ceiling, and all walls; images from a finite number of objects seem to stretch to infinity. There are many possible topologies, whatever the sign of k; the observational result — best modelled by considering many identical copies of a basic cell attached to each other in an infinitely repeating pattern — can be very like the real universe. In this case we would be able to see our own galaxy many times over, thus being able to observationally examine its historical evolution once we had identified which images of distant galaxies were in fact repeated images of our own galaxy.

It is possible the real universe is like this. Observational tests can be carried out by trying to identify the same cluster of galaxies, QSO's, or X-ray sources in different directions in the sky; or by detecting circles of identical temperature variation in the CBR sky. If no such circles are detected, this will be a reasonably convincing proof that we do not live in such a small universe — which has various philosophical advantages over the more conventional models with infinite spatial sections. Inter alia they can give reduced power on large scales so giving a simple explanation of the low CBR quadrupole anisotropy.

3.7 FLRW universes as cosmological models

These models are very successful in explaining the major features of the observed universe — its expansion from a hot big bang leading to the observed galactic redshifts and remnant black body radiation, tied in well with element abundance predictions and observations. However, these models do not describe the real universe well in an essential way, in that the highly idealized degree of symmetry does not correspond to the lumpy real universe. Thus, they can serve as basic models giving the largest-scale smoothed out features of the observable physical universe, but one needs to perturb them to get realistic ('almost-FLRW') universe models that can be used to examine the inhomogeneities and anisotropies arising during structure formation, and that can be compared in detail with observations.

However, there is a major underlying issue: because of their high symmetry, these models are infinitely improbable in the space of all possible universes. This high symmetry represents a very high degree of fine tuning of initial conditions, which is extraordinarily improbable, unless we can show physical reasons why it should develop from much more general conditions. In order to examine that question, one needs to look at much more general models and see if they do indeed evolve towards the FLRW models because of physical processes. Additionally, while the FLRW models seem good models for the observed universe at the present time, one can ask (a) are they the only possible models that will fit the observations? (b) does the universe necessarily have the same symmetries on very large scales (outside the particle horizon), or at very early and/or very late times?

To study these issues, we need to look at more general models, developing some understanding of their geometry and dynamics. Indeed there is a range of models in addition to the FLRW models that can fulfill all present day observational requirements. Nevertheless, it is important to state that the family of perturbed FLRW models can meet all present observational requirements, *provided* we allow suitable evolution of source properties back in the past. They also provide a powerful theoretical framework for considering the nature of and effects of cosmic evolution. Hence, they are justifiably the standard models of cosmology. No evidence stands solidly against them.

4 Dynamical evolution

We now consider aspects of more general models. Firstly, we have a good classification of symmetries of cosmological models; this is discussed in the Appendix. Secondly, we have a good analysis of some particular families of models in addition to the FLRW models discussed above, where we can characterise the geometry in detail and also obtain partial or full integration of the dynamical equations. The major such families of models are also briefly discussed in the Appendix.

4.1 Phase planes

Given the dynamical equation soutlined above, one can obtain dynamical systems representations for large families of cosmological models. This will be discussed by John Wainwright. In particular one can determine attractors, basins of attraction, and saddle points for these families, thus determining which are special and which are general models. A word of cauation here: proper assessment of probability and generality depends on a well-founded measure on the space of models. We do not have such a measure. Nevertheless the phase planes probably give reliable indications of the issue of speciality and generality.

I will just pick up two particular issues here. Firstly, it appears form the structure of the phase planes that the higher symmetry models provide a skeleton which guides the dynamical evolution of the lower symmetry models. This will be discussed by John Wainwright. Secondly, the FLRW models are saddle points in many of these phase planes. This means that intermediate isotropisation occurs: models that are quite unlike the FLRW models both at very early and at very late times can spend a very long time in a state very close to that of specific FLRW models, before diverging to a completely different anisotropic state. Such models can be good models of the real universe, as they will have the same observational properties as (perturbed) FLRW models if they have been close to a FLRW state since just before the epoch of nucleosynthesis.

4.2 Ensembles of models

Ensembles of models are characterised by distribution functions and measures on the space of models. Much has been written about ensembles in a hand-waving way but little has been done about such distribution functions except to a certain extent in the case of chaotic inflationary FLRW universe models. An in-depth examination of the evolution of distribution functions could be a rewarding exercise.

4.2.1 Basic singularity theorem

The issue of whether there was or was not a singularity in the early universe is of prime importance. Using the definition of S, the Raychaudhuri equation for a generic model can be rewritten in the form

$$3\,\frac{\ddot{S}}{S} = -\,2\,(\sigma^2 - \omega^2) +^3 \tilde{\nabla}_a \quad \dot{u}^a + (\dot{u}_a \dot{u}^a) - \frac{1}{2}\,(\mu + 3p) + \Lambda \ , \tag{50}$$

showing how the curvature of the curve $S(\tau)$ along each worldline (in terms of proper time τ along that worldline) is determined by the kinematical quantities, the total energy density and pressure in the combination $(\mu + 3p)$, and the cosmological constant Λ. This gives the basic

> **Singularity Theorem**: In a universe where $(\mu{+}3p) > 0$, $\Lambda \leq 0$, and $\dot{u}^a = \omega^a = 0$ at all times, at any instant when $H_0 = \frac{1}{3}\,\Theta_0 > 0$, there must have been a time $t_0 < 1/H_0$ ago such that $S \to 0$ as $t \to t_0$; a space-time singularity occurs there, where $\mu \to \infty$ and $p \to \infty$ for ordinary matter (with $(\mu + p) > 0$).

The further singularity theorems of Hawking and Penrose utilize this result (and its null version) as an essential part of their proofs. In effect this is the statement that if we follow the universe back far enough, under these conditions it must have entered a quantum gravity regime; quantum gravity may or then may not avoid the singularity.

However inflationary models violate the energy condition crucial to this theorem. Thus they can lead to singularity avoidance.

5 Singularities and non-singular models

We now look at the existence of singularity free solutions, see [10, 12], and pose the issue of the tension between very special initial conditions and the existence of singularities.

5.1 Eternal inflationary models

Here we consider closed models in which $K = +1$ and H can become zero. The models are simple, obey general relativity, and contain ordinary matter and (minimally coupled) scalar fields. Previous examples of closed inflationary models are, to our knowledge, either bouncing models or models in which inflation is preceded by deceleration. The $K = +1$ bouncing universe collapses from infinity and then turns around at t_* to expand in an inflationary phase, followed by a standard hot big bang evolution. The canonical model for such a bounce is the de Sitter universe in the $K = +1$ frame, with $a(t) = a_* \cosh Ht$. These coordinates cover the whole spacetime, which is geodesically complete. However, the bouncing models face a number of difficulties as realistic cosmologies. In particular, the initial state is hard to motivate (collapsing from infinite size without causal interaction), and it is also difficult to avoid nonlinearities in the collapse that prevent a regular bounce.

The models we present have a finite initial size and no bounce. The simplest versions are ever-expanding Eddington-Lemaître type models, with a finite amount of inflation occurring over an infinite time. The redshift and the total number of e-folds remain bounded through the expansion of the universe until the present day.

We assume that the early universe contains a scalar field ϕ with energy density $\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi)$ and pressure $p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)$, and possibly also matter with energy density ρ and pressure $p = w\rho$, where $-\frac{1}{3} \le w \le 1$. The cosmological constant is absorbed into the potential V. In the absence of interactions between matter and the scalar field, they separately obey the energy conservation and Klein-Gordon equations,

$$\dot{\rho} + 3(1+w)H\rho = 0\,, \tag{51}$$

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,. \tag{52}$$

The Raychaudhuri field equation

$$\frac{\ddot{a}}{a} = -\frac{8\pi G}{3}\left[\frac{1}{2}(1+3w)\rho + \dot{\phi}^2 - V(\phi)\right], \tag{53}$$

has first integral the Friedmann equation,

$$H^2 = \frac{8\pi G}{3}\left[\rho + \frac{1}{2}\dot{\phi}^2 + V(\phi)\right] - \frac{K}{a^2}\,, \tag{54}$$

and they imply

$$\dot{H} = -4\pi G\left[\dot{\phi}^2 + (1+w)\rho\right] + \frac{K}{a^2}\,. \tag{55}$$

The Raychaudhuri equation gives the condition for inflation,

$$\ddot{a} > 0 \;\Leftrightarrow\; \dot{\phi}^2 + \frac{1}{2}(1+3w)\rho < V(\phi)\,. \tag{56}$$

For a positive minimum in the inflationary scale factor, $a_* = a(t_*) > 0$, we require

$$H_* = 0 \;\Leftrightarrow\; \frac{1}{2}\dot{\phi}_*^2 + V(\phi_*) + \rho_* = \frac{3K}{8\pi G a_*^2}\,, \tag{57}$$

where the time t_* may be infinite. The only way to satisfy Eq. (57) with non-negative energy densities is if $K = +1$. Closed inflationary models admit a minimum scale factor if inflation occurs for long enough, since curvature will eventually win over a slow-rolling scalar field as we go back into the past The inflationary singularity theorems mentioned above exclude this case, since they either only consider $K \leq 0$, or explicitly exclude the possibility $H_* = 0$.

Closed models with a minimum scale factor $a_* > 0$ include both bouncing and ever-expanding cases. A simple ever-inflating model is the closed model with radiation ($w = \frac{1}{3}$) and cosmological constant. The exact solution is

$$a(t) = a_* \left[1 + \exp\left(\frac{\sqrt{2}\,t}{a_*}\right)\right]^{1/2}. \tag{58}$$

In the infinite past, $t \to -\infty$, the model is asymptotically Einstein static, $a \to a_*$. Inflation occurs for an infinite time to the past, but at any finite time $t_\mathrm{e} \gg a_*$, there is a finite number of e-folds,

$$N_\mathrm{e} = \ln\frac{a_\mathrm{e}}{a_*} \approx \frac{t_\mathrm{e}}{\sqrt{2}a_*}\,. \tag{59}$$

The curvature parameter at t_e is strongly suppressed by the de Sitter-like expansion:

$$\Omega_\mathrm{e} - 1 \approx 2e^{-N_\mathrm{e}}\,. \tag{60}$$

The exact model Eq. (58) is a simple example of Eddington-Lemaî tre type solutions. There are trajectories of this type in the classical phase space satisfying Einstein's equations. They are past-asymptotically Einstein static and ever-expanding. However, in these models inflation does not end. Below we discuss more realistic models than Eq. (58), based on inflationary potentials, which include Eddington-Lemaître type models that do exit from inflation.

In our scenario, the inflationary universe emerges from a small static state that has within it the seeds for the emergence of the macroscopic universe. We call this the "Emergent Universe" scenario. These universes are singularity-free, without particle horizons, and ever-expanding ($H \geq 0$). Even though Emergent Universes admit closed trapped surfaces, these do not lead to a singularity, since $K = +1$ and the weak energy condition is violated in the past .

The Einstein static universe is characterized by $K = +1$ and $a = a_\mathrm{i} =$ const. Equations (51)–(55) then imply that

$$\frac{1}{2}(1 - w_\mathrm{i})\rho_\mathrm{i} + V(\phi) = \frac{1}{4\pi G a_\mathrm{i}^2}\,, \tag{61}$$

$$(1 + w_\mathrm{i})\rho_\mathrm{i} + \dot{\phi}_\mathrm{i}^2 = \frac{1}{4\pi G a_\mathrm{i}^2}\,, \tag{62}$$

where $\dot{\rho}_{\rm i} = 0 = \ddot{\phi}_{\rm i}$ and $V'(\phi) = 0$. Thus V is constant, $V = V_{\rm i}$, and $V_{\rm i} = \Lambda_{\rm i}/8\pi G$ is the vacuum energy, where $\Lambda_{\rm i}$ is the primordial cosmological constant. If the scalar field kinetic energy vanishes, i.e. if $\dot{\phi}_{\rm i} = 0$, then $(1 + w_{\rm i})\rho_{\rm i} > 0$, so that there must be matter to keep the universe static. If the static universe has only a scalar field, i.e. if $\rho_{\rm i} = 0$, then the field must have nonzero (but constant) kinetic energy, as it rolls along the flat potential. (Dynamically, this case is equivalent to a stiff fluid, $w_{\rm i} = 1$, plus cosmological constant $\Lambda_{\rm i}$)

The radius $a_{\rm i}$ of the initial static universe (where $a_{\rm i} = a_*$) can be chosen to be a very small scale, but above the Planck scale,

$$a_{\rm i} > M_{\rm p}^{-1}\,, \tag{63}$$

by suitable choice of $V_{\rm i}$, $\dot{\phi}_{\rm i}^2$ and $\rho_{\rm i}$ (with all of them $\ll M_{\rm p}^4$). A simple way to realize the scenario of the Emergent Universe is the following.

Consider a potential that is asymptotically flat in the infinite past,

$$V(\phi) \to V_{\rm i} \text{ as } \phi \to -\infty\,,\ t \to -\infty\,, \tag{64}$$

but drops towards a minimum at a finite value $\phi_{\rm f}$. The scalar field kinetic energy density is asymptotic to the constant Einstein static value,

$$\frac{1}{2}\dot{\phi}^2 \to \frac{1}{2}V_{\rm i} = \frac{1}{8\pi G a^2{}_{\rm i}} \text{ as } \phi \to -\infty\,,\ t \to -\infty\,, \tag{65}$$

where we used Eqs. (61) and (62) with $\rho_{\rm i} = 0$. Because $\dot{\phi}_{\rm i} \neq 0$, no matter is needed to achieve the initial static state. The field rolls from the Einstein static state at $-\infty$ and the potential slowly drops from its Einstein static original value. Provided that $\dot{\phi}^2$ decreases more rapidly than the potential, we have $V - \dot{\phi}^2 > 0$, so that the universe accelerates, by Eq. (56). Since $\ddot{\phi} < 0$ and $V' < 0$, while $\dot{\phi} > 0$, the Klein-Gordon equation (52) shows that the universe is expanding ($H > 0$).

Inflation ends at time $t_{\rm e}$, where $V_{\rm e} = \dot{\phi}_{\rm e}^2$. Then reheating takes place as the field oscillates about the minimum at $\phi_{\rm f}$. In the asymptotic past, $V \to V_{\rm i}$, the primordial cosmological constant $\Lambda_{\rm i} = 8\pi G V_{\rm i}$ is given by Eq. (65) as

$$\Lambda_{\rm i} = \frac{2}{a_{\rm i}^2}\,, \tag{66}$$

so that $\Lambda_{\rm i}$ is large for a small initial radius. At the minimum, $V_{\rm f}$ defines the cosmological constant that dominates the late universe,

$$\Lambda = 8\pi G V_{\rm f} \ \ll\ \Lambda_{\rm i}\,. \tag{67}$$

A typical example of a potential is

$$V - V_{\rm f} = (V_{\rm i} - V_{\rm f})\left[\exp\left(\frac{\phi - \phi_{\rm f}}{\alpha}\right) - 1\right]^2, \tag{68}$$

where α is a constant energy scale.

The infinite time of inflation, from $t = -\infty$ to $t = t_e$ ($< t_f$), produces a finite amount of inflation (and a finite redshift to the initial Einstein static state). The potential produces expansion that is initially qualitatively similar to the exact solution in Eq. (58), so that the total number of e-folds from the initial expansion can be estimated, following Eq. (59), as

$$N_e = O\left(\frac{t_e}{a_i}\right) . \tag{69}$$

Provided that a_i is chosen small enough and t_e large enough, a very large number of e-folds can be produced. The parameters in the potential in Eq. (68) can be chosen so that the primordial universe is consistent with current observations.

As with standard inflationary models, fine-tuning is necessary to produce density perturbations at the $O(10^{-5})$ level, and to fix Λ so that $\Omega_{\Lambda 0} \sim 0.7$. The fine-tuning problem in the Emergent models requires in particular a suitable choice of initial radius a_i, or equivalently the primordial cosmological constant, Eq. (66). However, in a closed universe, there could be dynamical mechanisms that adjust the cosmological constant, determined by the size of the universe. It is also worth pointing out that Casimir effects in the early Einstein static state could play an important role. It is not inconceivable that the combination of these effects could lead to determination of a specific Einstein static radius that is more stable than any other.

The initial Einstein static state also has the attractive feature that it is neutrally stable against inhomogeneous linear perturbations when $\rho_i = 0$ (as in the simple Emergent model above), or when $\rho_i > 0$ and the sound speed of matter obeys $c_s^2 > \frac{1}{5}$. This stability property means that the Einstein static model can be a natural initial state in the space of closed universes. The Einstein static state is of course unstable to *homogeneous* perturbations, which break the balance between curvature and energy density. This instability is crucial for producing an inflationary era.

The further question of how probable the initial Einstein static model is in the space of all FRW models involves the unresolved issue of a suitable measure in this space. There are indications that it could be marginally preferred by measures based on maximum entropy principles. The Einstein static universe with radiation and vacuum energy maximizes the entropy, as shown by Gibbons. We are not claiming that the Einstein static initial state is preferred, but it does lead to an interesting and novel alternative to the standard models.

5.2 Special models (fine tuning) versus singularities

We pose the issue of the tension between very special initial conditions and the existence of singularities. The specific geometrical fine-tuning problem in the Emergent models is the requirement of a particular choice of the initial radius a, if we are given the cosmological constant, or equivalently a specific unique choice of the primordial cosmological constant, if we are given the initial radius (see Eq. (66)). The model has been strongly criticised by some because of this fine-tuning. However we point out firstly that this criticism does not deny these models are valid physical models, but rather claims they are not likely to occur in reality because they are improbable. But the claimed singularity theorems [2, 16, 29] are not based on probabilities, and these models do indeed show that the geometric conditions of those theorems need not be satisfied. Secondly, we point out that the force of the fine-tuning argument is based in philosophy rather than physics. There is no scientifically based proof whatever that the unique physical universe has to be probable. There are two major philosophical approaches in cosmology to explaining the current state of the universe. The presently fashionable one, following comments by Dingle in the 1930's and the pioneering work of Misner in the 1960's, and continued notably in the inflationary universe scenario, is to try to show that the present state of the universe is highly probable - physical processes make it very likely to have occurred. However the earlier tradition is based on an opposite view: the idea that Nature prefers symmetry, and the universe would have been likely to have originated in a highly symmetric state. This view was central to Einstein's paper on the Einstein static model in 1917 , and was developed into the full-blown philosophy of the Cosmological Principle used by McCrea, Bondi, and others to justify the Robertson-Walker metrics and the Perfect Cosmological Principle used by Bondi, Gold, and Hoyle to justify the Steady State Universe. A more informal version was used by Einstein and de Sitter to justify universe models with $K = 0$, see the discussion of this often used 'simplicity principle' by Dicke and Peebles.

Now the point is that there is no scientific proof that the one or the other of these approaches is the correct approach to use in cosmology. The underlying problem is the uniqueness of the universe, and all the scientific and philosophical difficulties that entails. One cannot apply statistics or probability to a unique object, unless one uses Bayesian statistics where one knows all the priors - but in the case of the existence of the Universe, they are unknowable, indeed they don't even exist in literal sense (because one is talking about the creation of the universe itself). One might have a scientific basis for use of probability in cosmology if we were certain that there exists an ensemble of universes - a multiverse - but this is a complex and controversial proposal, and there is no evidence whatever that it is correct, indeed it will almost certainly never be scientifically provable that this is the case . Consequently the choice between these two fundamentally different approaches to cosmological origins is of necessity a philosophical one. The fact that a philosophical approach based on probabilities is the currently fashionable one is irrelevant - that does not prove it is the principle which actually underlies the existence of the real physical universe. One can equally cogently suppose that whatever meta-physical process underlies the existence of reality may, as so many have supposed in the past, prefer symmetry to generality. Indeed there is some physical support for this proposal through Penrose' argument that a highly symmetric start to the universe is essential in order that the thermodynamic arrow can function as it does. An argument of this kind must presumably underlie the widespread use of inflationary universe models with $K = 0$, because this highly special case (where the magnitude of the kinetic term in the Friedmann equation precisely balances the energy density of the universe to infinite accuracy) is of precisely the same degree of speciality as is involved in the initial conditions for the Emergent Universe. Choice of this special case in so many inflationary universe studies prevents their investigation of important dynamical effects in the early universe, and that restrictive choice is made on philosophical rather than physical grounds (one cannot observationally prove that $\Omega = 1$ to the infinite accuracy required to establish that in reality $K = 0$).. We are quite entitled to use this argument in the case of the emergent universe - to consider the implications if whatever process caused the universe to come into being preferred the high-symmetry state of the Einstein static universe to any less ordered situation. Indeed the Einstein static universe is the unique highest symmetry non-empty Robertson-Walker universe model, being invariant under a 7-dimensional group of isometries. Thus it is a highly preferred initial state geometrically, and it is interesting to see what models may result if one assumes that this was (at least asymptotically) the initial state of the universe.

In summary: inflationary cosmologies exist in which the horizon problem is solved before inflation begins, there is no singularity, and the quantum gravity regime is avoided and no exotic physics is involved. These Emergent Universe models can be constructed with simple potentials (illustrated schematically in Fig. 1), leading to past-infinite inflation with a bounded number of e-folds and redshift. Explicit and simple forms of the potential can be found that are consistent with observations for suitable choice of parameters in the potential. The Emergent models illustrate the potentially strong primordial effects of positive spatial curvature, leading to a very different early universe than the standard models, while producing a late universe that can be observationally indistinguishable from the standard case. The initial Einstein static state has certain appealing stability properties, and provides, because of its compactness, an interesting arena for investigating the Casimir and other effects. The Emergent models also open up interesting issues arising from infinite time extent (as opposed to the issues arising from infinite spatial extent in the standard models).

5.3 Problems With Infinity

When speaking of multiverses or ensembles of universes – possible or realised – the issue of infinity often crops up. Researchers often envision an *infinite* set of universes, in which all possibilities are realised. Can there really be an infinite set of really existing universes? We suggest that, on the basis of well-known philosophical arguments, the answer is No. Furthermore there are problems with any universe modesl that entail either a spatial or temporal infinity.

There is no conceptual problem with an infinite set – countable or uncountable – of *possible* or *conceivable* universes. However, as stressed by David Hilbert (1964), it can be argued that a *really existing* infinite set is not possible. As he points out, following many others, the existence of the actually infinite inevitably leads to well-recognised unresolvable contradictions in set theory, and thus in definitions and deductive foundations of mathematics itself (Hilbert, pp.

141-142). His basic position therefore is that "Just as operations with the infinitely small were replaced by operations with the finite which yielded exactly the same results . . ., so in general must deductive methods based on the infinite be replaced by finite procedures which yield exactly the same results." (p. 135) He concludes, "Our principle result is that the infinite is nowhere to be found in reality. It neither exists in nature nor provides a legitimate basis for rational thought . . . The role that remains for the infinite to play is solely that of an idea . . . which transcends all experience and which completes the concrete as a totality . . ." (Hilbert, p. 151). Indeed realised infinite sets are not constructible – there is no procedure one can in principal implement to complete such a set – they are simply incompletable. But, if that is the case, then "infinity" cannot be arrived at, or realised. On the contrary, the concept itself implies its inability to be realised! This is precisely why a realised past infinity in time is not considered possible from this standpoint – since it involves an infinite set of completed events or moments. There is no way of constructing such a realised set, or actualising it.

Thus, it is important to recognise that infinity is not an actual number we can ever specify or reach – it is simply the code-word for " it continues without end" . Whenever infinities emerge in physics – such as in the case of singularities – we can be reasonably sure, as is usually recognised, that there has been a breakdown in our models. An achieved infinity in any physical parameter (temperature, density, spatial curvature) is almost certainly *not* a possible outcome of any physical process – simply because it means traversing in actuality an interval of values which never ends. We assume space extends forever in Euclidean geometry and in many cosmological models, but we can never prove that any realised 3-space in the real universe continues in this way - it is an untestable concept, and the real spatial geometry of the universe is almost certainly not Euclidean. Thus Euclidean space is an abstraction that is probably not realised in physical practice. In the physical universe spatial infinities can be avoided by compact spatial sections, either resultant from positive spatial curvature or from choice of compact topologies in universes that have zero or negative spatial curvature, (for example FLRW flat and open universes can have finite rather than infinite spatial sections). Future infinite time is never realised: rather the situation is that whatever time we reach, there is always more time available. Much the same applies to claims of a past infinity of time: there may be unbounded time available in the past in principle, but in what sense can it be attained in practice? The arguments against an infinite past time are strong – it's simply not constructible in terms of events or instants of time, besides being conceptually indefinite.

6 Appendices

7 General tetrad formalism

A tetrad is a set of four orthogonal unit basis vector fields $\{ \mathbf{e}_a \}$, $a = 0, 1, 2, 3$, which can be written in terms of a local coordinate basis by means of the *tetrad components* ${e_a}^i(x^j)$:

$$\mathbf{e}_a = {e_a}^i(x^j)\,\frac{\partial}{\partial x^i} \;\Leftrightarrow\; \mathbf{e}_a(f) = {e_a}^i(x^j)\,\frac{\partial f}{\partial x^i}\,,\; {e_a}^i = \mathbf{e}_a(x^i)\,, \tag{70}$$

(the latter stating that the i-th component of the a-th tetrad vector is just the directional derivative of the i-th coordinate in the direction $\mathbf{e}_a$). This can be thought of as just a general change of vector basis, leading to a change of tensor components of the standard tensorial form: ${T^{ab}}_{cd} = {e^a}_i\,{e^b}_j\,{e_c}^k\,{e_d}^l\,{T^{ij}}_{kl}$ with obvious inverse, where the inverse components ${e^a}_i(x^j)$ (note the placing of the indices!) are defined by

$${e_a}^i\,{e^a}_j = {\delta^i}_j \;\Leftrightarrow\; {e_a}^i\,{e^b}_i = {\delta^b}_a\,. \tag{71}$$

However, it is a change from an integrable basis to a non-integrable one, so non-tensorial relations (specifically: the form of the metric and connection components) are a bit different than when coordinate bases are used. A change of one tetrad basis to another will also lead to transformations of the standard tensor form for all tensorial quantities: if $\mathbf{e}_a = {\Lambda_a}^{a'}(x^i)\,\mathbf{e}_{a'}$ is a change of tetrad basis with inverse $\mathbf{e}_{a'} = {\Lambda_{a'}}^a(x^i)\,\mathbf{e}_a$ (each of these matrices representing a Lorentz transformation), then ${T^{ab}}_{cd} = {\Lambda_{a'}}^a\,{\Lambda_{b'}}^b\,{\Lambda_c}^{c'}\,{\Lambda_d}^{d'}\,{T^{a'b'}}_{c'd'}$. Again the inverse is obvious.

The metric tensor components in the tetrad form are given by

$$g_{ab} = g_{ij}\,{e_a}^i\,{e_b}^j = \mathbf{e}_a\cdot\mathbf{e}_b = \eta_{ab}\,, \tag{72}$$

where $\eta_{ab} = \mathrm{diag}(\,-1,\,+1,\,+1,\,+1\,)$, showing that the basis vectors are unit vectors orthogonal to each other (because the components g_{ab} are just the scalar products of these vectors with each other). The inverse equation

$$g_{ij}(x^k) = \eta_{ab}\,{e^a}_i(x^k)\,{e^b}_j(x^k) \tag{73}$$

explicitly constructs the coordinate components of the metric from the (inverse) tetrad components ${e^a}_i(x^j)$. We can raise and lower tetrad indices by use of the metric $g_{ab} = \eta_{ab}$ and its inverse $g^{ab} = \eta^{ab}$.

The *commutation functions* related to the tetrad are the quantities ${\gamma^a}_{bc}(x^i)$ defined by the *commutators* of the basis vectors:

$$[\,\mathbf{e}_a,\,\mathbf{e}_b\,] = {\gamma^c}_{ab}(x^i)\,\mathbf{e}_c \;\Rightarrow\; {\gamma^a}_{bc}(x^i) = -\,{\gamma^a}_{cb}(x^i)\,. \tag{74}$$

where the commutator of any two vectors X, Y is $[X, Y] = XY - YX$. It follows (apply this relation to the coordinate x^i) that in terms of the tetrad components,

$${\gamma^a}_{bc}(x^i) = {e^a}_i\,(\,{e_b}^j\,\partial_j{e_c}^i - {e_c}^j\,\partial_j{e_b}^i\,) = -\,2\,{e_b}^i\,{e_c}^j\,\nabla_{[i}{e^a}_{j]}\,. \tag{75}$$

These quantities vanish iff the basis $\{\mathbf{e}_a\}$ is a coordinate basis: that is, there exist coordinates x^i such that $\mathbf{e}_a = \delta_a{}^i\,\partial/\partial x^i$, iff $[\,\mathbf{e}_a,\,\mathbf{e}_b\,] = 0 \Leftrightarrow \gamma^a{}_{bc} = 0$.

The *connection components* $\Gamma^a{}_{bc}$ for the tetrad ('Ricci rotation coefficients') are defined by the relations

$$\nabla_{\mathbf{e}_b}\mathbf{e}_a = \Gamma^c{}_{ab}\,\mathbf{e}_c \;\Leftrightarrow\; \Gamma^c{}_{ab} = e^c{}_i\,e_b{}^j\,\nabla_j e_a{}^i\ , \tag{76}$$

i.e. it is the c-component of the covariant derivative in the b -direction of the a-vector. It follows that all covariant derivatives can be written out in tetrad components in a way completely analogous to the usual tensor form, for example $\nabla_a T_{bc} = \mathbf{e}_a(T_{bc}) - \Gamma^d{}_{ba}\,T_{dc} - \Gamma^d{}_{ca}\,T_{bd}$, where for any function f, $\mathbf{e}_a(f) = e_a{}^i\,\partial f/\partial x^i$ is the derivative of f in the direction $\mathbf{e}_a$. In particular, because $\mathbf{e}_a(g_{bc}) = 0$ for $g_{ab} = \eta_{ab}$, applying this to the metric gives

$$\nabla_a g_{bc} = 0 \;\Leftrightarrow\; -\Gamma^d{}_{ba}\,g_{dc} - \Gamma^d{}_{ca}\,g_{bd} = 0 \;\Leftrightarrow\; \Gamma_{(ab)c} = 0\ , \tag{77}$$

— the rotation coefficients are skew in their first two indices, when we raise and lower the first indices only. We obtain from this and the assumption of vanishing torsion the tetrad relations that are the analogue of the usual Christoffel relations:

$$\gamma^a{}_{bc} = -\,(\Gamma^a{}_{bc} - \Gamma^a{}_{cb})\ ,\ \Gamma_{abc} = \ \frac{1}{2}\,(\,g_{ad}\,\gamma^d{}_{cb} - g_{bd}\,\gamma^d{}_{ca} + g_{cd}\,\gamma^d{}_{ab}\,)\ . \tag{78}$$

This shows that the rotation coefficients and the commutation functions are each just linear combinations of the other.

Any set of vectors whatever must satisfy the *Jacobi identities*:

$$[\,X,\,[\,Y,\,Z\,]\,] + [\,Y,\,[\,Z,\,X]\,] + [\,Z,\,[\,X,\,Y\,]\,] = 0\ ,$$

which follow from the definition of a commutator. Applying this to the basis vectors $\mathbf{e}_a$, $\mathbf{e}_b$ and $\mathbf{e}_c$ gives the identities

$$\mathbf{e}_{[a}(\gamma^d{}_{bc]}) + \gamma^e{}_{[ab}\,\gamma^d{}_{c]e} = 0\ , \tag{79}$$

which are the integrability conditions that the $\gamma^a{}_{bc}(x^i)$ are the commutation functions for the set of vectors $\mathbf{e}_a$.

If we apply the Ricci identities to the tetrad basis vectors $\mathbf{e}_a$, we obtain the Riemann curvature tensor components in the form

$$R^a{}_{bcd} = \mathbf{e}_c(\Gamma^a{}_{bd}) - \mathbf{e}_d(\Gamma^a{}_{bc}) + \Gamma^a{}_{ec}\,\Gamma^e{}_{bd} - \Gamma^a{}_{ed}\,\Gamma^e{}_{bc} - \Gamma^a{}_{be}\,\gamma^e{}_{cd}\ . \tag{80}$$

Contracting this on a and c, one obtains the EFE (for $\Lambda = 0$) in the form

$$R_{bd} = \mathbf{e}_a(\Gamma^a{}_{bd}) - \mathbf{e}_d(\Gamma^a{}_{ba}) + \Gamma^a{}_{ea}\,\Gamma^e{}_{bd} - \Gamma^a{}_{de}\,\Gamma^e{}_{ba} = T_{bd} - \frac{1}{2}\,T\,g_{bd}\ . \tag{81}$$

It is not immediately obvious that this is symmetric, but this follows because (79) implies $R_{a[bcd]} = 0 \Rightarrow R_{ab} = R_{(ab)}$.

8 1+3 covariant propagation and constraint equations

There are three sets of equations to be considered, resulting from the EFE (1) and its associated integrability conditions.

8.1 Ricci identities

The first set arise from the *Ricci identities* for the vector field u^a, i.e.

$$2\,\nabla_{[a}\nabla_{b]}u^c = R_{ab}{}^c{}_d\,u^d\ . \qquad (82)$$

We obtain from this three propagation equations and three constraint equations. The *propagation equations* are,

1. The *Raychaudhuri equation*

$$\dot{\Theta} -{}^3\,\tilde{\nabla}_a\dot{u}^a = -\,\frac{1}{3}\,\Theta^2 + (\dot{u}_a\dot{u}^a) - 2\,\sigma^2 + 2\,\omega^2 - \frac{1}{2}\,(\mu+3p) + \Lambda\ , \qquad (83)$$

which is the *basic equation of gravitational attraction*, showing the repulsive nature of a positive cosmological constant, leading to identification of $(\mu+3p)$ as the active gravitational mass density, and underlying the basic singularity theorem (see below).

2. The *vorticity propagation equation*

$$\dot{\omega}^{\langle a\rangle} - \frac{1}{2}\,\eta^{abc\,3}\tilde{\nabla}_b\dot{u}_c = -\,\frac{2}{3}\,\Theta\,\omega^a + \sigma^a{}_b\,\omega^b\ ; \qquad (84)$$

together with (92) below, showing how vorticity conservation follows if there is a perfect fluid with acceleration potential Φ [3] since then, $\eta^{abc\,3}\tilde{\nabla}_b u_c = \eta^{abc3}\tilde{\nabla}^3_b\tilde{\nabla}_c\Phi = 2\,\omega^a\,\dot{\Phi}$.

3. The *shear propagation equation*

$$\dot{\sigma}^{\langle ab\rangle} -{}^3\,\tilde{\nabla}^{\langle a}\dot{u}^{b\rangle} = -\,\frac{2}{3}\,\Theta\,\sigma^{ab} + \dot{u}^{\langle a}\dot{u}^{b\rangle} - \sigma^{\langle a}{}_c\,\sigma^{b\rangle c} - \omega^{\langle a}\,\omega^{b\rangle} - (E^{ab} - \frac{1}{2}\,\pi^{ab})\ , \qquad (85)$$

the anisotropic pressure source term π_{ab} vanishing for a perfect fluid; this shows how the tidal gravitational field E_{ab} directly induces shear (which then feeds into the Raychaudhuri and vorticity propagation equations, thereby changing the nature of the fluid flow).

The *constraint equations* are,

1. The (0α)-*equation*

$$0 = (C_1)^a ={}^3\,\tilde{\nabla}_b\sigma^{ab} - \frac{2}{3}{}^3\,\tilde{\nabla}^a\Theta + \eta^{abc}\,[\ {}^3\tilde{\nabla}_b\omega_c + 2\,\dot{u}_b\,\omega_c\ ] + q^a\ , \qquad (86)$$

showing how the momentum flux (zero for a perfect fluid) relates to the spatial inhomogeneity of the expansion;

2. The *vorticity divergence identity*

$$0 = (C_2) =^3 \tilde{\nabla}_a \omega^a - (\dot{u}_a \omega^a) \; ; \tag{87}$$

3. The H_{ab}*-equation*

$$0 = (C_3)^{ab} = H^{ab} + 2\, \dot{u}^{\langle a}\, \omega^{b\rangle} +^3 \tilde{\nabla}^{\langle a} \omega^{b\rangle} - (\mathrm{curl}\, \sigma)^{ab} \; , \tag{88}$$

characterising the magnetic Weyl tensor as being constructed from the 'distortion' of the vorticity and the 'curl' of the shear, $(\mathrm{curl}\sigma)^{ab} = \eta^{cd\langle a\, 3}\tilde{\nabla}_c \sigma^{b\rangle}{}_d$.

8.2 Twice-contracted Bianchi identities

The second set of equations arise from the *twice-contracted Bianchi identities* which, by the EFE (1), imply the conservation equations (2). Projecting parallel and orthogonal to u^a, we obtain the propagation equations

$$\dot{\mu} +^3 \tilde{\nabla}_a q^a = -\,\Theta\, (\mu + p) - 2\, (\dot{u}_a q^a) - (\sigma^a{}_b \pi^b{}_a) \tag{89}$$

and

$$\dot{q}^{\langle a\rangle} +^3 \tilde{\nabla}^a p +^3 \tilde{\nabla}_b \pi^{ab} = -\,\frac{4}{3}\, \Theta\, q^a - \sigma^a{}_b\, q^b - (\mu + p)\, \dot{u}^a - \dot{u}_b\, \pi^{ab} - \eta^{abc}\, \omega_b\, q_c \; , \tag{90}$$

respectively. For perfect fluids, characterised by Eq. (16), these reduce to

$$\dot{\mu} = -\,\Theta\, (\mu + p) \; , \tag{91}$$

the *energy conservation equation*, and one constraint equation

$$0 =^3 \tilde{\nabla}_a p + (\mu + p)\, \dot{u}_a \; , \tag{92}$$

the *momentum conservation equation*. This shows that $(\mu + p)$ is the inertial mass density, and also governs the conservation of energy. It is clear that if this quantity is zero (an effective cosmological constant) or negative, the behaviour of matter will be anomalous.

8.3 Other Bianchi identities

The third set of equations arise from the *Bianchi identities*

$$\nabla_{[a} R_{bc]de} = 0 \; . \tag{93}$$

Double contraction gives Eq. (2), already considered. On using the splitting of R_{abcd} into R_{ab} and C_{abcd}, the above $1 + 3$ splitting of those quantities, and the EFE, the once-contracted Bianchi identities give two further propagation equations and two further constraint equations, which are similar in form to the Maxwell field equations in an expanding universe.

The *propagation equations* are,

$$(\dot{E}^{\langle ab\rangle} + \frac{1}{2}\,\dot{\pi}^{\langle ab\rangle}) - (\,\mathrm{curl}H)^{ab} + \frac{1}{2}\,{}^{3}\tilde{\nabla}^{\langle a}q^{b\rangle} \;=\; -\,\frac{1}{2}\,(\mu+p)\,\sigma^{ab} - \Theta\,(E^{ab} + \frac{1}{6}\,\pi^{ab}) \quad (94)$$
$$+\;3\,\sigma^{\langle a}{}_{c}\,(E^{b\rangle c} - \frac{1}{6}\,\pi^{b\rangle c}) - \dot{u}^{\langle a}\,q^{b\rangle} + \;\eta^{cd\langle a}\,[\;2\,\dot{u}_{c}\,H^{b\rangle}{}_{d} + \omega_{c}\,(E^{b\rangle}{}_{d} + \frac{1}{2}\,\pi^{b\rangle}{}_{d})\;]\;,$$

the $\dot{E}$-*equation*, and

$$\dot{H}^{\langle ab\rangle} + (\mathrm{curl}\,E)^{ab} - \frac{1}{2}(\mathrm{curl}\,\pi)^{ab} = -\,\Theta\,H^{ab} + 3\,\sigma^{\langle a}{}_{c}\,H^{b\rangle c} + \frac{3}{2}\,\omega^{\langle a}\,q^{b\rangle} \qquad (95)$$
$$-\;\eta^{cd\langle a}\,[\;2\,\dot{u}_{c}\,E^{b\rangle}{}_{d} - \frac{1}{2}\,\sigma^{b\rangle}{}_{c}\,q_{d} - \omega_{c}\,H^{b\rangle}{}_{d}\;]\;,$$

the $\dot{H}$-*equation*, where we have defined the 'curls'

$$(\mathrm{curl}H)^{ab} \;=\; \eta^{cd\langle a}\,{}^{3}\tilde{\nabla}_{c}H^{b\rangle}{}_{d}\;, \qquad (96)$$
$$(\mathrm{curl}\,E)^{ab} \;=\; \eta^{cd\langle a}\,{}^{3}\tilde{\nabla}_{c}E^{b\rangle}{}_{d}\;, \qquad (97)$$
$$(\mathrm{curl}\,\pi)^{ab} \;=\; \eta^{cd\langle a}\,{}^{3}\tilde{\nabla}_{c}\pi^{b\rangle}{}_{d}\;. \qquad (98)$$

These equations show how gravitational radiation arises: taking the time derivative of the $\dot{E}$-equation gives a term of the form (curl H) ; commuting the derivatives and substituting from the $\dot{H}$ -equation eliminates H, and results in a term in $\ddot{E}$ and a term of the form (curl curl E), which together give the wave operator acting on E; similarly the time derivative of the $\dot{H}$-equation gives a wave equation for H.

The *constraint equations* are

$$0 = (C_4)^{a} \;=\; \mathrm{curl}_{b}(E^{ab} + \frac{1}{2}\,\pi^{ab}) - \frac{1}{3}{}^{3}\tilde{\nabla}^{a}\mu + \frac{1}{3}\,\Theta\,q^{a} - \frac{1}{2}\,\sigma^{a}{}_{b}\,q^{b} - 3\,\omega_{b}\,H^{ab}$$
$$-\;\eta^{abc}\,[\;\sigma_{bd}\,H^{d}{}_{c} - \frac{3}{2}\,\omega_{b}\,q_{c}\;]\;, \quad (99)$$

the (div E)-*equation* with source the spatial gradient of the energy density, which can be regarded as a vector analogue of the Newtonian Poisson equation, enabling tidal action at a distance, and

$$0 = (C_5)^{a} \;=\; {}^{3}\tilde{\nabla}_{b}H^{ab} + (\mu+p)\,\omega^{a} + 3\,\omega_{b}\,(E^{ab} - \frac{1}{6}\,\pi^{ab})$$
$$+\;\eta^{abc}\,[\;\frac{1}{2}{}^{3}\tilde{\nabla}_{b}q_{c} + \sigma_{bd}\,(E^{d}{}_{c} + \;\frac{1}{2}\,\pi^{d}{}_{c})\;]\;, \qquad (100)$$

the (div H)-*equation*, with source the fluid vorticity. These equations show respectively that scalar modes will result in a non-zero divergence of E_{ab} (and hence a non-zero E-field), and vector modes in a non-zero divergence of H_{ab} (and hence a non-zero H-field).

8.4 Maxwell field equations

Finally, we turn for completeness to the $1+3$ decomposition of the *Maxwell field equations*

$$\nabla_b F^{ab} = j_{\rm e}^a \ , \ \nabla_{[a}F_{bc]} = 0 \ . \tag{101}$$

The *propagation equations* can be written as

$$\dot{E}^{<a>} - \eta^{abc}\,{}^3\tilde{\nabla}_b H_c \ = \ -\,j_{\rm e}^{<a>} - \frac{2}{3}\,\Theta\,E^a + \sigma^a{}_b\,E^b + \eta^{abc}\,[\ \dot{u}_b\,H_c + \omega_b\,E_c\ ] \tag{102}$$

$$\dot{H}^{<a>} + \eta^{abc}\,{}^3\tilde{\nabla}_b E_c \ = \ -\,\frac{2}{3}\,\Theta\,H^a + \sigma^a{}_b\,H^b - \eta^{abc}\,[\ \dot{u}_b\,E_c - \omega_b\,H_c\ ]\ , \tag{103}$$

while the *constraint equations* assume the form

$$0 \ = \ (C_E) \ = \ {}^3\tilde{\nabla}_a E^a - 2\,(\omega_a H^a) - \rho_{\rm e}\ , \tag{104}$$

$$0 \ = \ (C_H) \ = \ {}^3\tilde{\nabla}_a H^a + 2\,(\omega_a E^a)\ , \tag{105}$$

where $\rho_{\rm e} = (-j_{{\rm e}\,a}u^a)$.

9 Solutions with symmetries

Symmetries of a space or a space-time (generically, 'space') are transformations of the space into itself that leave the metric tensor and all physical and geometrical properties invariant. We deal here only with continuous symmetries, characterised by a continuous group of transformations and associated vector fields

9.1 Killing vectors

A space or space-time *symmetry*, or *isometry*, is a transformation that drags the metric along a certain congruence of curves into itself. The generating vector field ξ_i of such curves is called a *Killing vector (field)* (or 'KV'), and obeys Killing's equations,

$$(L_\xi g)_{ij} = 0 \ \Leftrightarrow\ \nabla_{(i}\xi_{j)} = 0 \ \Leftrightarrow\ \nabla_i\xi_j = -\,\nabla_j\xi_i\ , \tag{106}$$

where L_X is the *Lie derivative.* The set of all KV's forms a Lie algebra with a basis $\{\,\xi_a\,\}$, $a = 1, 2, \ldots, r$, of dimension $r \leq \frac{1}{2}\,n\,(n-1)$. ξ^i_a denote the components with respect to a local coordinate basis, a, b, c label the KV basis, and i, j, k the coordinate components. Any KV can be written in terms of this basis, with *constant coefficients*. Hence: if we take the commutator $[\,\xi_a,\,\xi_b\,]$ of two of the basis KV's, this is also a KV, and so can be written in terms of its components relative to the KV basis, which will be constants. We can write the constants as $C^c{}_{ab}$, obtaining

$$[\,\xi_a,\,\xi_b\,] = C^c{}_{ab}\,\xi_c,\ C^a{}_{bc} = C^a{}_{[bc]}\ . \tag{107}$$

By the Jacobi identities for the basis vectors, these structure constants must satisfy

$$C^a{}_{e[b}C^e{}_{cd]} = 0 \ , \tag{108}$$

(which is just equation (79) specialized to the case of a set of vectors with constant commutation functions). These are the integrability conditions that must be satisfied in order that the Lie algebra exist in a consistent way. The transformations generated by the Lie algebra form a Lie group of the same dimension.

9.2 Groups of isometries

The isometries of a space of dimension n must be a group, as the identity is an isometry, the inverse of an isometry is an isometry, and the composition of two isometries is an isometry. Continuous isometries are generated by the Lie algebra of KV's. The group structure is determined locally by the Lie algebra, in turn characterised by the structure constants. The action of the group is characterised by the nature of its orbits in space; this is only partially determined by the group structure (indeed the same group can act as a space-time symmetry group in quite different ways).

9.2.1 Dimensionality of groups and orbits

Most spaces have no KV's, but special spaces (with symmetries) have some. The group action defines orbits in the space where it acts, and the dimensionality of these orbits determines the kind of symmetry that is present.

The *orbit* of a point p is the set of all points into which p can be moved by the action of the isometries of a space. Orbits are necessarily homogeneous (all physical quantities are the same at each point). An *invariant variety* is a set of points moved into itself by the group. This will be bigger than (or equal to) all orbits it contains. The orbits are necessarily invariant varieties; indeed they are sometimes called *minimum invariant varieties*, because they are the smallest subspaces that are always moved into themselves by all the isometries in the group. *Fixed points* of a group of isometries are those points which are left invariant by the isometries (thus the orbit of such a point is just the point itself). These are the points where all KV's vanish (however, the derivatives of the KV's there are non-zero; the KV's generate isotropies about these points). *General points* are those where the dimension of the space spanned by the KV's (that is, the dimension of the orbit through the point) takes the value it has almost everywhere; *special points* are those where it has a lower dimension (e.g. fixed points). Consequently, the dimension of the orbits through special points is lower than that of orbits through general points. The dimension of the orbit and isotropy group is the same at each point of an orbit, because of the equivalence of the group action at all points on each orbit.

The group is *transitive on a surface S* (of whatever dimension) if it can move any point of S into any other point of S. Orbits are the largest surfaces through each point on which the group is transitive; they are therefore sometimes referred to as *surfaces of transitivity.* We define their dimension as follows, and determine limits from the maximal possible initial data for KV's:

dim surface of transitivity $= s$, where in a space of dimension n , $s \leq n$.

At each point we can also consider the dimension of the isotropy group (the group of isometries leaving that point fixed), generated by all those KV's that vanish at that point:

dim of isotropy group $= q$, where $q \leq \frac{1}{2}\,n\,(n-1)$.

The *dimension r of the group of symmetries* of a space of dimension n is $r = s + q$ (translations plus rotations). From the above limits , $0 \leq r \leq n + \frac{1}{2}\,n\,(n-1) = \frac{1}{2}\,n\,(n+1)$ (the maximal number of translations and of rotations). This shows the Lie algebra of KV's is finite dimensional.

Maximal dimensions: If $r = \frac{1}{2}\,n\,(n+1)$, we have a space(-time) of constant curvature (maximal symmetry for a space of dimension n). In this case,

$$R_{ijkl} = K\,(\,g_{ik}\,g_{jl} - g_{il}\,g_{jk}\,)\ , \tag{109}$$

with K a constant; and K necessarily *is* a constant if this equation is true and $n \geq 3$. One cannot get $q = \frac{1}{2}\,n\,(n-1) - 1$ so $r \neq \frac{1}{2}\,n\,(n+1) - 1$.

A group is *simply transitive* if $r = s \ \Leftrightarrow \ q = 0$ (no redundancy: dimensionality of group of isometries is just sufficient to move each point in a surface of transitivity into each other point). There is no continuous isotropy group.

A group is *multiply transitive* if $r > s \ \Leftrightarrow \ q > 0$ (there is redundancy in that the dimension of the group of isometries is larger than is needed to move each point in an orbit into each other point). There exist non-trivial isotropies.

9.3 Classification of cosmological symmetries

We consider non-empty perfect fluid models, i.e. (16) holds with $(\mu + p) > 0$.

For a cosmological model, because space-time is 4-dimensional, the possibilities for the dimension of the surface of transitivity are $s = 0, 1, 2, 3, 4$. As to isotropy, we assume $(\mu + p) \neq 0$; then $q = 3$, 1, or 0 because u^a is invariant and so the isotropy group at each point has to be a sub-group of the rotations acting orthogonally to u^a (and there is no 2-dimensional subgroup of $O(3)$.) The dimension q of the isotropy group can vary over the space (but not over an orbit): it can be greater at special points (e.g. an axis centre of symmetry) where the dimension s of the orbit is less, but r (the dimension of the total symmetry group) must stay the same everywhere. Thus the possibilities for isotropy at a general point are:

a) Isotropic: $q = 3$, the Weyl tensor vanishes, kinematical quantities vanish except Θ. All observations (at every point) are isotropic. This is the FLRW family of geometries;

b) Local Rotational Symmetry ('LRS'): $q = 1$, the Weyl tensor is of algebraic Petrov type D, kinematical quantities are rotationally symmetric about a preferred spatial direction. All observations at every general point are rotationally symmetric about this direction. All metrics are known in the case of dust and a perfect fluid .

c) Anisotropic: $q = 0$; there are no rotational symmetries. Observations in each direction are different from observations in each other direction.

Putting this together with the possibilities for the dimensions of the surfaces of transitivity, we have the following possibilities (see Figure 1):

9.3.1 Space-time homogeneous models

These models with $s = 4$ are unchanging in space and time, hence μ is a constant, so by the energy conservation equation (91) they cannot expand: $\Theta = 0$. They cannot produce an almost isotropic redshift, and are not useful as models of the real universe. Nevertheless they are of some interest.

The *isotropic case* $q = 3$ ($\Rightarrow r = 7$) is the Einstein static universe, the non-expanding FLRW model (briefly mentioned above) that was the first relativistic cosmological model found. It is not a viable cosmology inter alia because it has no redshifts, but it laid the foundation for the discovery of the expanding FLRW models.

The *LRS case* $q = 1$ ($\Rightarrow r = 5$) is the Gödel stationary rotating universe, also with no redshifts. This model was important because of the new understanding it brought as to the nature of time in General Relativity. Inter alia, it is a model in which causality is violated (there exist closed timelike lines through each space-time point) and there exists no cosmic time function whatsoever.

The anisotropic models $q = 0$ ($\Rightarrow r = 4$) are all known, but are interesting only for the light they shed on Mach's principle.

9.3.2 Spatially homogeneous universes

These models with $s = 3$ are the major models of theoretical cosmology, because they express mathematically the idea of the 'cosmological principle': all points of space at the same time are equivalent to each other .

The *isotropic case* $q = 3$ ($\Rightarrow r = 6$) is the family of FLRW models, the standard models of cosmology.

The *LRS case* $q = 1$ ($\Rightarrow r = 4$) is the family of Kantowski–Sachs universes plus the LRS orthogonal and tilted Bianchi models. The simplest are the Kantowski–Sachs family, with comoving metric form

$$ds^2 = -\,dt^2 + A^2(t)\,dr^2 + B^2(t)\,(\,d\theta^2 + f^2(\theta)\,d\phi^2\,)\ , \tag{110}$$

where $f(\theta)$ is given by (27).

The *anisotropic case* $q = 0$ ($\Rightarrow r = 3$) is the family of Bianchi universes with a group of isometries G_3 acting simply transitively on spacelike surfaces. They can be orthogonal or tilted; the simplest class is the Bianchi Type I family. There is only *one* essential dynamical coordinate (the time t), and the EFE reduce to ordinary differential equations, because the inhomogeneous degrees of freedom have been 'frozen out'. They are thus quite special in geometrical terms; nevertheless, they form a rich set of models where one can study the exact dynamics of the full non-linear field equations. The solutions to the EFE will depend on the matter in the space-time. In the case of a fluid (with uniquely defined flow lines), we have two different kinds of models:

Orthogonal models, with the fluid flow lines orthogonal to the surfaces of homogeneity;

Tilted models, with the fluid flow lines not orthogonal to the surfaces of homogeneity; the components of the fluid peculiar velocity enter as further variables. Rotating models *must* be tilted, and are much more complex than non-rotating models.

9.3.3 Spatially inhomogeneous universes

These models have $s \leq 2$.

The *LRS cases* ($q = 1 \Rightarrow s = 2, r = 3$) are the spherically symmetric family with comoving metric form

$$ds^2 = -\,C^2(t,r)\,dt^2 + A^2(t,r)\,dr^2 + B^2(t,r)\,(\,d\theta^2 + f^2(\theta)\,d\phi^2\,)\;, \tag{111}$$

where $f(\theta)$ is given by (27). In the dust case, we can set $C(t,r) = 1$ and can integrate the EFE analytically; for $k = +1$, these are the Lemaître–Tolman–Bondi ('LTB') spherically symmetric models [1]. They may have a centre of symmetry (a timelike worldline), and can even allow two such centres, but they cannot be isotropic about a general point (because isotropy everywhere implies spatial homogeneity; see the discussion of FLRW models). The *anisotropic cases* ($q = 0 \Rightarrow s \leq 2, r \leq 2$) include solutions admitting an Abelian or non-Abelian group of isometries G_2, and spatially self-similar models.

Solutions with no symmetries at all have $r = 0 \Rightarrow s = 0, q = 0$. The real universe, of course, belongs to this class; all the others are intended as approximations to this unique universe. Remarkably, we know some exact solutions without symmetries, specifically (a) the Szekeres quasi-spherical models, that are in a sense non-linear FLRW perturbations, with comoving metric form

$$ds^2 = -\,dt^2 + e^{2A}\,dx^2 + e^{2B}(dy^2 + dz^2)\;,\; A = A(t,x,y,z)\;,\; B = B(t,x,y,z)\;, \tag{112}$$

(b) Stephani's conformally flat models, and (c) Oleson's type N solutions (for a discussion of these and all the other inhomogeneous models, see Krasi ński [20] and Kramer et al [21]).

9.3.4 Swiss-Cheese models

Finally, an interesting family of inhomogeneous models is the Swiss-Cheese family of models, obtained by repeatedly cutting out a spherical region from a FLRW model and filling it in with another spherical model: Schwarzschild or LTB, for example. This requires:

(i) locating the 3-dimensional timelike *junction surfaces* $\Sigma_\pm$ in each of the two models;

(ii) defining a proposed *identification* Φ between Σ_+ and Σ_-;

(iii) determining the *junction conditions* that (a) the 3-dimensional metrics of Σ_+ and Σ_- (the first fundamental forms of these surfaces) be isometric under this identification, so that there be no discontinuity when we glue them together — we arrive at the same metric from both sides — and (b) the second fundamental forms of these surfaces must also be isometric when we make this identification, so that they too are continuous in the resultant space-time — equivalently, there is no discontinuity in the direction of the spacelike unit normal vector as we cross the junction surface Σ (this is the condition that there be no surface layer on Σ once we make the join).

(iv) Having determined that these junction conditions can be satisfied for some particular identification of points, one can then proceed to identify these corresponding points in the two surfaces Σ_+ and Σ_-, thus gluing an interior Schwarzschild part to an exterior FLRW part, for example.

(v) One can continue in this way, obtaining a family of holes of different sizes in a FLRW model with different interior fillings, with further FLRW model segments fitted into the interiors of some of these regions, obtaining a Swiss-Cheese model. One can even obtain a hierarchically structured family of spherically symmetric vacuum and non-vacuum regions in this way.

These models were originally developed by Einstein and Strauss to examine the effect of the expansion of the universe on the solar system. Subsequent uses of these models have included examining Oppenheimer–Snyder collapse in an expanding universe, examining gravitational lensing effects on area distances, investigating CBR anisotropies, modelling voids in large-scale structure, perhaps using surface-layers, modelling the universe as a patchwork of domains of different curvature $k = 0, \pm 1$.

References

[1] H Bondi, *Mon. Not. Roy. Astr. Soc.* **107**, 410 (1947).

[2] A. Borde and A. Vilenkin, Phys. Rev. Lett. **72**, 3305 (1994); Phys. Rev. D**56**, 717 (1997); A. Borde, A.H. Guth, and A. Vilenkin, gr-qc/0110012.

[3] J Ehlers, *Akad. Wiss. Lit. Mainz, Abhandl. Math.-Nat. Kl.* **11**, 793 (1961). Translation: J Ehlers, *Gen. Rel. Grav.* **25**, 1225 (1993).

[4] J Ehlers, P Geren and R K Sachs, *J. Math. Phys.* **9**, 1344 (1968).

[5] . J Ehlers and W Rindler, *Mon. Not. Roy. Astr. Soc.* **238**, 503 (1989).

[6] ell71 G F R Ellis, in *General Relativity and Cosmology, Proceedings of the XLVII Enrico Fermi Summer School*, Ed. R K Sachs (Academic Press, New York, 1971).

[7] G F R Ellis and M Bruni, *Phys. Rev.* D **40** , 1804 (1989).

[8] G F R Ellis, U Kirchner and W Stoeger (2003) To appear MNRAS astro-ph/0305292

[9] Ellis G F R and van Elst H 1999 Cosmological Models (Cargese Lectures 1998). In *Theoretical and Observational Cosmology* ed M Lachieze-Ray (Kluwer, Nato Series C: Mathematical and Physical Sciences, Vol 541, 1999) 1-116 (*Archive* gr-qc/9812046).

[10] George Ellis, Roy Maartens Class. Quant. Grav. 21 (2003) 223] gr-qc/0211082

[11]

[12] George F.R. Ellis, Jeff Murugan, Christos G. Tsagas, Class. Quant. Grav. 21 (2003) 233 gr-qc/0307112

[13] G F R Ellis and W R Stoeger, *Class. Quantum Grav.* **5**, 207 (1988).

[14] G F R Ellis and H van Elst, in *On Einstein's Path — Essays in Honor of Engelbert Schücking*, Ed. A Harvey, (Springer, New York, 1999). *gr-qc/9709060*.

[15] Fischer A E and Marsden J E 1979 The initial value problem and the dynamical formulation of general relativity in *General Relativity: An Einstein Centenary Survey* ed S W Hawking and W Israel (Cambridge: Cambridge University Press).

[16] A.H. Guth, astro-ph/0101507.

[17] Hewitt C G, Horwood J T and Wainwright J 2003 Asymptotic dynamics of the exceptional Bianchi cosmologies *Class. Quant. Grav.* **20** 1743-1756 (*Archive* gr-qc/0211071).

[18] Hilbert D 1964 On the Infinite in *Philosophy of Mathematics* ed P Benacerraf and H Putnam (Englewood Cliff, N. J.: Prentice Hall) pp 134-151.

[19] Horwood J T, Hancock M J, The D, and Wainwright J 2003 Late-time asymptotic dynamics of Bianchi VIII cosmologies *Class. Quant. Grav.* **20** 1757-1778 (*Archive* gr-qc/0210031).

[20] A Krasiński, *Physics in an Inhomogeneous Universe* (Cambridge University Press, Cambridge, 1996).

[21] D Kramer, H Stephani, M A H MacCallum and E Herlt, *Exact Solutions of Einstein's Field Equations* (Cambridge University Press, Cambridge, 1980).

[22] Lim W C, van Elst H, Uggla C, and Wainwright J 2003 Asymptotic isotropization in inhomogeneous cosmology (*Archive* gr-qc/0306118).

[23] M S Madsen and G F R Ellis, *Mon. Not. Roy. Astr. Soc.* **234**, 67 (1988).

[24] R Penrose, in *Relativity Groups and Topology*, Eds. C M DeWitt and B S DeWitt (Gordon and Breach, New York, 1963).

[25] Refsdal and Stabell

[26] W Rindler, *Mon. Not. Roy. Astr. Soc.* **116**, 662 (1956).

[27] F J Tipler, C J S Clarke and G F R Ellis, in *General Relativity and Gravitation: One Hundred Years after the Birth of Albert Einstein, Vol. 2*, Ed. A Held (Plenum Press, New York, 1980).

[28] Uggla C, van Elst H, Wainwright J and Ellis G F R 2003 The past attractor in inhomogeneous cosmology *Archive* gr-qc/0304002.

[29] A. Vilenkin, gr-qc/0204061.

[30] J Wainwright and G F R Ellis (Eds.), *Dynamical Systems in Cosmology* (Cambridge University Press, Cambridge, 1997).