Approaches to Ensembles of Universes

Ulrich Kirchner

Thesis Presented for the Degree of

Doctor of Philosophy

in the Department of Mathematics and Applied Mathematics
University of Cape Town

September 2003

Abstract

This thesis consists of three parts — each of them focusing on different aspects relating to ensembles of universes and causally disconnected regions.

In the first part I investigate possible measures over the space of FLRW models. Jaynes' principle is used to find new probability measures for ensembles of open and closed dust-FLRW models and the results generalized to a γ-equation of state. For big-bang solutions the measures are non-integrable at $\Omega_0 = 1$ and $\Omega_0 = \infty$. Together with additional restrictions on the possible values of the cosmological constant Λ these measures can solve the Flatness Problem.

Often the term "multiverse" is misleadingly used for multi-domain universes, i.e., universes which consist of many different domains separated by some period of inflation. It is often argued that the transition region between such domains should be of negligible spatial extend.

In the second part I examine the behaviour of such transition regions for spherically symmetric space-times. By using a new approach to evaluate the matching conditions I re-discover evolution equations for the junction surface and its "surface-matter" content. I give restrictions on the possible values for the surface-energy density and show that the time component of the Lanczos equation is always identically satisfied. Furthermore, I show that generic models do not allow timelike junctions without surface-layer and the behaviour for small surface-energies is discussed.

Special attention is paid to junctions between FLRW models. It is shown that it is geometrically impossible to join two distinct FLRW models (with γ-equation of state) along a junction surface which is comoving on both sides. I evaluate the restrictions resulting from a non-negativity condition on the surface-energy density. For the non-comoving case I give several numerical examples.

In the last part of this thesis I discuss philosophical, physical, and probabilistic issues related to the concept of a multiverse - an ensemble of universes. The difference between ensembles of really existing universes and ensembles of possible universes is emphasized. It is discussed why there is no *unique* multiverse in which "everything that can happen happens". The role of probability measures and distribution functions is emphasized and related problems with infinities discussed. The matter is illustrated by explicit examples of ensembles of FLRW models.

Recent CMB-data marginally indicates a closed geometry for our universe, which disagrees with predictions from Chaotic Inflation. In the appendix it is shown how this would limit the number of possible e-foldings that one could get out of inflation.

Contents

Prologue

The idea that our universe could be just one of an ensemble of universes is intriguing — many universes could be realized and we naturally live in one that admits life. This places anthropic reasoning on a new basis, where the 'observation of our existence' is just a selection criteria for the set of possible 'home' universes in the ensemble, not implying that the other universes are not realized.

Notwithstanding the recent interest in this kind of reasoning, it has to be admitted that its basis is very vague, and on closer inspection almost arbitrary: What is meant by a 'universe'? Which universes are in the ensemble? How many universes are there? What are the probability measures and distributions describing the ensemble? What is meant with 'existence' if the universes do not belong to the same space-time?

In this thesis I will deal with some of these questions — ranging from an investigation of possible minimum-information measures for ensembles of FLRW models, over the more practical issue on how one can 'glue' different FLRW sections along a spherical symmetric (timelike) junction surface (an often used model for multi-domain universes, in which different regions expand at different rates), to a discussion of the concept of a 'multiverse' in a cosmological context.

I start in chapter 1 with an investigation of minimum information measures for FLRW models. The present state of knowledge about how our universe was created and what determined its initial conditions is probably (and probably will be for a long time) best described by 'minimum knowledge' — for us there is only one observable universe and hence we cannot measure distribution functions or even identify which parameters could really vary within an ensemble, whether realized or not. Assuming that all (dust or γ-equation of state) FLRW models could be realized leaves the question which probability measures and distributions we should assume for the necessary parameters given the fact that we have minimum information.

Our only (assumed) knowledge about the ensemble is its mathematical structure, and in particular the allowed parameter ranges. In contrast to previous attempts I strictly use the constants of motion as parameters, such that a clear distinction is made between the model in the space of models and its state at a particular time in the state space.

Based on this I suggest new probability measures over the space of FLRW models. It should be noted that in the Bayesian interpretation of probability these measures are also meaningful if there is just one universe created, i.e., the ensemble is just an ensemble of possible models, but not necessarily realized ones. It turns out that these measures have a non-integrable divergency around $\Omega_0 = 1$, indicating that our almost-flat universe might not be so unlikely after all. This indeed would solve the Flatness Problem without any need for inflation.

Multi-domain universes like in Chaotic Inflation are an example of ensembles of universes in the wider sense (since the regions are causally connected in the past). The junction (bubble wall) between different domains is usually assumed to be 'thin' and (for simplicity) spherically symmetric. The matter content of the junction is modelled by a δ-function contribution to the energy-momentum tensor, which is related to the jump in extrinsic curvature across the junction surface.

In chapter 2 I investigate such spherically symmetric junction surfaces. My approach differs from previous investigations in the choice of coordinate system and variables. By re-scaling the time and radial coordinate we make all coordinates continuous at the junction and absorb the junction motion into the metric components.

This approach leads to a particularly pleasing form of the junction conditions. I use the developed formalism to investigate several aspects of timelike junction surfaces. In particular the behaviour for small surface energies is studied, conditions on the possible values of the surface energy density are evaluated, and it is studied under which conditions the surface energy density could vanish. Furthermore, it is shown that one of the junction conditions, the time component of the Lanczos equation, is identically satisfied for all spherically symmetric junctions. This appears to be an important new result, which supplements numerous previous studies of junctions between spherically symmetric space-time sections.

An interesting result is that in many cases a timelike junction might be initially possible, but the junction surface cannot be extended arbitrarily as a *timelike* surface. An illustrative example is the junction between a closed recollapsing inside region and an open inflating outside region. Due to the limited 'life time' of the inside region the junction surface has to end as well. Cases like these are given as numerical examples.

The idea of a multiverse – an ensemble of universes – has received increasing attention in cosmology, both as the outcome of the originating process that generated our own universe, and as an explanation for why our universe appears to be fine-tuned for life and consciousness. In chapter 3 I carefully consider how multiverses should be defined, stressing the distinction between the collection of all possible universes, and ensembles of really existing universes that are essential for an anthropic argument. It is shown that such realized multiverses are by no means unique. A proper measure on the space of all really existing universes or universe domains is needed, so that probabilities can be calculated, and major problems arise in terms of realized infinities. As an illustration I examine these issues in the case of the set of Friedmann-Lemaître-Robertson-Walker (FLRW) universes. Then I briefly summarise scenarios like chaotic inflation, which suggest how ensembles of universe domains may be generated, and point out that the regularities which must underlie any systematic description of truly disjoint multiverses must imply some kind of common generating mechanism. Finally, I discuss the issue of testability, which underlies the question of whether multiverse proposals are really scientific propositions.

One of the most accepted multi-domain theories is the Chaotic Inflation scenario by Andrei Linde, which proposes the existence of infinitely many flat or open domains. It is noteworthy that such a model is not available for closed geometries. With the recent release of the WMAP-data the quality of available microwave-background-radiation measurements improved drastically. While measurements are still in agreement with a flat spatial geometry, they seem to favour a closed spatial geometry. In appendix A I discuss how a closed spatial geometry (as indicated by the WMAP data) limits the amount of inflation that the universe can have undergone.

Publications

Parts of this thesis have been published, others are in the process of submission. The material presented in chapter 1 was published in [1].

The work on chapter 3 was done in collaboration with William R Stoeger and George F R Ellis.

Appendix A contains work which was done in collaboration with Jean-Phillipe Uzan (Paris) and George F R Ellis and was published in [2]. The results presented in chapters 2 and 3 are submitted for publication.

[1] 2003 *Class. Quantum Grav.* **20** 1199-1213

[2] 2003 *Month. Not. Roy. Astr. Soc.* accepted for publication

Chapter 1

Probability measures on the space of FLRW-models

1.1 Introduction

Recent measurements of the cosmic-microwave background radiation (CMB) [1, 2, 3] show that our universe is spatially almost isotropic and homogeneous, *i.e.*, it is well approximated by a Friedmann-Lemaître-Robertson-Walker (FLRW) model. Furthermore, the total density parameter Ω is found to be very close to one, while there is increasing evidence for a non-vanishing cosmological constant.

It is usually argued that a present density parameter so close to unity should have been extremely unlikely, and thus it is called the *Flatness Problem*.

There have been many attempts to resolve this fine-tuning problem. Firstly, there are inflationary models [4], which postulate the existence of one or more scalar fields which yield an epoch of exponential expansion. This indeed drives the density parameter exponentially close to one and might account for density perturbations through quantum fluctuations. Nevertheless, all attempts to identify the driving scalar fields have failed.

Secondly, there are anthropic arguments [5] which claim that life could only evolve because the density parameter Ω was very close to one. This implies that we have been either very lucky that there exists a universe with Ω close to one, or that there is a whole ensemble of universes [6] and we naturally exist in one which provides the right conditions for life.

A third intriguing possibility is that there is actually no Flatness Problem. One should note that it is the underlying assumption of a flat (non-divergent) probability distribution around $\Omega = 1$ which creates the Flatness Problem. A closer look at the evolution equations for the density parameter Ω reveals that this assumption might be unreasonable. Going backwards in time towards the big-bang [20] we find that Ω must converge towards one (if it is initially non-zero and we have an *exact* FLRW model) as long as the classical equations remain valid and hence all classical models have to start with this value. This is clearly in contradiction with a non-divergent probability distribution.

Here we will investigate the latter proposal by deriving a set of probability measures for different classes of cosmological dust-FLRW models. We start with a summary of Bayesian inference followed by a review of previous attempts to find such a measure and point out problems arising. In section 1.4.1 we identify a possible class of models and their parametrizations. Using the so called Jaynes' principle we find in section 1.4.2 a set of possible probability measures over the parameters. In section 1.4.3 we transform these measures into measures over the density parameters, which are cosmological observables, and in section 1.4.4 we extend the results to a γ-law equation of state. How additional information (like anthropic constraints) modify these measures accordig to Bayes' theorem is outlined in section 1.5. This is followed by the conclusion in section 1.6.

1.2 Bayesian Inference and Jaynes' Principle

We adopt the Bayesian interpretation of probability as the degree of consistent belief, which allows us to assign probabilities to unique events [7]. Probabilities always depend on given information. We denote by $P(A|C)$ the probability that A is true given the information that C is true.

Bayes' theorem

$$P(A|BC) = \frac{P(B|AC)}{P(B|C)} P(A|C) \tag{1.1}$$

tells us how the probability changes when more information (the truth of proposition B) is acquired. One calls $P(A|C)$ the prior probability and $P(A|BC)$ the posterior probability. It should be noted that the rules for calculating with probabilities (including (1.1)) follow alone from consistency requirements [8].

The prior probabilities are assigned using the maximum-entropy principle [9, 10], which appears to be the only consistent variational principle for this task [11, 12, 13, 14]. This ensures that given the same information C different individuals assign the same prior probability $P(A|C)$. For a discrete distribution the probabilities are assigned by maximising the information entropy

$$-\sum_i p_i \log p_i \tag{1.2}$$

subject to given constraints.

This can be generalized to continuous probability densities [9]. The information entropy becomes

$$S = -\int \mathrm{d}x P(x) \log\Big(\frac{P(x)}{\mu(x)}\Big), \tag{1.3}$$

where $\mu(x)$ is a measure, which is necessary to make (1.3) invariant under a change of integration variables.

If we have no information about the system, *i.e.*, no constraints for the probability density, then the information entropy is maximized for $P(x) = \mu(x)$, which corresponds to a constant probability density over some 'natural' background space.

In order to identify the measure $\mu(x)$ Jaynes' suggested a 'transformation group method' [15], which is known today as Jaynes' principle. One identifies symmetry transformations $x'(x)$ and demands that $\mu(x)\mathrm{d}x = \mu(x')\mathrm{d}x'$.

A location parameter z, *i.e.*, a parameter such that in the state of 'complete ignorance' there is no prefered value (see pp 379 in [16]), can take all real values and the probability measure should be invariant under the transformation

$z \to z' = z + \alpha$. Jaynes' principle demands $\mu(z)\mathrm{d}z = \mu(z')\mathrm{d}z'$ and hence $\mu(z) = \mu(z+\alpha)$ for all $\alpha \in \mathbf{R}$. This has the unique solution $\mu(z) = \text{constant}$.

On the other hand, 'complete ignorance' of scale is expressed by the invariance of the probability measure under scaling. If u is a positive parameter which serves as a scale parameter (*e.g.*, expressing the freedom of choice of units, not a scale factor in the cosmological sense) then scale invariance demands that $\mu(u)\mathrm{d}u = \mu(\lambda u)\lambda\mathrm{d}u$ and hence

$$\mu(u) \propto \frac{1}{u}. \tag{1.4}$$

We note that above considerations do not apply if $u = 0$ is allowed, because $\lambda > 0$. Nevertheless, given a parameter v, which can take values in $[0, \infty)$ we can define a new parameter $w \in \mathbf{R}$ such that $v = w^2$. Applying Jaynes' principle yields

$$\mu \propto \mathrm{d}w \propto \frac{\mathrm{d}v}{\sqrt{v}}. \tag{1.5}$$

Similarly for a parameter s which takes values in a compact interval $[a, b]$ we can define $s = a + \frac{b-a}{2}(1 + \sin(t))$. Now t can take all real values and we conclude

$$\mu \propto \mathrm{d}t \propto \frac{\mathrm{d}s}{\sqrt{s-a}\sqrt{b-s}}. \tag{1.6}$$

If the allowed parameter range is non-compact then the measure will in general be non-normalisable. Nevertheless, these improper measures can be used as priors in Bayes theorem (1.1) because the normalization factor cancels.

Note that the derivations of the last two measures depend on a somehow arbitrary choice of new parametrisation. These transformations are not unique, and hence one could find many different measures. Nevertheless, in the state of minimum information we don't know what the natural parametrization is for the possibility space and different measures correspond to different guesses. Surprisingly Jaynes' principle is "relatively invariant" under simple parametrization changes. For example, introducing a new parametrization for a positive quantity G by $G = m^n$ for positive m, or by $G = \exp(\lambda)$ for real λ, will give the same measure.

1.3 Measures on the space of models

1.3.1 The measure of Gibbons, Hawking, and Stewart

In [17] Gibbons, Hawking, and Stewart suggested a measure over an ensemble of FLRW universes with scalar field, which they called the 'canonical measure'. It is based on the assumption that the probability measure should be proportional to the phase-space volume if energy and time are kept constant. This measure is conserved along flow-lines and hence assigns the same probability at all times for the same model. However, the measure applies only to ensembles of systems with an at least four-dimensional phase-space.

Nevertheless, the simplest FLRW models have only one degree of freedom — the scale factor a — and hence it is not possible to use this measure in these cases.

In [17, 18] this measure was discussed for FLRW models with a massive scalar field which satisfies the Klein-Gordon equation. It was shown in [18, 19] that the measure diverges for inflating and non-inflating models and hence leaves the ratio undefined. Furthermore, the conservation of the measure along the flow-lines depends on the validity of the Klein-Gordon equation and hence it does not apply to a perfect fluid, *i.e.*, the universe at recent times.

1.3.2 The measure of Evrard and Coles

Evrard and Coles suggested a measure [20] based on Jaynes' principle. They use the scale factor a and the constant of motion $\chi \stackrel{\text{def}}{=} \frac{\kappa\rho a^3}{3}$ to parametrize dust-FLRW models for fixed values of the cosmological density parameter Ω_Λ [1]. It is argued that because χ and a take only positive values the measure takes the form

$$\mu \propto \frac{\mathrm{d}a\mathrm{d}\chi}{a\chi}. \tag{1.7}$$

In [20] it was claimed that this measure can be expressed in terms of the observable Hubble parameter H and density parameter Ω (assuming $\Omega_\Lambda = 0$) as

$$\mu \propto \frac{\mathrm{d}H\mathrm{d}\Omega}{H\Omega|1-\Omega|}. \tag{1.8}$$

This measure is non-integrable around $\Omega = 0$ and $\Omega = 1$ — predicting almost all universes to be created around these points. This in fact could solve the flatness problem (though there is an ambiguity about which divergence is 'stronger'), but there are problems with this measure and its derivation.

The two parameters a and χ do not only specify the model itself, but also the time during the evolution of the model, *i.e.*, they completely define a point in state-space, rather than a trajectory. Hence (1.7) is a measure over an ensemble of possible states of all dust-FLRW models and not over the models itself. Nevertheless, one would expect that a probability is associated with the likelihood that a certain model is created and hence we should rather look for a measure over models than a measure over the state-space of all models.

The probability that our universe has a scale factor a_0 and matter constant χ_0 in given intervals ,*i.e.*, $a_0 \in [a_1, a_2], \chi_0 \in [\chi_1, \chi_2]$, is (up to a normalization constant) given by

$$\int_{a_1}^{a_2} \int_{\chi_1}^{\chi_2} \mu(a,\chi) \propto \ln\left(\frac{a_2}{a_1}\right) \ln\left(\frac{\chi_2}{\chi_1}\right). \tag{1.9}$$

While χ is constant during the evolution of the FLRW model, the scale factor a changes with time. The first factor is only time-independent for $a_1 = \lambda a_2$, where λ is a time-independent real constant; this will be true for the scale-free $k = 0$ models. Nevertheless, for open and closed models in the Friedmann equation the curvature constant k was rescaled to ± 1, and hence there remains no freedom to rescale the scale factor a and first factor of (1.9) will be time-dependent when we consider this set of models at different times. Hence the measure assigns different probabilities to the same set of universe models at different

[1] In their derivation they assume a vanishing cosmological constant, but the same result holds if $\Omega_\Lambda \neq 0$ is taken as a known quantity.

times. Nevertheless, one would expect that the *a priori* probability expresses the likelihood of initial-values, and therefore should be time-independent.

For certain classes of FLRW models the scale factor *cannot* take all positive values, and hence one should not expect that the measure is constant in the $\ln a$-parametrization. Furthermore, χ can vanish for empty models, but $\chi = 0$ is not covered in the $\ln \chi$ parametrization.

It should be noted that (1.7) is not a measure over all possible dust-FLRW models since it assumes a fixed value for the cosmological density parameter Ω_Λ, indeed it is proposed specifically for the case $\Omega_\Lambda = 0$. It would be desirable to have a measure over all parameters which are necessary to uniquely identify the model including Ω_Λ.

Furthermore, it should be mentioned that the parameters a, χ are not in a one-to-one correspondence to H, Ω because they do not uniquely define the geometry of the universe (see figure 1.1). In effect one needs separate probability distributions in the three cases of open, closed, and flat space-time geometry.

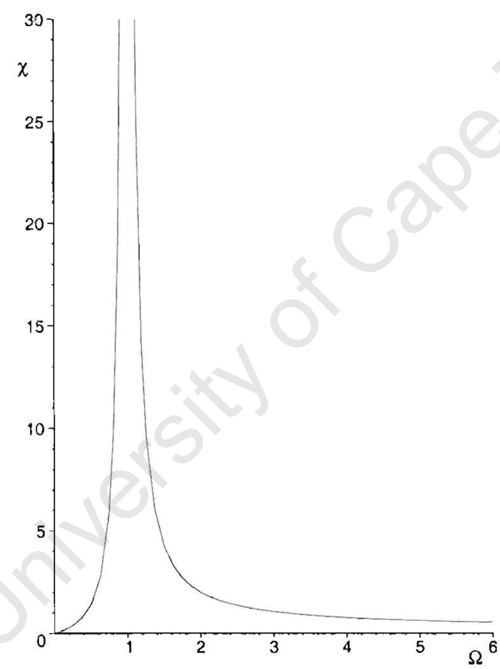


Figure 1.1: Relation between χ and Ω for a constant Hubble parameter H and vanishing cosmological constant Λ.

1.4 A new measure

1.4.1 Selecting a parametrization

Open and closed dust-FLRW models have three constants of motion: the cosmological constant Λ, the matter constant $\chi = \kappa\rho a^3/3$, and the curvature constant k. The latter one can be normalized to $+1$ for closed and -1 for open models.

For open models the Friedmann equation becomes with $y = \frac{1}{a}$

$$g(y) = H^2 = \chi y^3 + y^2 + \frac{\Lambda}{3}. \tag{1.10}$$

For any values of χ and Λ there exists a unique (continuously connected) range of scale factors where $H^2 > 0$, *i.e.*, a uniquely identified model. For $\Lambda < 0$ the models expand to a maximum scale factor and re-collapse, otherwise they expand forever and the Hubble parameter H approaches asymptotically the value $\sqrt{\frac{\Lambda}{3}}$.

While χ and Λ uniquely characterize open models, this is not true for closed models. In this case the Friedmann equation becomes with $y \stackrel{\text{def}}{=} \frac{1}{a}$

$$f(y) = H^2 = \chi y^3 - y^2 + \frac{\Lambda}{3}. \tag{1.11}$$

This function attains a maximum at $(0, \frac{\Lambda}{3})$ (which corresponds to $a = \infty$) and a minimum at $(\frac{2}{3\chi}, \frac{\Lambda}{3} - \frac{4}{27\chi^2})$. The distinct possible cases which are identified by Λ and χ are shown in figure 1.2.

I: The universe starts with a big-bang ($a = 0, y = \infty$), expands until $H = 0$, and then re-collapses (leading to a big-crunch). All values of Λ and χ which yield a negative minimum for $f(y)$, *i.e.*, $9\Lambda\chi^2 - 4 < 0$, allow this solution.

II: Like in I the universe starts with a big-bang, but now the Hubble parameter H never reaches zero. These *closed* models expand forever and are uniquely identified by values of Λ and χ which do yield a positive minimum for $f(y)$, *i.e.*, $9\Lambda\chi^2 - 4 > 0$.

III: The universe starts flat ($a = \infty, y = 0$), but with a negative Hubble parameter $H < 0$. The universe then contracts to some minimum scale factor where $H = 0$, and then re-expands. This so called *bounce* solution is distinct in that the universe has to start with some non-zero (or infinite) scale factor, *i.e.*, this solution is not continuously connected to a big-bang. Bounce solutions exist whenever case **I** exists with a non-vanishing positive cosmological constant $\Lambda > 0$.

IV: If the minimum touches the $H^2 = f(y) = 0$ axis, *i.e.*, for $\Lambda = \frac{4}{27\chi^2}$, we find three distinct solutions. For $a = 3\chi/2$ (*i.e.*, where $H = 0$) we find the static Einstein-de Sitter solution. For $a > 3\chi/2$ the bounce model 'degenerates' into two solutions — an initially flat universe with negative Hubble parameter approaching asymptotically the Einstein de-Sitter solution and its time reversed model. Similarly for $a < \chi/2$ the type-I solution degenerates into an forever expanding model, which approaches asymptotically the Einstein-de Sitter model, and its time-reversed model.

We conclude that values of Λ and χ which allow bounce solutions represent two distinct closed dust-FLRW models of type **I** and **III**, respectively.

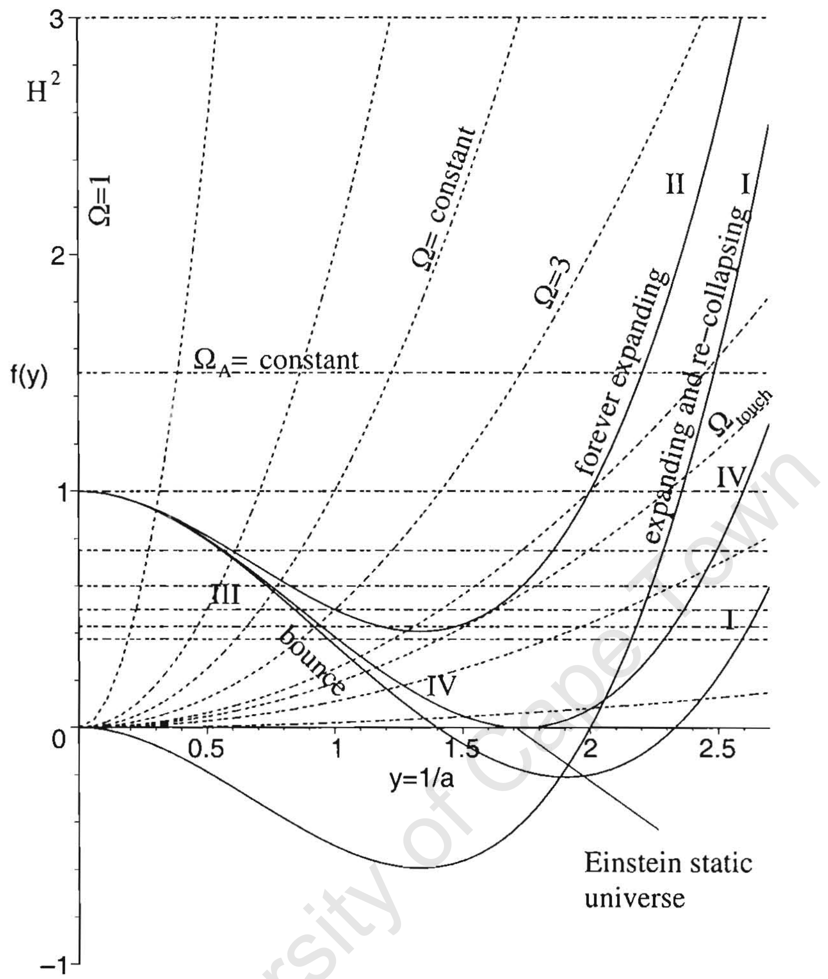


Figure 1.2: Evolution of closed models according to the Friedmann equation $f(y) = H^2$. The initial big-bang corresponds to $y = \infty$. The curves of constant Ω are given by the Friedmann equation $H^2 = \frac{1}{\Omega-1}y^2$.

1.4.2 The measure in $\chi - \Lambda$-parametrization

We have seen that for each of the families of open, closed-type I+II, and closed-type III models the constants of motion χ and Λ give a suitable continuous parametrization. In each case one can use Jaynes' principle to obtain a minimum-information measure.

Open and closed Big-Bang Models For open and closed-type I+II models the parameters take values in the the ranges $\chi \in [0, \infty)$, $\Lambda \in \mathbf{R}$ (with $a > 3\chi/2$ for closed-type I models). We introduce a new parameter [2] $z \in \mathbf{R}$ such that

[2]Note that this choice is not unique. Nevertheless, to use a quadratic relation appears to be the simplest choice. In fact, any relation of the kind $\chi = z^{2n}$ (with $n \in \mathbf{N}$) leads to the same non-integrabilities. In general the relation might be of the form $\chi = f(z)$, where f can be taken such that it is differentiable, has the range $[0, \infty)$, satisfies $f(0) = 0$ and is even.

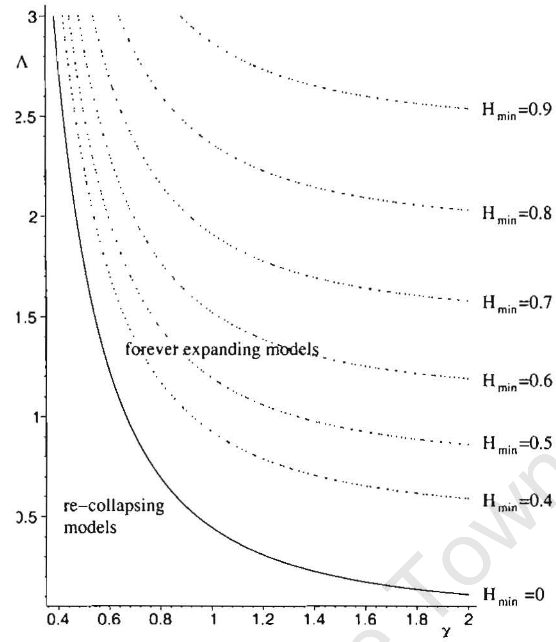


Figure 1.3: Lines of constant minimum Hubble parameter H in the χ-Λ-plane for closed models.

$\chi = z^2$. Applying Jaynes' principle suggests the minimum information measure

$$\mu \propto \mathrm{d}z\mathrm{d}\Lambda \propto \frac{\mathrm{d}\chi\mathrm{d}\Lambda}{\sqrt{\chi}}. \tag{1.12}$$

This measure is a volume element in probability space and as such necessarily covariant under the change of coordinates in probability space.

We note that for fixed cosmological constant Λ this measure does not diverge at $\chi = 0$, but at $\chi = \infty$. Hence it predicts that any given finite range of χ corresponds to a set of measure zero over the universes in the ensemble, *i.e.*, a typical value for χ will be a very 'big' real number. As can be seen in figure 1.4 for any fixed Hubble parameter H high values of χ correspond to density parameters close to unity, *i.e.*, $\Omega \approx 1$. In this way the measure resolves the flatness problem.

The flat probability distribution for the cosmological constant Λ is a result of our assumption of minimum information — to derive the measure we have only used the allowed parameter ranges. There are two points to note. Firstly, (1.12) is only the *a priori* probability distribution. By Bayes' theorem, incorporating additional information, *e.g.*, from anthropic arguments or astronomical observations, will modify the probability distribution such as to reflect the additional knowledge.

Secondly, the failure of the measure to predict the observed very small, but non-zero, value of the cosmological constant might be taken as a sign: maybe

Then the first Taylor term is the quadratic term. This guarantees the right behaviour around $\chi = 0$, *i.e.*, the integrability of the measure. However, in such a case the behaviour for lage values of χ might be different.

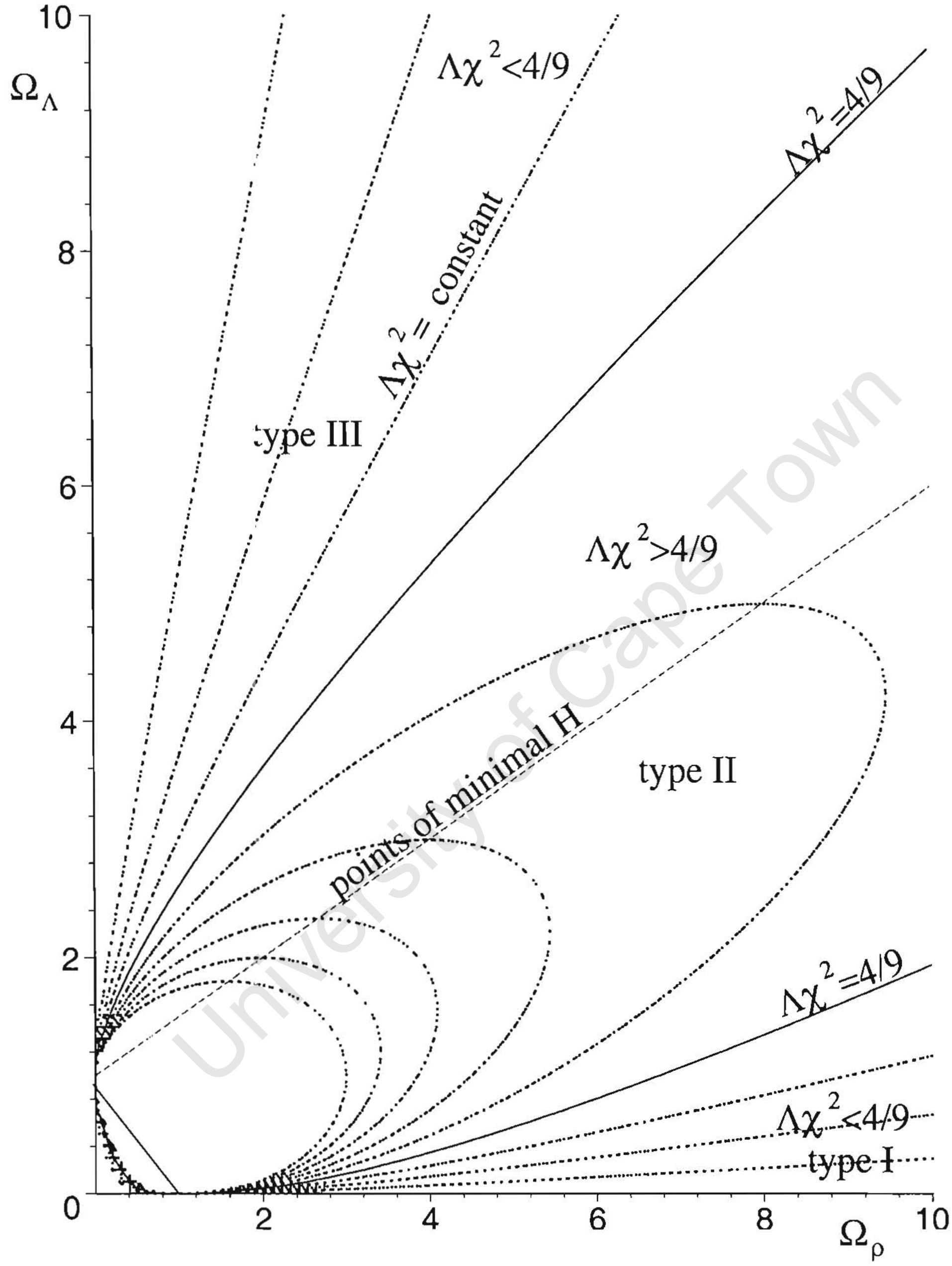


Figure 1.4: The trajectories of $(\Omega_\rho, \Omega_\Lambda)$ for closed FLRW models. The trajectories are characterized by the constant of motion $\Lambda\chi^2$. The solid line $\Lambda\chi^2 = 4/9$ divides the type I/III and type II models. The Hubble parameter H is constant along horizontal lines.

the assumption that Λ can take all real values is not valid. Indeed, if the cosmological constant would be restricted to positive real values, *i.e.*, $\Lambda \in \mathbf{R}^+$, then the measure would take the form $\frac{\mathrm{d}\chi \mathrm{d}\Lambda}{\sqrt{\chi\Lambda}}$, which favours very small and very large values of Λ. Such a restriction might be enforced by some higher theory of quantum-gravity, of which general relativity is the low-energy limit.

Bounce Models For closed-type III models and the corresponding degenerated type-IV model, *i.e.*, bounce models, $\Lambda \in \mathbf{R}^+$, $\chi \in [0, \frac{2}{3\sqrt{\Lambda}}]$, and the scale factor a is restricted to $a \geq \frac{3\chi}{2}$.

We can parametrize χ by setting

$$\chi = \frac{1}{3\sqrt{\Lambda}}[1 - \cos(z)] \tag{1.13}$$

with $z \in \mathbf{R}$. z can take all real values and hence Jaynes' principle yields the

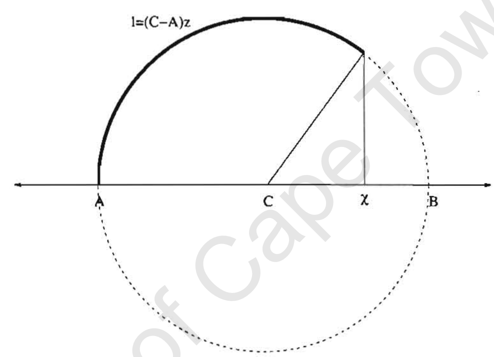


Figure 1.5: The measure (1.14) corresponds to a constant probability distribution on a circle in a fictional plane cutting the χ-axis. $l = C(\arcsin(\chi - C) + \pi/2)$.

minimum-information measure

$$\mu_{\text{bounce}} \propto \left(\frac{1}{\sqrt{\chi}} \frac{1}{\sqrt{\frac{2}{3} - \chi\sqrt{\Lambda}}} \right) \frac{\mathrm{d}\chi \mathrm{d}\Lambda}{\Lambda^{3/4}}. \tag{1.14}$$

This measure is non-integrable for $\Lambda \to \infty$.

1.4.3 Transformation to $\Omega - \Omega_\Lambda$-parametrization

With (1.12) and (1.14) we found measures over the different trajectories in phase space identified by the constants of motion and the nature of the solution. The advantage of these 'co-moving coordinates' in phase-space is that the measure is necessarily time-independent. Nevertheless, χ and Λ are not cosmological observables.

The three observables which uniquely identify the FLRW model of the universe and its age are the Hubble parameter H and the density-parameters Ω_ρ and Ω_Λ. We want a measure over the different trajectories and hence we need to select a hypersurface in phase space which cuts every trajectory, or at least a well defined subset of them, only once. Our measures then translate into measures over this hypersurface.

Such a hypersurface can be defined by fixing the Hubble parameter $H = H_0$ at its present-day value. The resulting measure will then be parametrized by the present-day density parameters $\Omega_{\rho 0}$ and $\Omega_{\Lambda 0}$. The interpretation is clear: given the knowledge of the Hubble parameter H_0 we estimate likelihoods for finding the density parameters in certain ranges. It should be noted that the following considerations hold for any choice of the Hubble parameter H and hence the results will be valid at any time.

We note that this parametrization will automatically exclude all models which never reach this Hubble parameter, *i.e.*, the set of open models which approach $\Lambda/3 \geq H_0$, the set of closed-type II models with $H^2_{\text{min}} = \Lambda/3 - 4/(27\chi^2) > H_0^2$, and the set of bounce models with $\Lambda/3 \leq H_0$.

Furthermore, the issue is complicated by the fact that some closed models of type II (forever expanding) might have $H = H_0$ twice during their evolution (see figure 1.2 where a horizontal line corresponds to a constant Hubble parameter). Hence two different points in the $(\Omega_{\rho 0}, \Omega_{\Lambda 0})$-plane correspond to the same model. This issue can be resolved by excluding the intersection point which is left of the minimum (in figure 1.2).

Hence for closed-type I+II models we restrict ourself to the intersection points right (in figure 1.2) of the minimum, *i.e.*, $a < 3\chi/2$. In terms of the present-day density parameters this condition becomes

$$\Omega_{\Lambda 0} < \frac{\Omega_{\rho 0}}{2} + 1 \tag{1.15}$$

or $\Omega_{\Lambda 0} < (\Omega_0 + 2)/3$, where $\Omega_0 \stackrel{\text{def}}{=} \Omega_{\Lambda 0} + \Omega_{\rho 0}$ is the total density parameter.

For a constant Hubble parameter $H = H_0$ the constants of motion χ and Λ are expressed in terms of the present-day density parameters $\Omega_{\rho 0}$ and $\Omega_{\Lambda 0}$ by

$$\begin{pmatrix} \chi \\ \Lambda \end{pmatrix} = \begin{pmatrix} \Omega_{\rho 0}/(H_0|\Omega_0 - 1|^{3/2}) \\ 3H_0^2\Omega_{\Lambda 0} \end{pmatrix}. \tag{1.16}$$

The determinant of the Jacobian of this transformation is

$$\frac{\mathrm{d}\chi}{\mathrm{d}\Omega_{\rho 0}} \frac{\mathrm{d}\Lambda}{\mathrm{d}\Omega_{\Lambda 0}} = 3H_0^2\chi \left[\frac{1}{\Omega_{\rho 0}} - \frac{3/2}{\Omega_0 - 1}\right]. \tag{1.17}$$

It is interesting to note that the curves of constant-Ω_0 in the $\chi - \Lambda$-plane are straight lines as shown in figure 1.6. In the closed case this lines intersect, which corresponds to the non-uniqueness of the $\chi - \Lambda$-parametrization.

Open and closed Big-Bang models For the open and closed Big-Bang models the measure (1.12) gives

$$\mu \propto \sqrt{\frac{\Omega_{\rho 0}}{|\Omega_0 - 1|^{3/2}}} \left|\frac{1}{\Omega_{\rho 0}} - \frac{3/2}{\Omega_0 - 1}\right| \mathrm{d}\Omega_{\rho 0}\mathrm{d}\Omega_{\Lambda 0}. \tag{1.18}$$

Note that this measure approaches zero as we approach equality of (1.15). Figure 1.8(a) illustrates this measure in the $\Omega_{\rho 0} - \Omega_{\Lambda 0}$-plane. There is a non-integrable divergency along the line $\Omega_{\rho 0} + \Omega_{\Lambda 0} = 1$. For fixed $\Omega_{\Lambda 0}$ the measure is integrable at infinity.

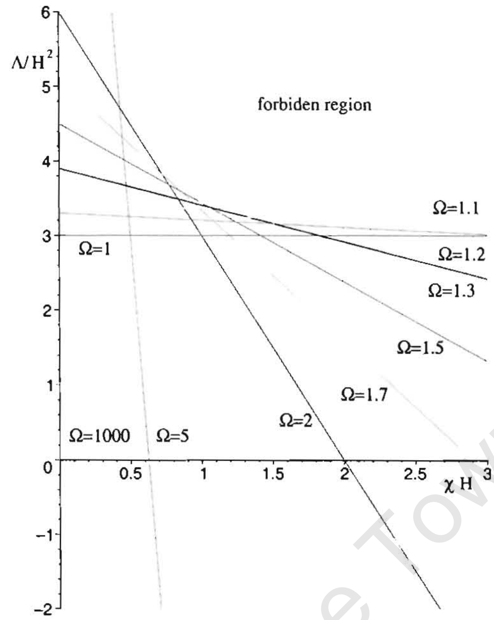


Figure 1.6: Lines of constant density parameter Ω in the $\chi-\Lambda$-plane for closed models.

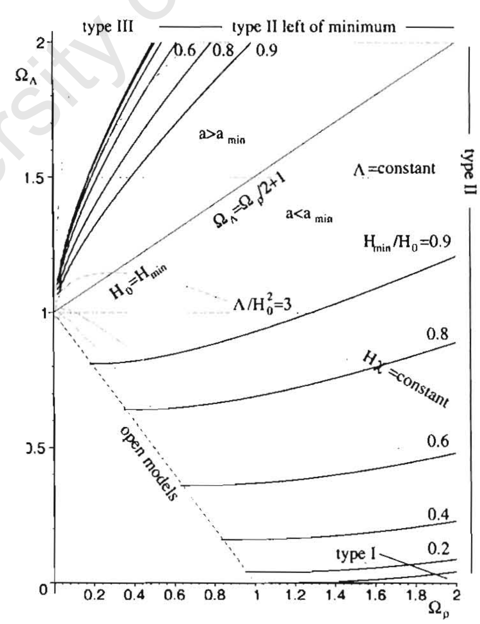
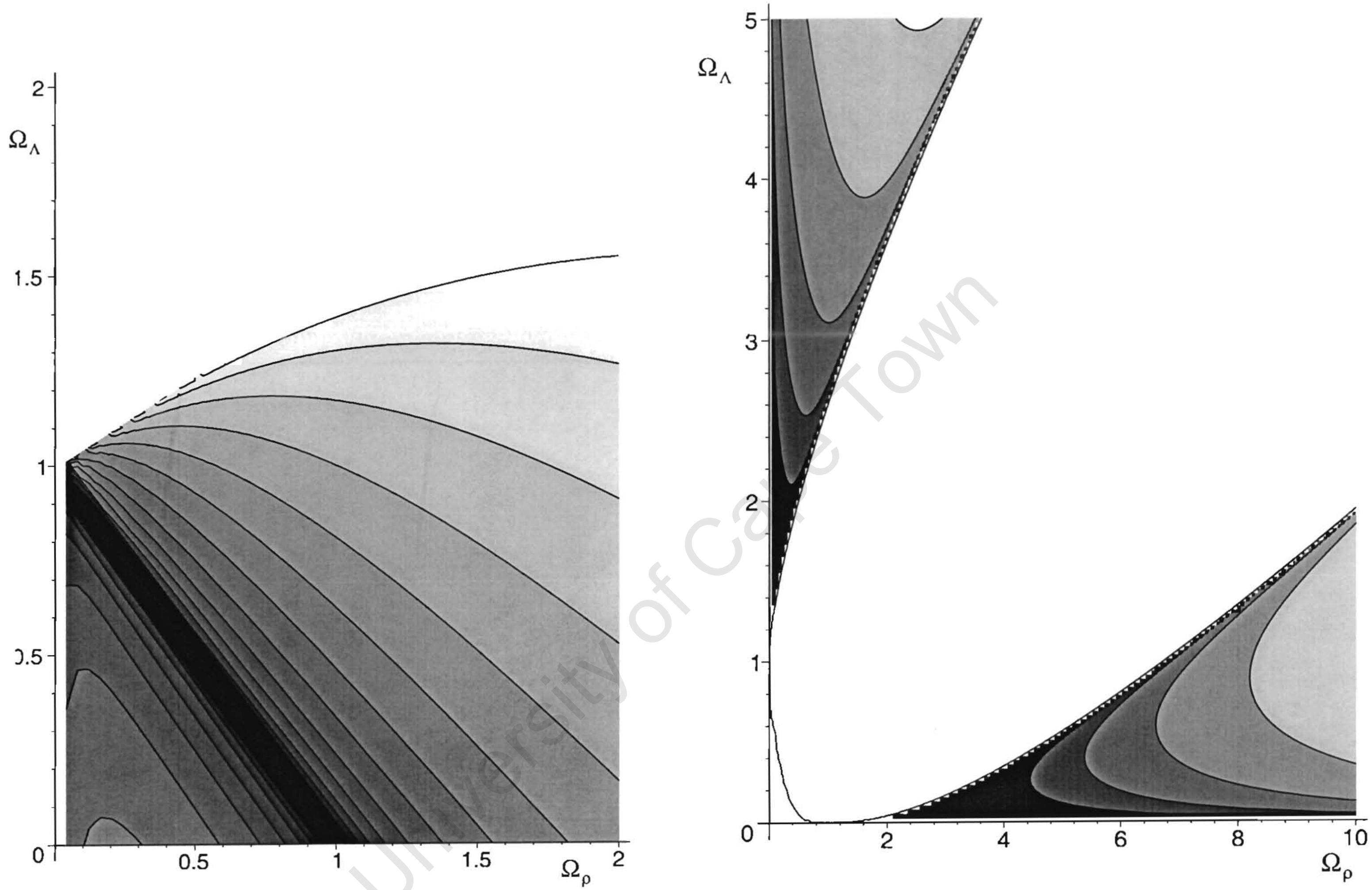


Figure 1.7: Curves of constant minimum Hubble parameter in the $\Omega_\Lambda-\Omega_\rho$-plane for closed models.



(a) Lines of constant probability measure in the $\Omega_\rho - \Omega_\Lambda$-plane for big-bang models. Note that the measure vanishes on the line $\Omega_\Lambda = 1 + \Omega_\rho/2$.

(b) Lines of constant probability measure for the bounce solutions in the $\Omega_\rho - \Omega_\Lambda$-plane.

Figure 1.8: Illustration of the measures in $\Omega_\rho - \Omega_\Lambda$ parametrization.

Closed Bounce Models The measure (1.14) differs from (1.12) only by a factor $(\Lambda\sqrt{\frac{2}{3\sqrt{\Lambda}}-\chi})^{-1}$ so that the measure becomes in $\Omega_{\rho 0}, \Omega_{\Lambda 0}$-parametrization

$$\mu_{\text{bounce}} \propto \left(2\sqrt{3} - 9\Omega_{\rho 0}\sqrt{\frac{\Omega_{\Lambda 0}}{(\Omega_0 - 1)^3}}\right)^{-1/2} \Omega_{\Lambda 0}^{-3/4}\mu, \tag{1.19}$$

where μ is given by (1.18).

The contour-map of this measure is given in figure 1.8(b). Note that the bounce solution only exists for

$$\frac{\Omega_{\Lambda 0}\Omega_{\rho 0}^2}{(\Omega_0 - 1)^3} \leq \frac{4}{27}, \tag{1.20}$$

where we included the limiting case (type IV).

1.4.4 Extension to γ-law models and non-interacting matter components

Above results hold for open and closed dust FLRW models. Nevertheless, the measures for big-bang models are easily extended to all FLRW models with a γ-law equation of state, *i.e.*, models where pressure and energy density are related by a linear relation

$$p = (\gamma - 1)\rho. \tag{1.21}$$

Radiation dominated models have $\gamma = 4/3$, while the dust case is recovered by setting $\gamma = 1$.

The constants of motion for these models are the cosmological constant Λ and the matter constant $\chi_\gamma = \kappa\rho a^{3\gamma}/3$. Using these quantities as parameters for the big-bang models we find in analogy to (1.12)

$$\mu_\gamma \propto \frac{\mathrm{d}\chi_\gamma \mathrm{d}\Lambda}{\sqrt{\chi_\gamma}}. \tag{1.22}$$

and in $\Omega_{\rho 0} - \Omega_{\Lambda 0}$ parametrization

$$\mu_\gamma \propto \frac{\Omega_{\rho 0}^{1/2}}{|\Omega_0 - 1|^{3\gamma/4}} \left|\frac{1}{\Omega_{\rho 0}} - \frac{3\gamma}{2}\frac{1}{\Omega_0 - 1}\right| \mathrm{d}\Omega_{\rho 0}\mathrm{d}\Omega_{\Lambda 0}. \tag{1.23}$$

In particular we find for radiation dominated models with $\gamma = 4/3$

$$\mu_{\text{rad}} \propto \frac{\Omega_{\rho 0}^{1/2}}{|\Omega_0 - 1|} \left|\frac{1}{\Omega_{\rho 0}} - \frac{2}{\Omega_0 - 1}\right| \mathrm{d}\Omega_{\rho 0}\mathrm{d}\Omega_{\Lambda 0}. \tag{1.24}$$

Similarly the measure is easily extended to two non-interacting matter components, each with γ-equation of state. In this case each matter component has a constant of motion $\chi_{1/2} = \frac{\kappa a^{3\gamma_{1/2}}\rho_{1/2}}{3}$, where $\gamma_{1/2}$ is describes the equation of state for each component. The probability space has now three dimensions and a similar reasoning as above suggests the minimum information measure

$$\mu_{\gamma_1,\gamma_2} \propto \frac{\mathrm{d}\chi_1 \mathrm{d}\chi_2 \mathrm{d}\Lambda}{\sqrt{\chi_1\chi_2}}. \tag{1.25}$$

Translating this into $\Omega_{\rho 0} - \Omega_{\Lambda 0}$ parametrization gives the generalization of (1.23)

$$\mu_{\gamma_1,\gamma_2} \propto \frac{(\Omega_{10}\Omega_{20})^{1/2}}{|\Omega_0 - 1|^{3(\gamma_1+\gamma_2)/4}} \left| \frac{1}{\Omega_{10}\Omega_{20}} - \frac{3(\gamma_1/\Omega_{20} + \gamma_2/\Omega_{10})}{2(\Omega_0 - 1)} \right| \mathrm{d}\Omega_{\rho 0}\mathrm{d}\Omega_{\Lambda 0}, \quad (1.26)$$

where Ω_{10} and Ω_{20} are the present-day density parameters for the two matter components.

It is intersting to note that the measure (1.26) is non-integrable around $\Omega_0 = 1$, but has an integrable divergence around $\Omega_{10}\Omega_{20} = 0$. The measure does not supply any information about the relation between the densities of the two matter components.

1.5 Additional Constraints and the Flatness Problem

Anthropic and experimental constraints on the possible values of the cosmological parameters do not change the form of the minimal-information measures as derived from Jaynes' principle, because they do not restrict the in principle allowed parameter range but rather measure the actual value for our universe. Instead they provide additional information which updates the probability measure according to Bayes' theorem.

Let us assume an anthropic constraint on the value of the cosmological constant which restricts the corresponding density parameter (at the present time) to some finite interval I_2, *i.e.*, $\Omega_{\Lambda 0} \in I_2$. According to Bayes' theorem the probability that the value of the cosmological constant lies in some interval I_1 becomes now (here C represents the minimal information)

$$P(\Omega_{\Lambda 0} \in I_1 | \Omega_{\Lambda 0} \in I_2, C) = \frac{P(\Omega_{\Lambda 0} \in I_1 \cap I_2 | C)}{P(\Omega_{\Lambda 0} \in I_2 | C)} P(\Omega_{\Lambda 0} \in I_2 | \Omega_{\Lambda 0} \in I_1 \cap I_2, C). \quad (1.27)$$

The second factor is unity for $I_1 \cap I_2 \neq \emptyset$ and otherwise zero. Assuming for the moment that the probability measure is integrable for $\Omega_{\Lambda 0} \in I_2$ and $\Omega_{\rho 0} \in \mathbf{R}^+$ we can express the remaining probabilities as integrals over the *a priori* measure and obtain

$$\frac{\int_{\Omega_{\rho 0} \in \mathbf{R}^+} \int_{\Omega_{\Lambda 0} \in I_1 \cap I_2} \mu}{\int_{\Omega_{\rho 0} \in \mathbf{R}^+} \int_{\Omega_{\Lambda 0} \in I_2} \mu} = \int_{\Omega_{\rho 0} \in \mathbf{R}^+} \int_{\Omega_{\Lambda 0} \in I_1 \cap I_2} \mu^*, \quad (1.28)$$

where μ^* is a new measure given by the normalized old measure restricted to the anthropically allowed region with $\Omega_{\Lambda 0} \in I_2$ (zero anywhere else). Since the *a priori* measure μ given by (1.18) is non-integrable around $\Omega_{\rho 0} + \Omega_{\Lambda 0} = 1$ some of the above integrals might be undefined. However, the denominator in (1.28) is merely a normalization factor for the new measure μ^* and above arguments still hold if we accept that μ^* might be non-normalisable too.

Because the interval I_2 is finite the new measure μ^* is only non-integrable for $\Omega_0 = \Omega_{\rho 0} + \Omega_{\Lambda 0} = 1$ *i.e.*, the non-integrabilities at infinity are removed. This is exactly what we need to solve the flatness problem. We conclude that our *a priori* measure (1.18) can solve the flatness problem together with constraints which restrict the cosmological constant to a finite range.

1.6 Conclusion

We discussed the measures suggested by Gibbons, Hawking, and Stewart [17] and Evrard and Coles [20] and propose here an alternative.

We investigated the parametrization of dust-FLRW models by the constants of motion and suggested new measures based on Jaynes' principle for each class of dust models with a continuous parametrization.

For open and closed big-bang models the minimum information measure is given by $\frac{d\chi d\Lambda}{\sqrt{\chi}}$. Though the measure is non-integrable, it is integrable in any finite region around $\chi = 0$ and hence predicts that almost all universes lay outside this region.

The uniform probability distribution for the cosmological constant Λ is a direct result of missing constraints on its parameter range. Additional information will modify the measure according to Bayes' theorem such as to reflect more knowledge, like anthropic or astronomical constraints. Nevertheless, it is only the *a priori* probability which allows us to decide whether our universe is a likely one or not and it could be argued that the failure of above measure to favour small, non-zero values of the cosmological constant is a hint that Λ is restricted by some underlying theory.

It should be noted that the measures (1.12) and (1.18) (without the $d\Lambda$ and $d\Omega_{\Lambda 0}$) can be derived under the assumption of a fixed cosmological constant Λ. The corresponding measure in Ω_0 parametrization is then integrable for $\Omega_0 \to 0$ and $\Omega_0 \to \infty$ but not for $\Omega_0 \to 1$ for any fixed value of Λ. Hence if Λ would not be taken as a random variable this implies that almost all models have a density parameter Ω_0 infinitesimally close to unity. This in fact would solve the Flatness Problem, but it is disturbing that Λ has to be treated on a different footing.

With (1.18) and (1.19) we expressed the measures in terms of observable quantities by transforming it into $\Omega_{\rho 0} - \Omega_{\Lambda 0}$-parametrization for a constant Hubble parameter H_0. It diverges only for $\Omega_0 \to 1$, but it is non-integrable for $\Omega_0 \to 1$ and $\Omega_0 \to \infty$. Nevertheless, if $\Omega_{\Lambda 0}$ is restricted by additional information to a finite range then our measure has only one non-integrable divergency at $\Omega_0 = 1$, and not like in [20] a second at $\Omega_0 = 0$. We conclude that the measures alone cannot resolve the Flatness Problem. Nevertheless, the given measures are *a priori* measures and together with some additional information on the possible range of Λ it can resolve the flatness problem.

Furthermore, it should be mentioned that the divergent nature of the measures, which is a direct consequence of the unbounded parameter ranges, can lead to ambiguities for the ratio of *a priori* probabilities. For example the ratio of the *a priori* probabilities for closed models of type II and I leads to a ratio of two undefined limits. Similar to the Flatness Problem this could be resolved by giving more information.

For open and closed big-bang models the measures were generalized to FLRW models with a γ-equation of state and are given by (1.23). For two non-interacting matter components the measure took the form (1.26).

References

[1] Melchiorri A 2002 *Preprint* astro-ph/0204017

[2] de Bernardis P *et al* 2002 *Astrophys. J.* **564** 559-566 (*Preprint* astro-ph/0105296)

[3] Stompor R *et al* 2001 *Astrophys. J.* **561** L7-L10 (*Preprint* astro-ph/0105062)

[4] Guth A H 1981 *Phys. Rev.* D **23** 347

[5] Barrow J D and Tipler F J 1986 *The Anthropic Cosmological Principle*, (New York: Oxford University Press)

[6] Rees M J 2001 *Preprint* astro-ph/0101268

[7] Garrett A J M and Coles P 1993 *Comments on Astrophys.* **17** 23-47

[8] Cox R T 1946 *Am. J. Phys.* **14** 1

[9] Garrett A J M 1991 in *Maximum Entropy in Action*, ed B Buck and V A Macaulay (Oxford: Clarendon Press) pp 139-170

[10] Jaynes E T 1983 *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics* ed R D Rosenkrantz, *(Synthese Series vol 158)* (Dordrecht: Reidel)

[11] Shannon C E 1948 *Bell Syst. Tech. J.* **27** 379-423 and 623-659

[12] Jaynes E T 1982 *Proceedings of the IEEE* **70** 939-952

[13] Shore J E and Johnson R W 1980 *IEEE Trans. Inf. Theory* **26** 26-37

[14] Gull S F and Skilling J 1984 *IEE Proceedings* **131F** 646-659

[15] Jaynes E T 1968 *IEEE Trans. Syst. Sci. Cybern.* **4** 227-241 (reprinted in [10])

[16] Jaynes E T 2003 *Probability Theory* (Cambridge: Cambridge University Press)

[17] Gibbons G W, Hawking S W and Stewart J M 1987 *Nucl. Phys.* B 736-751

[18] Hawking S W and Page D N 1988 *Nucl. Phys.* B **298** 789-809

[19] Page D N 1987 *Phys. Rev.* D **36** 1607

[20] Evrard G and Coles P 1995 *Class. Quantum Grav.* **12** L93-L97 (*Preprint* astro-ph/9507020)

Chapter 2

Spherically symmetric junction surfaces and the matching of FLRW regions

2.1 Introduction

Recent measurements of the microwave background radiation support the idea that our universe is highly isotropic and homogeneous [1]. Cosmological models with these properties are uniquely represented by the class of Friedmann-Lemaître-Robertson-Walker (FLRW) models. The observed flatness of the universe can then be explained by employing an inflationary model [2], which suggests an exponential expansion of the early universe driven by a scalar field – the inflaton field.

Nevertheless, it is often speculated that this might only be the local geometry, while over a larger scale the universe is inhomogeneous and anisotropic, i.e., the matter content and geometry vary. In particular, it appears as if the parameters necessary for life are highly fine-tuned and in order to solve this "fine-tuning problem" it was suggested that we live in one of many different FLRW regions – most of them might be unsuitable for life. The most prominent example is Linde's Chaotic Inflation Scenario [3, 4], in which the different FLRW regions originate from different almost homogeneous Planck-sized regions which experience a period of exponential expansion.

When such models are discussed it is usually assumed that the transition region between two almost FLRW regions is very small and can be approximated by a timelike junction hypersurface, the so-called thin bubble wall. To find the motion of this hypersurface one has to find the matching surface to the two solutions of the Einstein-field equation representing the space-time on each side.

The matching conditions are of two different types. On the one hand there is the purely geometric necessity that 'things fit together' — distances on the junction surface should have the same length when measured 'inside' or 'outside'. On a mathematical level this reduces to a matching of the tangential metric components.

On the other hand there are the matching conditions which result from the assumed validity of certain physical laws, in particular the energy-momentum conservation across the junction surface and the validity of the Einstein-field equation on each side. These conditions have been evaluated by C. Lanczos [5], R. Dautcourt [6], and in a ground-breaking work by W. Israel [7].

While these equations are in principle valid for any matching of two space-times satisfying the Einstein-field equation, it is in practice impossible to handle their complexity except for highly symmetric cases. In particular the matching of spherically symmetric space-times has been studied extensively (e.g., [7, 8, 9, 10]).

Our aim here is to present a new approach to junctions between spherically symmetric space-times and to apply the formalism to junctions between FLRW models — analytically and numerically.

The approach will focus on the geometrical quantities describing the situation, i.e., the distance of the junction surface from the centre of symmetry, and not alone on the junction surface radius. In contrast to most other studies, we do not evaluate the junction conditions (in particular the Lanczos equation) in the original coordinate system or in Gaussian normal coordinates based on the junction surface. Instead we introduce new coordinates such that the junction surface is at a fixed (new) 'radial' coordinate and all coordinates are continuous at the junction surface. The motion of the junction surface is now absorbed into the metric components and the continuity of the radial coordinate ensures the unambiguous identification of the normals on each side of the junction surface.

In spherically symmetric cases the Lanczos equation has two non-trivial independent components - an angular and a time component. While the first one leads to the well-known evolution equations, there seems to be uncertainty about the interpretation of the time component, which is a second order (in time) differential equation for the junction surface motion. It has been known that for certain particular cases this equation reduces to an identity [11]. Nevertheless, other authors[1] suggested that this equation acts as a surface equation of state [9], i.e., determines the surface pressure. Using the presented approach we will show that the time component of the Lanczos equation is in fact an identity for all junctions between spherically symmetric solutions of the Einstein-field equation.

It should be pointed out that there are special cases of junctions which could be examined without employing junction conditions. If the γ-equation of state and the cosmological constant have no discontinuity at the junction surface then the spherically symmetric space-time can be described in terms of the Lemaître-Tolman model. However, the really intersting question is how the junction behaves if the inside and outside region have different dynamical behaviour, i.e., different equation of state and cosmological constant. For these cases one cannot avoid the use of junction conditions and all numerical examples given in this Chapter will be of this kind.

This chapter will be structured as follows: In section 2.2 we re-examine junction conditions for the matching of generic spherically symmetric sections. We will re-derive the evolution and constraint equations using a new approach, based on a coordinate transformation which makes *all* coordinates continuous at the junction surface. We investigate the behaviour for small values of the surface-energy density and discuss the special case of vanishing surface-energy

[1] In particular this was suggested for the junction between FLRW models.

density. In section 2.3 we pay particular attention to constraints on the surface-energy density and in section 2.4 we show that the time-component of the Lanczos equation is an identity. This is followed by an application to the matching of FLRW sections in section 2.5 and numerous numerical examples in section 2.6.

2.1.1 Notations and Definitions

The Einstein summation convention will be used, i.e., any index which appears twice (once up and once down) in a term is understood to be summed over. We label the coordinates with greek indices, running from 0 to 3, where 0 represents the time coordinate. Latin indices label coordinates on the three-dimensional junction surface hypersurface. As will be seen below, with our choice of coordinates the junction surface is located at a fixed radial coordinate $R = 1$, and hence latin indices take the values $0, 2$, and 3 (or equivalently t, θ, and ϕ).

We will use the convention that round/square brackets around indices represent the symmetrization/antisymmetrization operator, i.e., $T_{(\mu\nu)} \stackrel{\text{def}}{=} \frac{1}{2}(T_{\mu\nu} + T_{\nu\mu})$ and $T_{[\mu\nu]} \stackrel{\text{def}}{=} \frac{1}{2}(T_{\mu\nu} - T_{\nu\mu})$.

We use a metric signature $(- + ++)$, i.e., $u_\mu u^\mu < 0$ if and only if u_μ is a timelike vector. The Christoffel symbols are given in terms of the metric by

$$\Gamma_{\mu\nu\lambda} \stackrel{\text{def}}{=} \frac{1}{2}(g_{\mu\nu,\lambda} + g_{\lambda\mu,\nu} - g_{\nu\lambda,\mu}),$$

where a comma indicates the ordinary partial derivative with respect to the coordinates, e.g., $g_{\mu\nu,\lambda} = \frac{\partial}{\partial x^\lambda} g_{\mu\nu}$. The covariant derivative of a vector u^μ is then given by

$$u^\mu{}_{;\nu} = u^\mu{}_{,\nu} + u^\lambda \Gamma^\mu{}_{\nu\lambda} \qquad u_{\mu;\nu} = u_{\mu,\nu} - u_\lambda \Gamma^\lambda{}_{\nu\mu}.$$

When considering quantities defined for certain hypersurfaces (like the junction surface) it will be convenient to have a covariant derivative for this subspace. We will use a vertical bar (like in $K_{ab|c}$) to refer to this covariant hypersurface derivative, which is evaluated in the same way as above covariant derivative, but with all quantities replaced by the corresponding hypersurface quantities. Any quantity referring to the three-dimensional hypersurface of the junction surface will have a superscript '(3)', e.g., ${}^{(3)}R$ and ${}^{(3)}g_{ab}$.

The Riemann curvature tensor is defined in terms of the Christoffel symbols by

$$R_\mu{}^\nu{}_{\lambda\kappa} \stackrel{\text{def}}{=} \Gamma^\nu{}_{\lambda\mu,\kappa} - \Gamma^\nu{}_{\kappa\mu,\lambda} + \Gamma^\alpha{}_{\lambda\mu}\Gamma^\nu{}_{\kappa\alpha} - \Gamma^\alpha{}_{\kappa\mu}\Gamma^\nu{}_{\lambda\alpha}.$$

The Ricci tensor and scalar are defined as contractions of the Riemann curvature tensor by $R_{\mu\nu} \stackrel{\text{def}}{=} R^\lambda{}_{\mu\lambda\nu}$ and $R \stackrel{\text{def}}{=} R^\mu{}_\mu$.

We will use a subscript $-$ or $+$ sign to indicate whether quantities refer to the in- or outside region, respectively.

2.2 Matching of generic O(3)-symmetric sections

2.2.1 The coordinate system

Any spherically symmetric space-time (i.e., having O(3) symmetry) allows coordinates such that the metric takes the form

$$\mathrm{d}s^2 = -\mathfrak{N}^2(\tau, r)\mathrm{d}\tau^2 + l^2(\tau, r)\{\mathrm{d}r^2 + f^2(\tau, r)\mathrm{d}\Omega^2\}, \tag{2.1}$$

where $\mathfrak{N}(\tau, r)$ is the so-called lapse function, and $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\phi^2$ the line-element on the two-dimensional unit-sphere.

We are interested in the matching of two spherically symmetric space-times, each having a metric of the form (2.1). Generally, the coordinates will not match up at the junction surface and the manifold is described by two different coordinate charts – one for the inside region (subscript $-$) and one for the outside region (subscript $+$). At any coordinate time (in the inside or outside region) the junction surface itself is assumed to be a two-sphere which is described by its coordinate radius in the inside and outside region, $\alpha_-(\tau_-)$ and $\alpha_+(\tau_+)$.

Since the inside and outside regions are originally described by a metric in the form (2.1) we will use the convention that a dot/prime refers to the proper time/radial derivative with respect to the metric (2.1) at the junction surface, e.g.,

$$\dot{l}_+ \stackrel{\mathrm{def}}{=} \frac{1}{\mathfrak{N}_+}\frac{\partial}{\partial\tau}l_+(\tau, r)\bigg|_{r=\alpha_+(\tau_+)} \quad \text{and} \quad l'_+ \stackrel{\mathrm{def}}{=} \frac{1}{l_+}\frac{\partial}{\partial r}l_+(\tau, r)\bigg|_{r=\alpha_+(\tau_+)}.$$

While the angular coordinates θ and ϕ can be chosen such that they are continuous at the junction surface, this is generally not true for the radial and time coordinates r and t – the same point on the junction surface is represented by different coordinate values on each side. In order to describe the motion of the junction surface one usually tries to find the evolution of the junction surface radius in each coordinate system.

We want to suggest a different approach: we introduce a new coordinate system, such that *all* the coordinates are continuous at the junction surface while only the transverse metric components are discontinuous at the junction surface. The junction surface motion is now described by the evolution of the metric components.

Constructing a continuous time coordinate Let us assume that the inside and outside spaces are given in terms of their metrics, which take the form (2.1). Generally the time coordinates for the inside and outside region will not match up at the junction hypersurface. Nevertheless, it is possible to re-scale the time coordinate on each side of the junction surface by setting

$$\mathrm{d}\tau = \eta(t)\mathrm{d}t$$

in such a way that the new time coordinates t_+ and t_- match up at the junction surface. This gives us a 'global time coordinate' t.

If the original lapse function was constant then this re-scaling results in a new time dependent lapse function which contains information about the junction surface motion. For example, this will be the case for the matching of FLRW models.

If on the other hand the original lapse function $\mathfrak{N}(\tau, r)$, where τ is the original time coordinate, depends on the radial coordinate then we can write the new lapse function (which makes the time coordinate continuous) as

$$N(t,r) = \eta(t)\mathfrak{N}(\tau, r)\,,$$

where $\tau = \int_0^t \eta(t')\mathrm{d}t'$ and t is the new time coordinate. It is now $\eta(t)$ which contains information about the junction surface motion, while $\mathfrak{N}$ contains information about the background space. In particular the quantity

$$\frac{N'(t,r)}{N(t,r)} = \frac{\mathfrak{N}'(\tau,r)}{\mathfrak{N}(\tau,r)}$$

(taken as a function of proper time) does not depend on η. We will keep these issues in mind when we use the lapse function in the following calculations.

Constructing a continuous radial coordinate Now we want to construct new radial coordinates such that the junction surface is at a fixed radial coordinate $R = 1$. This can be achieved by setting $r_\pm = \alpha_\pm(t)R$, where the subscript $+$ refers to outside ($R > 1$) and $-$ to the inside region, and $\alpha_\pm(t)$ is the coordinate radius of the junction surface at time t. The relation between the old and the new coordinates is illustrated in figure 2.1. With this new radial coordinate the metrics for the inside and outside region take the form

$$\mathrm{d}s^2 = -N^2\{1 - R^2\dot{\alpha}^2 l^2\}\mathrm{d}t^2 + 2\alpha N\dot{\alpha}l^2 R\mathrm{d}R\,\mathrm{d}t + l^2\{\alpha^2\mathrm{d}R^2 + f^2(\alpha R)\mathrm{d}\Omega^2\}, \quad (2.2)$$

where a dot indicates the proper time derivative along paths of constant r, θ, ϕ, i.e., $\dot{\alpha} \stackrel{\text{def}}{=} \frac{1}{N}\frac{\partial}{\partial t}\alpha$. It will be useful to have the inverse metric at hand, which takes the form

$$g^{\mu\nu} = \begin{pmatrix} -\frac{1}{N^2} & \frac{\dot{\alpha}R}{\alpha N} & 0 & 0 \\ \frac{\dot{\alpha}R}{\alpha N} & \frac{1-R^2\dot{\alpha}^2 l^2}{l^2\alpha^2} & 0 & 0 \\ 0 & 0 & \frac{1}{l^2 f^2} & 0 \\ 0 & 0 & 0 & \frac{1}{l^2 f^2 \sin^2(\theta)} \end{pmatrix}. \quad (2.3)$$

2.2.2 Geometric matching conditions

At the matching surface $R = 1$ the tangential metric components must be continuous. This gives us the two matching conditions

$$[lf] = 0. \quad (2.4)$$

and

$$\left[N^2(\dot{\alpha}^2 l^2 - 1)\right] = 0, \quad (2.5)$$

where

$$[g(R)] \stackrel{\text{def}}{=} \lim_{R\to 1^+}(g(R)) - \lim_{R\to 1^-}(g(R)) = \lim_{r\to\alpha(t)^+}(g(r)) - \lim_{r\to\alpha(t)^-}(g(r)).$$

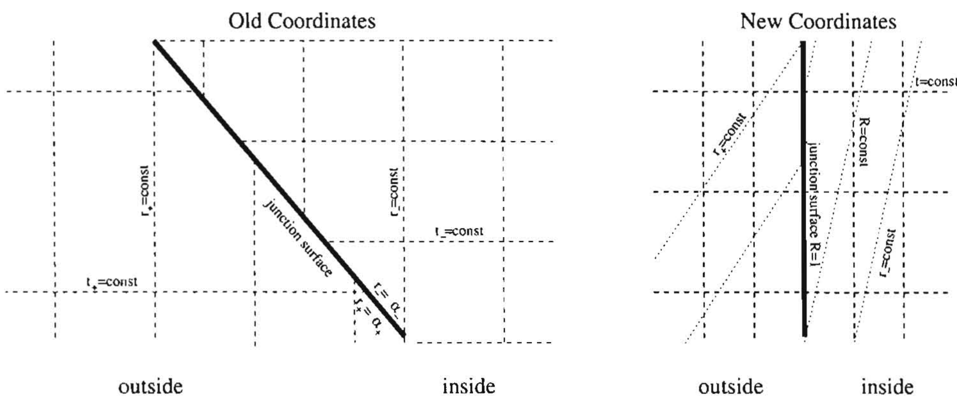


Figure 2.1: An illustration of the coordinate transformation. The original radial coordinates r_+ and r_- are rescaled such that in the new coordinate system the junction surface is at a fixed 'radial' coordinate $R = 1$. The time-coordinates are rescaled such that they match up at the junction surface.

These relations identify two quantities which are continuous across the junction surface and we define

$$k(t,R) \stackrel{\text{def}}{=} N\sqrt{1-\dot{\alpha}^2 l^2 R^2} \quad , \quad k \stackrel{\text{def}}{=} k(t,1) = N\sqrt{1-\dot{\alpha}^2 l^2}\Big|_{R=1} \tag{2.6}$$

and

$$w \stackrel{\text{def}}{=} lf, \tag{2.7}$$

which are the tangential metric components. Note that $k(t, R)$ becomes complex for large $\dot{\alpha}^2 l^2 R^2$. However, this is not of relevance for the problem at hand, since we are only interested in the behaviour around the junction surface at $R = 1$ where $k(t. R)$ is real for timelike junction surfaces.

It follows from (2.5) that if the surface appears from one side as timelike, it will so from the other side. From now on let us assume that the junction surface is a timelike surface, *i.e.*, $\dot{\alpha}_\pm l_\pm < 1$.

There is a remaining gauge freedom: we can rescale the global time coordinate, i.e., multiply the lapse functions with a time dependent factor. One particularly useful choice is to rescale the time such that the tangential metric component in the timelike direction parallel to the junction surface becomes unity, i.e.,

$$k = N\sqrt{1-\dot{\alpha}^2 l^2}\Big|_{R=1} = 1. \tag{2.8}$$

To maintain generality we will not assume this choice until explicitly stated (in section 2.2.4).

The two conditions (2.4) and (2.5) are of pure geometric character – they have to be satisfied independently of the evolution equations at all times. Taking the total derivative of (2.7) with respect to coordinate time we obtain the corresponding restriction on the junction surface motion

$$\frac{\mathrm{d}w}{\mathrm{d}t} = N\{(lf)^\bullet + (lf)'\dot{\alpha}l\} \qquad \left[\frac{\mathrm{d}w}{\mathrm{d}t}\right] = 0. \tag{2.9}$$

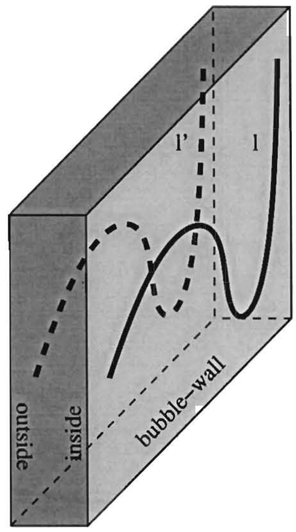


Figure 2.2: The geometric matching condition ensures that curves on the hypersurface measure the same length on each side of the junction surface.

It should be noted that here N is not independent, but depends via (2.8) on the junction surface motion. Equations (2.8) and (2.9) generally have two solutions for the junction surface motion in terms of the surface radius evolution.

Let us note here that the metric of the timelike hypersurface representing the junction surface is given by

$$^{(3)}g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \stackrel{\mathrm{def}}{=} \mathrm{d}s_\Sigma^2 = -k^2\mathrm{d}t^2 + w^2\mathrm{d}\Omega^2.$$

The intrinsic geometry of the junction hypersurface is completely defined by k and w. Nevertheless, w might not uniquely identify the position of the junction surface as is illustrated in figure 2.3. Only if l_+f_+ *and* l_-f_- are invertible functions of $\alpha_\pm$ then the position of the junction surface is indirectly given by the value of w. It can easily be seen that the derivatives of lf are given by

$$-(lf)\frac{\partial(lf)}{\partial x^\mu} = \Gamma_{\mu\theta\theta},$$

and in particular the evolution of the junction surface radius with respect to coordinate time is

$$\frac{\partial w}{\partial t} = -\frac{1}{w}\Gamma_{t\theta\theta},$$

which is continuous across the junction surface.

2.2.3 Lanczos equation and Israel junction conditions

Let us split the energy-momentum tensor in a regular and a δ-function part, so that

$$T_{\mu\nu} = \delta(\eta)S_{\mu\nu} + \tilde{T}_{\mu\nu}, \tag{2.10}$$

where $\tilde{T}_{\mu\nu}$ contains the regular part and η is a function of the coordinates which vanishes on the junction surface, is non-zero everywhere else, and on the junction

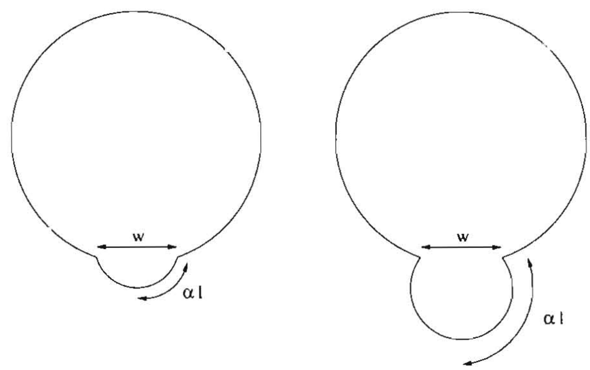


Figure 2.3: The junction surface radius w does not necessarily identify the position of the junction surface, represented by αl. Here the two matching 'surfaces' have the same surface radius $w/2$, but different positions αl.

surface its gradient is a unit vector. The tensor $S_{\mu\nu}$ is called the surface stress-energy (or energy-momentum) tensor. The δ-function restricts its influence to the junction surface and we assume that it only depends on coordinates on the junction surface, i.e., in our case this tensor does not depend on R. The Lanczos equation [5] relates the surface energy-momentum tensor $S_{\mu\nu}$ to the jump in the extrinsic curvature $K_{\mu\nu}$ of the junction surface by

$$\kappa S^{\mu\nu} = {}^{(3)}g^{\mu\nu}[K] - [K^{\mu\nu}], \tag{2.11}$$

or equivalently (after taking the trace and substituting back for K)

$$\frac{1}{\kappa}[K^{\mu\nu}] = -S^{\mu\nu} + \frac{1}{2}{}^{(3)}g^{\mu\nu}S, \tag{2.12}$$

where $K \stackrel{\text{def}}{=} K^{\mu}{}_{\mu}$, $S \stackrel{\text{def}}{=} S^{\mu}{}_{\mu}$, and $\kappa \stackrel{\text{def}}{=} 8\pi G$. These two equations imply that the presence of a surface layer is equivalent to a jump in the extrinsic curvature, *i.e.*, $\gamma^{\mu\nu} \stackrel{\text{def}}{=} [K^{\mu\nu}] \neq 0$.

To relate this conditions to the metric components on both sides of the junction surface we have to find expressions for the extrinsic curvature. We start with the normal to the junction surface, which is given by

$$n_{\mu} = \delta^{R}_{\mu} \frac{l\alpha}{\sqrt{1 - R^2\dot{\alpha}^2 l^2}}\Bigg|_{R=1}, \tag{2.13}$$

and the unique timelike unit-vector tangential to the junction surface and orthogonal to the spherical symmetric subspace, which is given by

$$u^{\mu} = \frac{1}{k}\delta^{\mu}_{t}. \tag{2.14}$$

The projection tensor for the junction surface becomes

$$h^{\mu\nu} = g^{\mu\nu} - n^{\mu}n^{\nu} = u^{\mu}u^{\nu} + \frac{1}{l^2 f^2}(\delta^{\mu}_{\theta}\delta^{\nu}_{\theta} + \sin^{-2}(\theta)\delta^{\mu}_{\phi}\delta^{\nu}_{\phi})$$

and the extrinsic curvature on each side of the junction surface is given by

$$K_{\mu\nu} = \frac{1}{2}(h_{\mu\nu;\lambda}n^{\lambda} + n_{\mu;\nu} + n_{\nu;\mu}) = n_{(\lambda;\kappa)}h^{\lambda}{}_{\mu}h^{\kappa}{}_{\nu}. \tag{2.15}$$

It is easy to verify that the extrinsic curvature is symmetric and tangential, i.e., $K_{\mu\nu} = K_{(\mu\nu)}$ and $K_{\mu\nu}n^{\mu} = 0$. From (2.15) we derive convenient expressions for the extrinsic curvature of the junction surface in terms of unit-vector derivatives and Christoffel-symbols. In particular, the two independent components are given by

$$\begin{aligned} K_{\mu\nu}u^{\mu}u^{\nu} &= -n_{\mu}u^{\mu}{}_{;\nu}u^{\nu} = \frac{l\alpha N}{k^3}\Gamma^{R}{}_{tt} && (2.16)\\ K_{\theta\theta} &= n_{\rho}\Gamma^{\rho}{}_{\theta\theta} = \frac{Nlf}{k}\Gamma^{R}{}_{\theta\theta}. && (2.17)\end{aligned}$$

Setting $k = 1$ (for all times) and using the coordinate time derivative of (2.8) we find

$$K_{\mu\nu}u^{\mu}u^{\nu} = \frac{1}{Nl\alpha}\Gamma_{Rtt} = -\frac{\dot{N} + 2N'l\dot{\alpha} + \dot{\alpha}^2 l^2 N(\dot{l}/l)}{\dot{\alpha}l}. \tag{2.18}$$

$$K^{\theta}{}_{\theta} = \frac{N}{w}(\dot{\alpha}l(lf)^{\bullet} + (lf)'). \tag{2.19}$$

This form of the extrinsic curvature implies that the surface energy-momentum tensor $S^{\mu\nu}$ is diagonal and of perfect-fluid form (in the three-dimensional hypersurface space). We introduce the surface energy density ρ_s and pressure p_s such that

$$S^{\mu\nu} = (\rho_s + p_s)u^{\mu}u^{\nu} + p_s{}^{(3)}g^{\mu\nu}. \tag{2.20}$$

The Lanczos equations (2.12) are now given by the two independent equations

$$\begin{aligned} [N\{\dot{\alpha}l(fl)^{\bullet} + (fl)'\}] &= -\frac{\kappa}{2}\rho_s wk && (2.21)\\ \left[\dot{N} + 2N'l\dot{\alpha} + \dot{\alpha}^2 l^2 N(\dot{l}/l)\right] &= \kappa\left(\frac{1}{2}\rho_s + p_s\right)k^3. && (2.22)\end{aligned}$$

The Gauss-Codazzi equations relate the curvature at one point to the extrinsic and intrinsic curvature of a hypersurface which passes through this point. Let us first define coordinate basis vectors tangential to the junction surface by

$$e_a{}^{\mu} \stackrel{\text{def}}{=} \frac{\partial x^{\mu}}{\partial \xi^a},$$

where ξ^a are the coordinates covering the junction surface (in our case t, θ, ϕ). The Gauss-Codazzi equations are then given by

$$R_{\mu\nu\lambda\kappa}e_a{}^{\mu}e_b{}^{\nu}e_c{}^{\lambda}e_d{}^{\kappa} = {}^{(3)}R_{abcd} + \epsilon(K_{ad}K_{bc} - K_{ac}K_{bd})$$

and

$$R_{\mu\nu\lambda\kappa}n^{\mu}e_b{}^{\nu}e_c{}^{\lambda}e_d{}^{\kappa} = K_{bc|d} - K_{bd|c},$$

where a vertical bar denotes the covariant derivative with respect to the induced hypersurface metric $h_{\mu\nu} = g_{\mu\nu} - n_{\mu}n_{\nu}$, and $\epsilon \stackrel{\text{def}}{=} n^{\mu}n_{\mu}$ equals $+1$ for timelike hypersurfaces, and -1 for spacelike hypersurfaces. For a timelike hypersurface these equations lead to [12]

$$2G_{\mu\nu}n^{\mu}n^{\mu} = -{}^{(3)}R + K^2 - K_{\mu\nu}K^{\mu\nu} \tag{2.23}$$

$$G_{\mu\nu}e_a{}^{\mu}n^{\nu} = K_a{}^b{}_{|b} - K_{|a}, \tag{2.24}$$

where $G_{\mu\nu} \stackrel{\text{def}}{=} R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is the Einstein tensor. Substituting (2.12) into (2.23) and (2.24) gives

$$[G_{\mu\nu}e_a{}^{\mu}n^{\nu}] = -\kappa S_a{}^b{}_{|b} \tag{2.25}$$

$$[G_{\mu\nu}n^{\mu}n^{\nu}] = \frac{1}{2}[K^2 - K_{\mu\nu}K^{\mu\nu}] = \kappa S^{\mu\nu}\bar{K}_{\mu\nu}, \tag{2.26}$$

where $\bar{K}_{\mu\nu} \stackrel{\text{def}}{=} \frac{1}{2}(\lim_{R\to 1^+} K_{\mu\nu} + \lim_{R\to 1^-} K_{\mu\nu})$. The second of these equations represents nothing else than the definition of the surface stress-energy tensor $S^{\mu\nu}$ (which was substituted) and hence is redundant. Using (2.20) to evaluate (2.25) we find that the only non-vanishing component is the time component, which is given by (using $e_0{}^{\mu} = u^{\mu}$)

$$ku^{\mu}\rho_{s;\mu} + k(\rho_s + p_s)u^a{}_{|a} = \kappa^{-1}[G_{\mu\nu}u^{\mu}n^{\nu}] = [T_{\mu\nu}u^{\mu}n^{\nu}], \tag{2.27}$$

where the last equality follows from the Einstein equation and $u^a{}_{|a} = u^{\mu}(2\ln(lf))_{;\mu}$. Together with an equation of state for the surface energy and pressure densities this equation describes the evolution of the 'matter' on the surface.

On the other hand (2.26) becomes after substituting (2.20) and (2.16)

$$-(\rho_s + p_s)\overline{n_{\mu}u^{\mu}{}_{;\nu}u^{\nu}} + p_s\bar{K}. = \kappa^{-1}[G_{\mu\nu}n^{\mu}n^{\nu}] = [T_{\mu\nu}n^{\mu}n^{\nu} - \Lambda/\kappa].$$

2.2.4 Matching with surface-layer

The second geometric matching condition shows that $k \stackrel{\text{def}}{=} N\sqrt{1 - \dot{\alpha}^2 l^2}$ is continuous across the junction surface. Hence we can express the junction surface motion in terms of the lapse function by

$$\dot{\alpha}l = \pm\sqrt{1 - \left(\frac{k}{N}\right)^2}. \tag{2.28}$$

For convenience we will choose now $k = 1$, i.e., the coordinate time corresponds to proper time along the curves $R = 1, \theta, \phi$ constant on the junction surface. Setting $u \stackrel{\text{def}}{=} N/k$ and $j_{\pm} \stackrel{\text{def}}{=} \text{sign}(\dot{\alpha}_{\pm})$ the angular component of the extrinsic curvature (2.19) on both sides of the junction surface and equation (2.9), the derivative of the junction surface radius with respect to coordinate time, are given by

$$\begin{aligned} K^{\theta}{}_{\theta} &= (uw' + j\dot{w}\sqrt{u^2 - 1})/w && (2.29) \\ L &\stackrel{\text{def}}{=} \frac{\mathrm{d}w}{\mathrm{d}t} = u\dot{w} + jw'\sqrt{u^2 - 1} && (2.30) \end{aligned}$$

A valid matching between two spherically symmetric sections, each satisfying the Einstein-field equations, must satisfy the two geometric matching conditions ((2.4) and (2.5)) and the two independent components of the Lanczos equation ((2.21) and (2.22)). We first note, that if the first geometric matching condition (matching of the surface radius) (2.4) is satisfied initially, then it is sufficient to demand that its coordinate time derivative (2.9) is satisfied at all times. Secondly, with our choice of variables the second geometric matching condition (2.5) is nothing more than an identity — with (2.28) it has already been used to eliminate one variable. Thirdly, as will be shown in section 2.4, equation (2.22), the time-component of the Lanczos equation, is in fact identically satisfied if all the other matching conditions are satisfied, the Einstein-field equations are valid on each side, and the surface-matter evolution is given by (2.27).

We conclude that the matching conditions are completely represented by (2.9), the coordinate time derivative of the first geometric matching condition, and the angular component of the Lanczos equation (2.21) together with an initial matching of the proper surface radius $w = lf$. With our choice of variables these equations take the (surprisingly symmetric) form

$$\begin{aligned} [L] &= [u\dot{w} + jw'\sqrt{u^2-1}] = 0 && (2.31)\\ [wK^\theta{}_\theta] &= [uw' + j\dot{w}\sqrt{u^2-1}] = -E, && (2.32) \end{aligned}$$

where $w \stackrel{\text{def}}{=} lf$ is the proper surface radius of the spherical junction surface on each side and

$$E \stackrel{\text{def}}{=} \frac{\kappa\rho_s w}{2} \tag{2.33}$$

quantifies the energy-content of the layer. To find a relation between $K^\theta{}_\theta$ and L we square (2.29) and substitute L^2 from the square of (2.30) and obtain

$$(wK^\theta{}_\theta)^2 = L^2 + a, \tag{2.34}$$

where $a \stackrel{\text{def}}{=} w'^2 - \dot{w}^2$. Versions of this equation have been given in [9] and [13]. For $E \neq 0$ we can express $K^\theta{}_\theta$ in terms of L by using an algebraic identity as

$$wK^\theta{}_\theta = \frac{[(wK^\theta{}_\theta)^2] \pm [wK^\theta{}_\theta]^2}{2[wK^\theta{}_\theta]} = \frac{b \pm E^2}{-2E}, \tag{2.35}$$

where $b \stackrel{\text{def}}{=} a_+ - a_-$. The explicit expression for L in terms of E takes the form

$$L^2 = \left(\frac{b-E^2}{2E}\right)^2 - a_- = \left(\frac{b+E^2}{2E}\right)^2 - a_+ = (E^4 - 2E^2(a_+ + a_-) + b^2)/4E^2, \tag{2.36}$$

Note that we find by differentiating (2.29) and (2.30) with respect to u (taking $w, \dot{w}$, and w' to be independent of u) the helpful relations

$$j\sqrt{u^2-1}\frac{\partial L}{\partial u} = wK^\theta{}_\theta \quad \text{and} \quad j\sqrt{u^2-1}\frac{\partial wK^\theta{}_\theta}{\partial u} = L, \tag{2.37}$$

which are valid on each side of the junction surface.

The geometric matching condition (2.30) can be solved for $u_\pm$ and $j_\pm$ in terms of the time derivative of the surface radius L and the extrinsic curvature component $K^\theta{}_\theta$

$$u = \frac{\dot{w}L - w'wK^\theta{}_\theta}{-a}. \tag{2.38}$$

We note that differentiating (2.38) with respect to u and using (2.37) yields

$$j\sqrt{u^2-1} = \frac{w'L - \dot{w}wK^\theta{}_\theta}{a}, \tag{2.39}$$

what also determines the sign of j and hence the radial direction of motion of the junction surface for each side.

2.2.5 The 'no surface-layer' case

If all tangential components of the extrinsic curvature are continuous at the junction surface then it follows from (2.32) that the surface-energy density ρ_s vanishes. In this case the junction hypersurface is called a boundary-surface.

It is an immediate consequence of (2.34) that in this case

$$b = [a] = [w'^2 - \dot{w}^2] = 0.$$

Feasible solutions need to satisfy the two geometric matching conditions (2.4) and (2.5), the derivative of the first matching condition (2.9) and the matching of the extrinsic curvature (2.21). Recognizing the similar structure of (2.9) and (2.21) we form two new equivalent equations by adding and subtracting the two equations. The result reads

$$[N(1-\dot{\alpha}l)(w' - \dot{w})] = 0 \qquad [N(1+\dot{\alpha}l)(w' + \dot{w})] = 0.$$

Using the factorized form of the second geometric matching condition (2.5) and defining $q = N(1-\dot{\alpha}l)$ this becomes

$$[q(w' - \dot{w})] = 0 \qquad [\frac{1}{q}(w' + \dot{w})] = 0.$$

If both $w' - \dot{w}$ and $w' + \dot{w}$ vanish separately on both sides, then the system becomes an identity and the junction surface motion remains undefined. Let us assume now that this is not the case.

We note that for $[w'] = [\dot{w}] = 0$ the system is solved for any $q_+ = q_-$. If the angular component of the metric and its first order proper time and radial derivatives are continuous, then the junction surface motion does not follow from the matching conditions. In particular, this is the case for the trivial matching of two identical space-times, were we have an 'imaginary junction surface', which could be placed anywhere.

Let us from now on assume that at least one of the proper derivatives of w is not continuous at the junction surface. It is easy to see from above system that there are no solutions to the system if $w' \pm \dot{w} \neq 0$ on both sides.

If $w' \pm \dot{w}$ is zero on one side, it has to be zero on the other side too (otherwise no matching is possible) and the solutions have to satisfy

$$\frac{N_-\sqrt{1-\dot{\alpha}_-^2 l_-^2}}{N_+\sqrt{1-\dot{\alpha}_+^2 l_+^2}} = \frac{w'_\pm}{w'_\mp}.$$

We conclude that a matching without surface-layer is only possible for some exceptional situations. The junction-motion is only defined if either $w' + \dot{w} = 0$ or $w' - \dot{w} = 0$ (but not both) on each sides. If w' and $\dot{w}$ are continuous at the junction surface then the motion remains undefined – 'there is no real junction'.

2.2.6 Expansion for small surface-energy densities

As will be seen in the numerical examples given later, in many cases the surface-energy density (and hence E) approaches zero at some finite coordinate time. In this case the dynamic quantities can be approximated by a series expansion in terms of E. We start by re-writing the exact expression for the extrinsic curvature (2.35) as

$$wK^{\theta}_{\pm\theta} = -\frac{b}{2E} \mp \frac{E}{2}.$$

It follows then from (2.36) that

$$\pm L = \frac{|b|}{2E} - \frac{a_+ + a_-}{2|b|}E + O(E^3),$$

where $O(E^3)$ represents terms of the order E^3 or smaller. Furthermore, from (2.38) we find for $u = N/k$ the expansion

$$u_\pm = -\frac{w'_\pm b + \operatorname{sign}(L)\dot{w}_\pm |b|}{2a_\pm}\frac{1}{E} + \frac{\operatorname{sign}(L)\dot{w}_\pm(a_+ + a_-)/|b| \mp w'_\pm}{2a_\pm}E + O(E^3).$$

As the surface-energy density approaches zero the lapse functions (given by u_+ and u_-) diverge and the proper speed of the junction surface approaches the speed of light quadratically since

$$|\dot{\alpha}_\pm l_\pm| = 1 - \frac{2\alpha_\pm l_\pm^2}{w'_\pm b + \operatorname{sign}(L)\dot{w}_\pm|b|}E^2 + O(E^4).$$

To examine if and how E approaches zero we finally expand the evolution equation (2.27)

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\kappa}{2}\left(w[T_{\mu\nu}u^\mu n^\nu] - \frac{|b|}{w}\left(\gamma_s - \frac{1}{2}\right)\right) + O(E^2).$$

Generally the first term diverges as we approach the speed of light — for example in the case of a perfect fluid one finds

$$|T_{\mu\nu}u^\mu n^\nu = \left(\frac{w'b + \operatorname{sign}(L)\dot{w}|b|}{2a}\right)^2 (\rho + p)\frac{1}{E^2} + O(E^0).$$

We conclude that if the energy-momentum contribution $[T_{\mu\nu}u^\mu n^\nu]$ has a sign opposite to E, then E accelerates towards zero. In many cases E will reach zero at some finite coordinate time t_0. Close to this point and assuming that the time dependence of all other terms is negligible we have $E \propto (t_0 - t)^{1/3}$. This implies that the lapse functions are integrable and the junction surface reaches the speed of light (on each side) within a finite proper time. Here our formalism breaks down and one would need a separate treatment of these singular cases. We want to speculate here that at these points the junction surface turns spacelike.

On the other hand, if the sign of $[T_{\mu\nu}u^\mu n^\nu]$ is the same as the sign of E then E cannot get arbitrarily close to zero. In some cases (see figure 2.11) E will oscillate around some value (which is itself time dependent). Even in these cases we can encounter divergencies resulting from diverging $a_\pm$ and b. This can lead to non-integrable lapse functions - from each side the junction seems to exist forever, but an observer who moves along the junction encounters a singular point after a finite time. At this point the surface energy density is zero and again the formalism breaks down.

It should be noted that in the case of a perfect fluid on both sides of the junction the sign of the stress-energy contribution $[T_{\mu\nu}u^\mu n^\nu]$ depends on the energy density ρ and pressure p on both sides, i.e., on the equations of state. Hence whether a particular junction reaches the speed of light within a finite time or not might depend on the equation of state on each side. In section 2.6 we give a numerical example for such a case.

Because points on the junction surface are not causally connected a spacelike junction surface has a very different physical interpretation . In such cases the junction surface cannot be treated as an 'evolving system' on its own, but rather as some kind of (spacelike) transition surface which is generated by the physics underlying the cosmological model.

Usually a timelike junction surface is used to model the time evolution of a spatially *localized* inhomogeneity. If a junction surface turns spacelike a breakdown in the thin wall approximation must have occurred.

2.3 Necessary and sufficient conditions for a possible matching

2.3.1 Demanding real solutions for L

From (2.36) we find with $L^2 \geq 0$ a necessary condition for the existence of solutions which restricts the allowed values for E, such that

$$E^4 - 2E^2(a_+ + a_-) + b^2 \geq 0. \qquad (2.40)$$

The roots of the quadratic polynomial (in E^2) on the left-hand side are given by $a_+ + a_- \pm 2\sqrt{a_+a_-}$. For $a_+a_- < 0$ the quadratic has no root and all values of E are feasible. If on the other hand, both a_+ and a_- are negative then (2.40) has only negative roots for E^2 and hence all real values of E are feasible. For $a_+, a_- \geq 0$ the roots become $|\sqrt{a}_+ \pm \sqrt{a}_-|$ and the feasible values for E are

$$0 \leq |E| \leq |\sqrt{a_+} - \sqrt{a_-}| \quad \text{or} \quad \sqrt{a_+} + \sqrt{a}_- \leq |E|. \qquad (2.41)$$

The shape of the forbidden region in the $E - a_+ - a_-$ and $E - \sqrt{a_+}$ space is illustrated in figures 2.4 and 2.5, respectively.

For $a_+, a_- \geq 0$ it is easily verified that

$$0 \leq |\sqrt{a_+} - \sqrt{a_-}| \leq \sqrt{|a_+ - a_-|} \leq \sqrt{a_+} + \sqrt{a_-}$$

and hence the two disjoint regions allowed for E given by (2.41) are easily distinguished by $E^2 \leq |b|$ and $E^2 \geq |b|$. Furthermore, from (2.35) we find $\mathrm{sign}(K^{\theta}_{-\theta}) = \mathrm{sign}(K^{\theta}_{+\theta}) = -\mathrm{sign}(b/E)$ for $E^2 < |b|$ and $\mathrm{sign}(K^{\theta}_{-\theta}) = \mathrm{sign}(E), \mathrm{sign}(K^{\theta}_{+\theta}) = -\mathrm{sign}(E)$ for $E^2 > |b|$.

2.3.2 Proper time relations

By setting $k = 1$ it follows from (2.6) the condition

$$N = u \geq +1, \qquad (2.42)$$

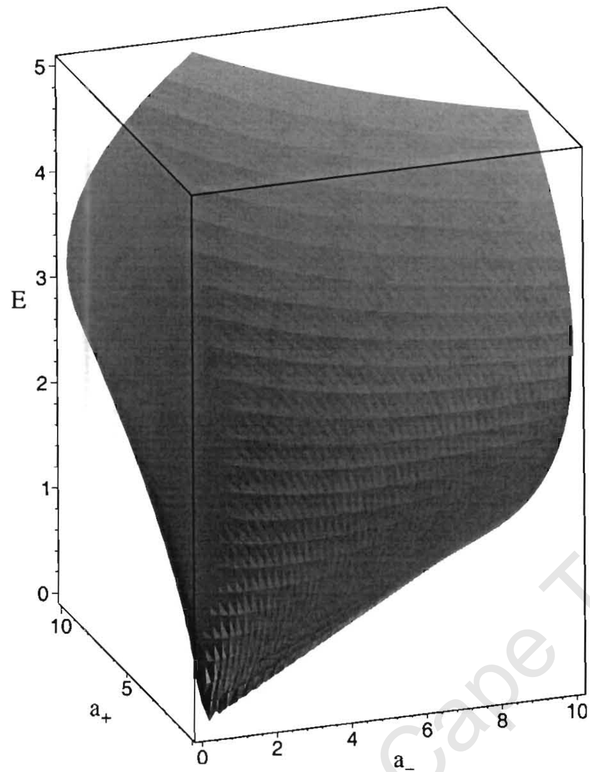


Figure 2.4: The shaded volume represents the region in the $E - a_+ - a_-$-space where there is no solution for the junction surface motion.

i.e., on each side of the junction surface proper time must proceed faster (with respect to the time coordinate) than on the junction surface. To investigate the resulting constraints on the surface energy density we start by noting that

$$|w'(wK^\theta{}_\theta) - a| - |\dot{w}L| \begin{cases} \geq 0 \text{ for } a \geq 0 \\ \leq 0 \text{ for } a \leq 0 \end{cases}, \tag{2.43}$$

which can be easily verified by squaring and substituting from (2.34). Substituting (2.38) the inequality (2.42) takes the form

$$\frac{(w'(wK^\theta{}_\theta) - a) - \dot{w}L}{a} \geq 0.$$

It follows from (2.43) that for $a > 0$ we need

$$w'(wK^\theta{}_\theta) - a \geq 0,$$

while for $a < 0$

$$\dot{w}L \geq 0. \tag{2.44}$$

Let us first consider the case $a > 0$. Using (2.35) the condition becomes

$$w'\frac{b + \sigma E^2}{-2E} - a \geq 0,$$

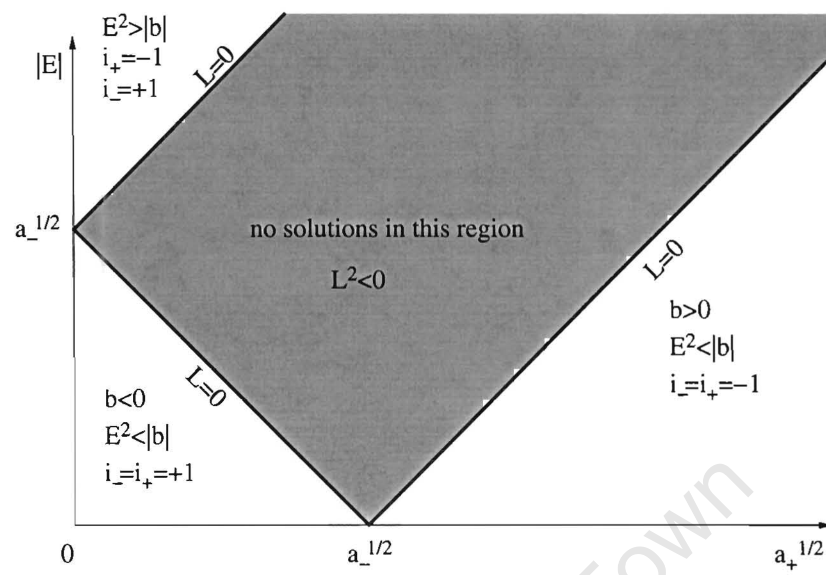


Figure 2.5: Region in the $E - \sqrt{a_+}$-plane where there is no solution for the junction surface motion for positive a_+ and a_-. On the bold line we have $E = |\sqrt{a_+} \pm \sqrt{a_-}|$ and hence $L = 0$. Each of the three different allowed regions is distinguished by the signs of the extrinsic curvature and the surface energy density $i_\pm \stackrel{\text{def}}{=} \text{sign}(K^\theta_{\pm\theta}) \cdot \text{sign}(E)$.

where $\sigma = \pm 1$ corresponding to the outside (+) and inside (−) case. In the following let us use the convention that if a, w refer to the quantities on one side, then a_*, w_* refer to the quantities on the other side of the junction. By setting $x \stackrel{\text{def}}{=} \text{sign}(b)E/\sqrt{|b|}$, $\epsilon_\pm \stackrel{\text{def}}{=} \pm\text{sign}(b) = a - a_*$, and

$$s_\pm \stackrel{\text{def}}{=} -\frac{a_\pm}{w'_\pm\sqrt{|b|}}$$

we bring the inequality in the form

$$\frac{1}{x} + \epsilon_\pm x \begin{cases} \leq 2s_\pm \text{ for } w'_\pm > 0 \\ \geq 2s_\pm \text{ for } w'_\pm < 0 \end{cases}. \tag{2.45}$$

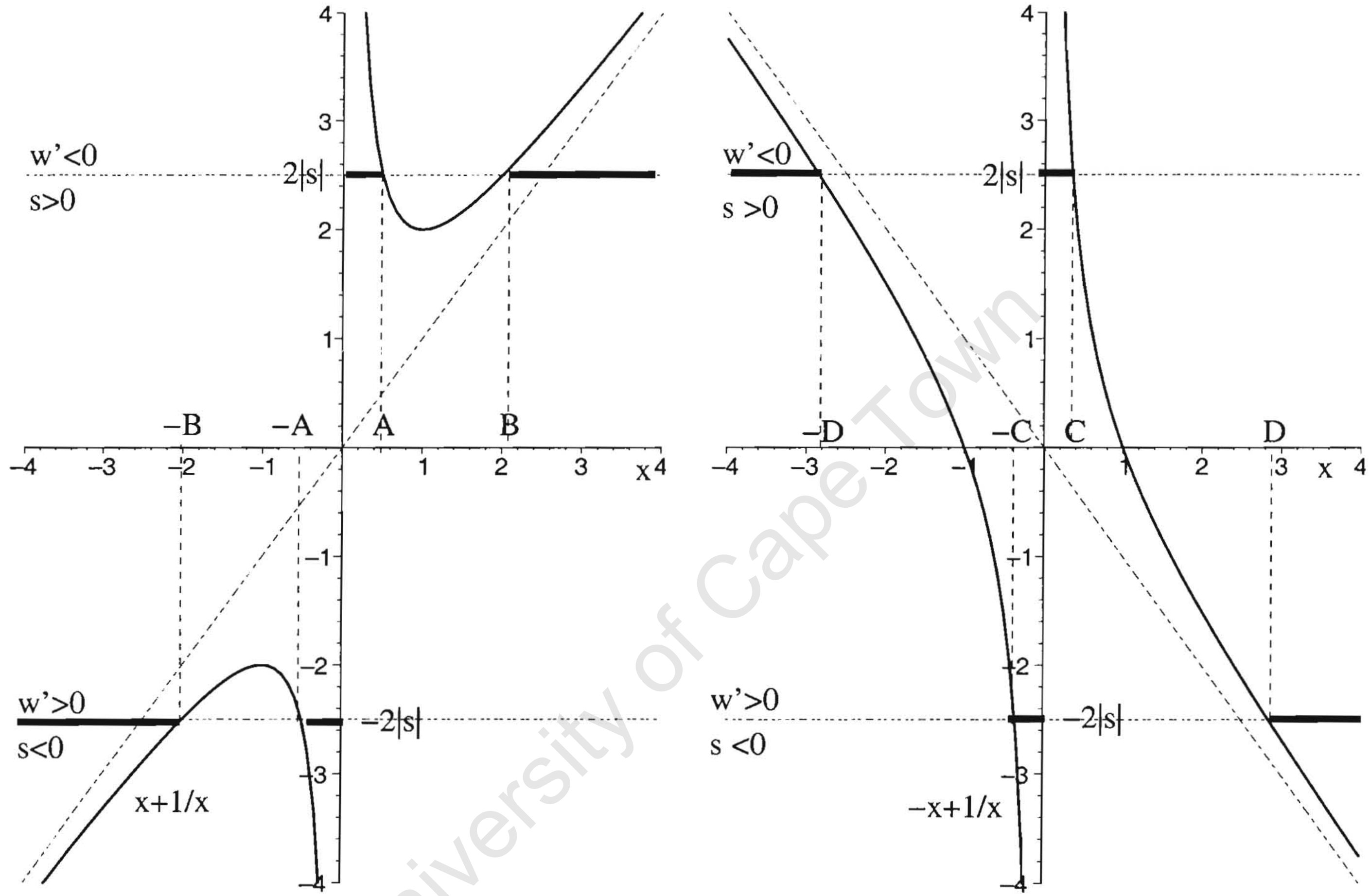
The allowed ranges for x are illustrated in figure 2.6.

The case of $\epsilon > 0$ Let us note that if $\epsilon_+ > 0$ then $\epsilon_- < 0$ and vice versa ($\epsilon_* < 0$). The sign of the surface energy density is now determined by

$$\text{sign}(\rho_s) = -\text{sign}(b)\text{sign}(w'). \tag{2.46}$$

Furthermore, if $|s| = |a/(w'\sqrt{|b|})| \leq 1$ then no restrictions are placed on $|x|$ (but further restrictions could come from $a_* > 0$). For $|s| > 1$ the allowed range can be found by setting $|x| = e^z$ and hence $|1/x + x| = 2\cosh(z)$. Using $\cosh^{-1}|s| = \ln(|s| + \sqrt{s^2 - 1})$ one obtains the ranges

$$0 < |x| \leq |s| - \sqrt{s^2 - 1} \quad \text{or} \quad |s| + \sqrt{s^2 - 1} \leq |x|.$$



(a) Allowed ranges for E if $\epsilon > 0$.

(b) Allowed ranges for E if $\epsilon < 0$.

Figure 2.6: Visualization of the inequality (2.45) for $\epsilon > 0$ (figure 2.6(a)) and $\epsilon < 0$ (figure 2.6(b)). The lower and upper horizontal line correspond to $w'_\pm > 0$ and $w'_\pm < 0$, respectively. Here $s_\pm \stackrel{\text{def}}{=} -a_\pm/(w'_\pm\sqrt{|b|})$ are the terms on the right-hand side of (2.45). The allowed ranges are indicated by a bold line. The end-points are given by (A and B are only defined for $|s| > 1$) $A = |s| - \sqrt{s^2-1}, B = |s| + \sqrt{s^2-1}, C = -|s| + \sqrt{s^2+1}, D = |s| + \sqrt{s^2+1}$.

The case of $\epsilon < 0$ Similarly to the last case, if $\epsilon < 0$ then $\epsilon_* > 0$. Hence each case will occur once at the junction. A similar procedure as above yields the restrictions

$$|s| - \sqrt{s^2+1} \le \mathrm{sign}(w')x < 0 \quad \text{or} \quad |s| + \sqrt{s^2+1} \le \mathrm{sign}(w')x.$$

The case of $a > 0$ on both sides If a is positive on both sides then x can only take values which lay in the intersection of the allowed ranges on each side. The allowed intervals differ, depending on the signs of w'_+ and w'_-. All possible cases are shown in table 2.1. We note that in every case the allowed values for E have the same sign.

As a particularly important case (for the matching of FLRW models) and an illustrative example we evaluate the restrictions on the surface energy density for $a_+, a_- > 0$ and $w'_+, w'_- > 0$. In this case x has to be negative and hence

$$\mathrm{sign}(\rho_s) = \mathrm{sign}(E) = -\mathrm{sign}(b).$$

If one assumes on physical grounds that ρ_s should be positive then no matching will be possible if b is positive.

The case of $a < 0$ If $a < 0$ on one side of the junction surface then (2.44) implies with

$$\mathrm{sign}(L) = \mathrm{sign}(\dot{w})$$

the sign for L, the coordinate time derivative of the surface radius, which was left undefined in (2.36). For the case that a_+ and a_- are negative this condition must hold on both sides and hence a matching is only possible if

$$\mathrm{sign}(\dot{w}_+) = \mathrm{sign}(\dot{w}_-) \quad \text{for } a_+, a_- < 0. \tag{2.47}$$

2.4 The time-component of the Lanczos equation

So far we have only considered matching of the metric and of the angular components of the extrinsic curvature of the junction surface. The remaining matching condition comes from the time-component of the extrinsic curvature (2.22), which contains a second order time derivative of the junction coordinate radius, or equivalently a first-order time derivative of the lapse function.

Rewriting the time-component of the extrinsic curvature in terms of our variable $u = N/k$ yields

$$-K^{\mu\nu}u_\mu u_\nu = \frac{j}{\sqrt{u^2-1}}\frac{\mathrm{d}u}{\mathrm{d}t} + j\sqrt{u^2-1}\frac{\dot{l}}{l} + u\frac{N'}{N}, \tag{2.48}$$

where the factor N'/N is independent of the junction surface motion. Taking the coordinate time derivative of (2.35) and using (2.29) and (2.37) we obtain

$$\frac{j_\pm L}{\sqrt{u_\pm^2-1}}\frac{\mathrm{d}u_\pm}{\mathrm{d}t} = -\frac{1}{2E}\frac{\mathrm{d}b}{\mathrm{d}t} - \frac{\mathrm{d}E}{\mathrm{d}t}\frac{1}{E}wK^{\theta}_{\mp\theta} - z, \tag{2.49}$$

	$b > 0$ $\Rightarrow \epsilon_+ = +1, \epsilon_- = -1$	$b < 0$ $\Rightarrow \epsilon_+ = -1, \epsilon_- = +1$
$w'_+ > 0$ $w'_- > 0$	$\mathrm{sign}(\rho_s) = -1$ $\sqrt{\lvert b\rvert}\max(-A_+, -C_-) \leq E < 0$ for $\lvert s_+\rvert > 1$ $-\sqrt{\lvert b\rvert}C_- \leq E < 0$ for $\lvert s_+\rvert \leq 1$	$\mathrm{sign}(\rho_s) = +1$ $0 < E \leq \sqrt{\lvert b\rvert}\min(A_-, C_+)$ for $\lvert s_+\rvert > 1$ $0 < E \leq \sqrt{\lvert b\rvert}C_+$ for $\lvert s_+\rvert \leq 1$
$w'_+ > 0$ $w'_- < 0$	$\mathrm{sign}(\rho_s) = -1$ $E \leq \sqrt{\lvert b\rvert}\min(-B_+, -D_-)$ for $\lvert s_+\rvert > 1$ $E \leq -\sqrt{\lvert b\rvert}D_-$ for $\lvert s_+\rvert \leq 1$	$\mathrm{sign}(\rho_s) = -1$ $E \leq \sqrt{\lvert b\rvert}\min(-B_-, -D_+)$ for $\lvert s_-\rvert > 1$ $E \leq -\sqrt{\lvert b\rvert}D_+$ for $\lvert s_-\rvert \leq 1$
$w'_+ < 0$ $w'_- > 0$	$\mathrm{sign}(\rho_s) = +1$ $\sqrt{\lvert b\rvert}\max(B_+, D_-) \leq E$ for $\lvert s_+\rvert > 1$ $\sqrt{\lvert b\rvert}D_- \leq E$ for $\lvert s_+\rvert \leq 1$	$\mathrm{sign}(\rho_s) = +1$ $\sqrt{\lvert b\rvert}\max(B_-, D_+) \leq E$ for $\lvert s_-\rvert > 1$ $\sqrt{\lvert b\rvert}D_+ \leq E$ for $\lvert s_-\rvert \leq 1$
$w'_+ < 0$ $w'_- < 0$	$\mathrm{sign}(\rho_s) = +1$ $0 < E \leq \sqrt{\lvert b\rvert}\min(A_+, C_-)$ for $\lvert s_+\rvert > 1$ $0 < E \leq \sqrt{\lvert b\rvert}C_-$ for $\lvert s_+\rvert \leq 1$	$\mathrm{sign}(\rho_s) = -1$ $\sqrt{\lvert b\rvert}\max(-A_-, -C_+) \leq E < 0$ for $\lvert s_-\rvert > 1$ $-\sqrt{\lvert b\rvert}C_+ \leq E < 0$ for $\lvert s_-\rvert \leq 1$

Table 2.1: For $a_+ > 0$ and $a_- > 0$ this table shows the allowed region for E for all possible combinations of $\mathrm{sign}(w'_+)$ and $\mathrm{sign}(w'_-)$. Here $A_\pm = |s_\pm| - \sqrt{s_\pm^2 - 1} \leq 1; B_\pm = |s_\pm| + \sqrt{s_\pm^2 - 1} \geq 1; C_\pm = -|s_\pm| + \sqrt{s_\pm^2 + 1} \leq 1; D_\pm = |s_\pm| + \sqrt{s_\pm^2 + 1} \geq 1$. and $s_\pm = -a_\pm/(w'_\pm\sqrt{|b|})$.

where (note that $(\mathrm{d}/\mathrm{d}t)f(t, \alpha(t)R) = u\dot{f} + j\sqrt{u^2-1}f'$)

$$z \stackrel{\text{def}}{=} u\frac{\mathrm{d}w'}{\mathrm{d}t} + j\sqrt{u^2-1}\frac{\mathrm{d}\dot{w}}{\mathrm{d}t} = ju\sqrt{u^2-1}\{w'' + \ddot{w}\} + u^2(w')^\bullet + (u^2-1)(\dot{w})'.$$

Differentiating (2.33) and using (2.27) with $u^a{}_{|a} = 2L/w$ yields

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\kappa L\left(\frac{\rho_s}{2} + p_s\right) + \kappa\frac{w}{2}[T_{\mu\nu}u^\mu n^\nu],$$

which expresses the coordinate time derivative of E. Substituting for the first term in (2.48) allows us to evaluate the remaining junction condition $[K^{\mu\nu}u_\mu u_\nu] = -\kappa(\rho_s/2 + p_s)$. The terms containing the surface pressure and density cancel each other and we obtain

$$0 = \frac{\kappa}{2}w[T^{\mu\nu}u_\mu n_\nu] + [z] - L\left[j\sqrt{u^2-1}\frac{\dot{l}}{l}\right] - L\left[u\frac{N'}{N}\right]. \tag{2.50}$$

The first term can be expressed in terms of the Einstein-tensor with respect to the original metric (2.1) as

$$\kappa T_{\mu\nu}u^\mu n^\nu = G_{\mu\nu}u^\mu n^\nu = ju\sqrt{u^2-1}\left(\frac{G_{tt}}{N^2} + \frac{G_{rr}}{l^2}\right) + (2u^2-1)\left(\frac{G_{tr}}{Nl}\right),$$

where the relevant components of the Einstein-tensor are given by

$$\begin{aligned}
G_{tt} &= \frac{2N^2}{w}\left(-w'' + \dot{w}\frac{\dot{l}}{l} + \frac{1-a}{2w}\right)\\
G_{rr} &= \frac{2l^2}{w}\left(\ddot{w} - w'\frac{N'}{N} + \frac{1-a}{2w}\right)\\
G_{tr} &= \frac{2Nl}{w}\left(\dot{w}\frac{N'}{N} - (w')^\bullet\right),
\end{aligned}$$

and L is given by (2.30). Substituting into (2.50) and using the relation

$$(\dot{w})' - (w')^\bullet = \frac{\dot{l}}{l}w' - \frac{N'}{N}\dot{w}$$

shows that (2.50) is an identity, satisfied *for all* spherically symmetric junctions between solutions of the Einstein-field equations if the geometric matching conditions ((2.4) and (2.5)) together with the angular component of the Lanczos equation (2.32) are satisfied. While it was well-known that for certain cases the time-component of the Lanczos equation is identically satisfied (e.g. [8, 11]), it seems to be a new result for the generic spherically symmetric case.

It was suggested that for the matching of FLRW models the time-component of the Lanczos equation determines the pressure [9]. In light of the above result this cannot be the case and one needs to supplement the model with an equation of state for the surface-matter.

2.5 Matching of FLRW sections

We want to turn our attention now to the special case of the matching of two distinct FLRW regions. Such junctions are encountered in cosmological models which approximate universes containing many FLRW domains (multidomain universes). The most prominent example is Linde's Chaotic Inflation scenario [3, 4].

Junctions of this type have been studied in [9, 13]. Our treatment will serve as an illustration for the introduced method and as a source for numerical examples. Here it is not our aim to investigate physical processes which could lead the to creation of a "bubble" and we refer the interested reader to the vast literature (see, e.g., [9, 14, 15, 16, 17]). Instead we want to focus on the generic geometrical and mathematical aspects.

2.5.1 FLRW models and their parametrization

The metric of FLRW models can be written in the form

$$\mathrm{d}s^2 = -N^2(t)\mathrm{d}t^2 + l^2(t)\{\mathrm{d}r^2 + f^2(r)\mathrm{d}\Omega^2\}, \tag{2.51}$$

where $l(t)$ is the scale factor, $N(t)$ the so-called lapse function, $\mathrm{d}\Omega$ the line-element on the two-dimensional unit-sphere, and

$$f(r) = \begin{cases} \sin(r) & \text{for closed models} \\ r & \text{for flat models} \\ \sinh(r) & \text{for open models} \end{cases} .$$

Note that the FLRW metric (2.51) has the same form as the general metric for spherical symmetric spaces (2.1), but with $l' = 0, N'/N = 0$ and $\dot{f} = 0$.

The evolution of FLRW models is described by the Friedmann equation - the dynamic part of the Einstein-Field equations -

$$\left(\frac{\dot{l}}{l}\right)^2 - \frac{\kappa\rho + \Lambda}{3} = -\frac{\zeta}{l^2}, \tag{2.52}$$

where $\zeta = 0, +1, -1$ for flat, closed, and open models, respectively, and a dot indicates the derivative with respect to *proper* time t, *i.e.*, $\dot{l} \stackrel{\text{def}}{=} \frac{1}{N}\frac{\mathrm{d}l}{\mathrm{d}t}$.

The matter is described by an energy-momentum tensor of perfect fluid type. The unit tangent vectors to the fluid flow lines are given by

$$v^\mu = \frac{1}{N}\delta^\mu_t - \frac{\dot{\alpha}R}{\alpha}\delta^\mu_R,$$

where we used the coordinates introduced in subsection 2.2.1, and the energy-momentum tensor takes the form

$$T^{\mu\nu} = (\rho + p)v^\mu v^\nu + pg^{\mu\nu},$$

where ρ is the energy density and p the pressure. The matter evolution is then described by the energy-conservation equation

$$\dot{\rho} + (\rho + p)3H = 0,$$

where $H \stackrel{\text{def}}{=} \dot{l}/l$ is the Hubble parameter.

We restrict ourself to models with a γ-law equation of state, i.e., models in which energy density ρ and pressure p are related by[2]

$$p = (\gamma - 1)\rho \qquad \gamma \in (2/3, 2].$$

In this case $\chi_\gamma \stackrel{\text{def}}{=} \frac{\kappa}{3}\rho l^{3\gamma}$ is a constant of motion proportional to the entropy per coordinate volume. This allows us to eliminate the energy density ρ from the Friedmann equation, so that the evolution of the scale factor l is described in terms of the constants of motion by

$$H^2 = \chi_\gamma l^{-3\gamma} + \frac{1}{3}\Lambda - \frac{\zeta}{l^2}. \tag{2.53}$$

Taking the proper-time derivative of this equation we recover the Raychaudhuri equation, which takes the form

$$\ddot{l} = \frac{2-3\gamma}{2}\chi_\gamma l^{-3\gamma+1} + \frac{\Lambda}{3}l \tag{2.54}$$

for $\dot{l} \neq 0$.

2.5.2 Modelling transition regions

Often the space-time is modelled by two almost-FLRW sections which are continuously joined through a transition region as illustrated in figure 2.7(a). Each FLRW section is characterized by its constants of motion.

The non-FLRW region which constitutes the transition region allows for the change of these constants from one side to the other.

Even under the assumption of spherical symmetry such models are of a rather complex nature — there are many possible choices for the geometry of the transition region, which should asymptotically approach the FLRW geometry on each side. Finding an exact solution to the Einstein field equation which models a particular transition region seems impossible.

Assuming that the thickness of the transition region is small compared to the length scale of the FLRW regions (Hubble length) one can model such a transition region by a junction surface as shown in figure 2.7(b). The surface-energy density then represents the 'collective effect' of the non-FLRW region.

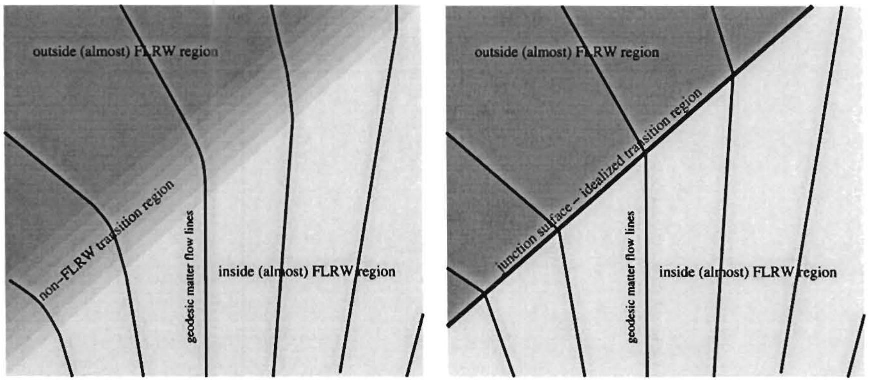
2.5.3 Identification of models

It will be of interest whether there are different models with the same evolution of the Hubble parameter H. For this purpose let us only consider models with a γ-law equation of state.

Let us first note that the Hubble parameter can be written as $H = \frac{\dot{l}}{l} = (\ln(l))^\bullet$. Hence models with the same evolution of H are models whose scale factors differ by a *constant* factor.

Let us assume we are given a solution to the Friedmann equation with $\gamma = \gamma_-, \chi = \chi_-$, $\Lambda = \Lambda_-$, and $\zeta = \zeta_-$. The question is now, whether there are constants γ_+, Λ_+, $\chi_{\gamma+}$ and ζ_+ such that $l_+ = \lambda l_-$ is a solution for constant

[2] The case $\gamma = 0$ gives an effective cosmological constant. We can exclude this case here because the cosmological constant is included separately.



(a) Two (almost) FLRW regions are joined by a 'thin' non-FLRW region.

(b) The non-FLRW region is approximated as a surface.

Figure 2.7: Approximating a transition region by a junction surface simplifies the treatment.

$\lambda > 0$. Both solutions would have the same Hubble parameter $H \stackrel{\text{def}}{=} \frac{\dot{l}_+}{l_+} = \frac{\dot{l}_-}{l_-}$. Substituting in each case from the Friedmann equation (2.53) gives

$$\chi_{\gamma_-} l_-^{-3\gamma_-} + \frac{1}{3}\Lambda_- - \frac{\zeta_-}{l_-^2} = \chi_{\gamma_+} l_-^{-3\gamma_+} \lambda^{-3\gamma_+} + \frac{1}{3}\Lambda_+ - \frac{\zeta_+}{l_-^2}\frac{1}{\lambda^2},$$

which has to be satisfied for all values of l_-. On both sides all terms contain different powers of l (note that $-3\gamma \in (-2, 6]$). In order that both sides contain the same powers of l we need $\gamma_- = \gamma_+$. Comparing the coefficients gives then $\lambda = 1, \zeta_+ = \zeta_-, \Lambda_+ = \Lambda_-$, and $\chi_{\gamma_+} = \chi_{\gamma_-}$. Hence the solutions are identical.

We conclude that the evolution of the Hubble parameter H uniquely identifies a FLRW model with a γ-equation of state.

2.5.4 Matching without surface-layer

From the results of the previous section it is clear that it is impossible to join generic spherically symmetric sections along a timelike hypersurface without a surface-layer, i.e., $\rho_s = 0$, except the metric satisfies some extraordinary conditions. In realistic models this could happen at most in singular situations and not along the whole trajectory of the junction hypersurface.

2.5.5 Comoving junction surface

Let us first investigate whether there could be a comoving junction surface, *i.e.*, a junction surface at a fixed value of the comoving radial coordinate r inside and outside the bubble. In this case $\dot{\alpha}_\pm = 0$ and from the geometric matching condition (specialized to the FLRW metric) (2.5) we find $N_+ = 1$. On the other hand (2.4) gives

$$l_+ = \underbrace{\frac{f_-(\alpha_-)}{f_+(\alpha_+)}}_{\text{const.}} l_-, \tag{2.55}$$

where the first factor is now time-independent. Given the result from the preceeding section we conclude that if the inside and outside of the bubble are evolving according to the Friedmann equation (2.52) with a γ-law equation of state then the inside and outside region must be identical in order to satisfy (2.55).

We note that this result follows alone from the geometric matching condition (2.5) – it does not depend on the presence of a surface layer.

2.5.6 Matching of FLRW regions with surface-layer

The FLRW metric (2.51) implies $w = l(t)f(r)$ and by taking proper-time and radial derivatives we derive

$$\dot{w} = Hlf = Hw \quad \text{and} \quad w' = \frac{\mathrm{d}f}{\mathrm{d}r}.$$

The component of the energy-momentum tensor which is needed to evaluate the surface-matter evolution according to (2.27) is easily found to be (for completeness we include k, which is set to unity)

$$[T^{\mu\nu}u_\mu n_\nu] = \left[\frac{N^2\dot{\alpha}l(\rho+p)}{k^2}\right].$$

We proceed now by expressing all quantities related to the metric and its derivatives ($a, b, w', \dot{w}$ etc.) in terms of FLRW model quantities. With $(\mathrm{d}f/\mathrm{d}r)^2 = 1 - \zeta f^2$ we obtain the expressions

$$a = 1 - f^2(H^2l^2 + \zeta) \overset{(2.52)}{=} 1 - \frac{\kappa\rho + \Lambda}{3}w^2 = 1 - w^2H^2\Omega.$$

$$b = a_+ - a_- = -\frac{w^2}{3}[\kappa\rho + \Lambda] = -w^2[H^2\Omega] = -w^2\left[\frac{\Omega}{l^2|\Omega - 1|}\right].$$

First we want to examine which kind of bubbles could exist if there can only be a positive surface-energy density on the junction surface, i.e., $E > 0$. For reasonably small bubbles we have

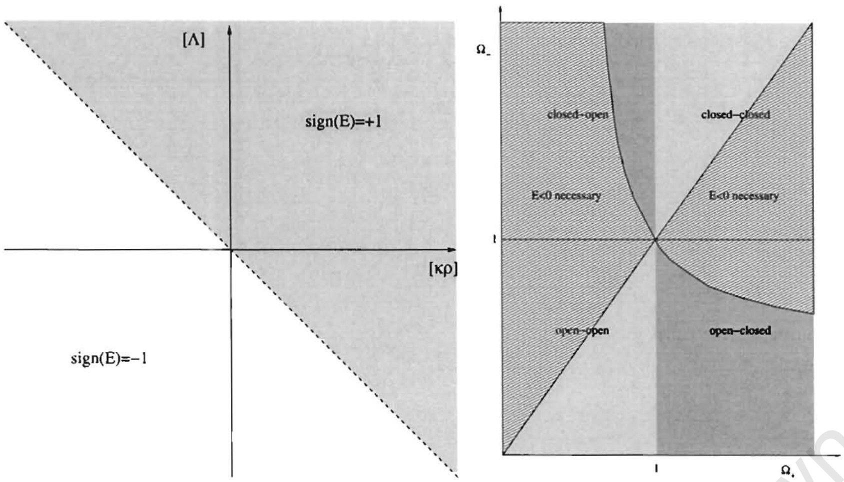
$$a_+ > 0 \qquad a_- > 0 \qquad w' > 0.$$

We find from table 2.1 that in this case

$$\mathrm{sign}(E) = -\mathrm{sign}(b) = \mathrm{sign}([\kappa\rho + \Lambda]),$$

and hence junctions are only possible if the inside FLRW region has a smaller total[3] energy density than the outside region. This is illustrated in figure 2.8(a). Figure 2.8(b) shows the allowed cases in the $\Omega_- - \Omega_+$-plane for equal scale-factors. The case $\kappa\rho_+ + \Lambda_+ = \kappa\rho_- + \Lambda_-$ corresponds to the 'no surface-layer' case.

[3]The cosmological constant represents the vacuum energy density.



(a) Points in the $\Lambda - \kappa\rho$-plane for which a matching requires a positive/negative surface-energy density.

(b) Points in the $\Omega_- - \Omega_+$-plane for which a matching requires a negative surface-energy density. This plot is for the case of equal scale-factors $l_+ = l_-$ and $w^2 H^2 \Omega < 1$ on both sides.

Figure 2.8: Only certain matchings are possible if there can only be a positive surface energy density.

2.5.7 A particular example

To understand the behaviour of the junction surface it is instructive to consider a particularly simple example for which the evolution equations are known. One such example is the junction between a non-inflating closed geometry with radiation inside ($\Lambda_- = 0, \zeta_- = +1$) and an inflating empty open geometry outside ($\chi_+ = 0, \zeta_+ = -1$).

For these cases the Friedmann equation (2.53) is easily integrated and one finds the well-known solutions

$$l_-(\tau_-) = \sqrt{2\tau_-\chi_- - \tau_-^2} \qquad l_+(\tau_+) = \sqrt{\frac{3}{\Lambda_+}} \sinh\left(\sqrt{\frac{\Lambda_+}{3}}\tau_+\right), \tag{2.56}$$

where τ_+ and τ_- are the proper times outside and inside, respectively. For the inside model the scale factor l_- grows until it reaches a maximum at $\tau_- = \chi_-$, and then declines until it reaches zero at $\tau_- = 2\chi$. On the contrary the outside model expands exponentially forever.

Since for each time t the proper time measured along the junction surface is always less or equal to the proper time in the inside and outside model ($u_\pm \geq 1$) it is clear that such a boundary can only exist for a finite proper time measured along the junction — the (timelike) junction surface must be 'terminated' at some time.

There are four possible solutions. Firstly, it is possible that the junction

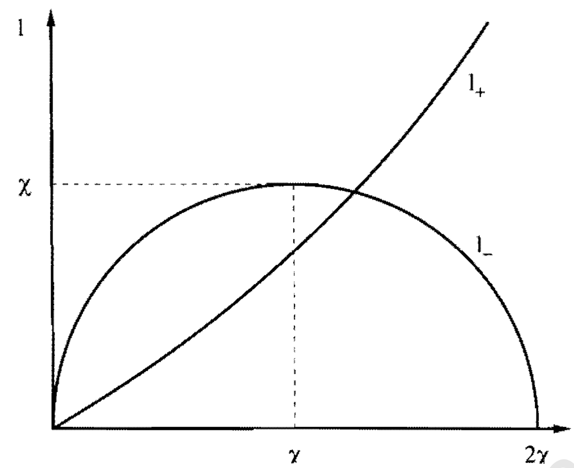


Figure 2.9: Evolution of the scale factor in a closed non-inflating FLRW model (l_-) and in an empty open inflating model (l_+).

surface exists forever (in terms of the proper time) in the outer region while the proper time along the junction surface is bounded. In our formalism this corresponds to a non-integrable divergence in the lapse function $N_+ = u_+$ for the outer region.

Secondly, the junction surface can contract to a point such that the inner region is eliminated. This case is characterized by α_+, α_- and w approaching zero at some finite time.

Thirdly, the closed surface might detach from the open geometry — the birth of a child universe. In this case the radial coordinate for the closed geometry α_- approaches π, while α_+ and w vanish (see figure 2.10).

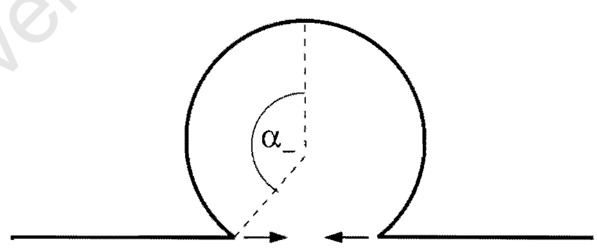


Figure 2.10: The closed inside geometry can detach from the open outside geometry. This happens when $\alpha_- \to \pi$ within a finite time.

As a last option the outside region might be eliminated. In the case of a closed outside geometry this is surely a possible solution, but in the cases of flat and open outside geometries this generally requires that the junction surface turns spacelike[4].

[4]One might speculate that for a closed inside geometry there has to be a finite volume and hence the junction surface has to turn backwards in time and become timelike again.

In our formalism such a behaviour would yield a diverging, but integrable, lapse function. At the singularity we reach the speed of light and our formalism breaks down. Nevertheless, of the three options, to become super luminal, to continue at the speed of light, or to decelerate, the first one seems most convincing, also with view on the results of section 2.2.6.

Generally, one of these cases has to occur before we reach the singularity in the inside region as can be seen from the following argument. Let us assume that for physical reasons only positive surface energies are allowed. Since the outside geometry is open we have $w'_+ > 0$. As the closed inside geometry approaches the 'big crunch' singularity the energy density grows without bound. Hence $b = w^2(\kappa\rho_- - \Lambda_+)$ has to become positive at some stage during the contraction phase. However, from table 2.1 one can see that there is no solution possible with $w'_+ > 0$ and $b > 0$ if a_+ and a_- are positive. Let us note that when the inside region is contracting we have $\dot{w}_- < 0$ and $\dot{w}_+ > 0$. Hence according to (2.47) a_+ and a_- cannot be both negative. If a_+ is positive (and a_- negative) then (2.46) implies that the surface energy density is negative, which is in contradiction with our assumption. If on the other hand $a_+ < 0$ then this implies $L > 0$ and hence the proper surface radius would increase. This just helps driving $a_- = 1 - \kappa\rho_- w^2/3$ closer to zero, which eventually has to turn negative due to the diverging energy density. Again we reach a point where no solution is possible without negative surface energies.

Note that if one allows negative surface energies then the above argument shows that if the junction starts with a positive surface-energy density then at some point the junction must have a vanishing surface-energy density. For a generic situation this implies that at this point the junction moves with the speed of light.

Figure 2.14 shows the results of a numerical integration of this particular model. It appears as if the speed of light is reached within a finite time (integrable lapse functions) on both sides. This strongly suggests to us that the junction turned spacelike. Note that this happens even far before the inner closed region enters the contracting phase.

With this example we want to emphasize that there might be junctions which are possible initially, but which evolve to some singular point. As can be seen from the numerous examples in the next section this behaviour appears to be rather common.

2.6 Numerical Results

A computer program has been written to integrate the evolution equations for several FLRW junctions numerically. To achieve better accuracy around the singularities a variable step-width was used. All examples given here are for positive surface-energy densities. Cases with negative surface energy can easily be constructed by exchanging the inside and outside region. The graphs on the following pages illustrate the results and will be discussed one-by-one below.

- *Open inside, inflating closed geometry outside* Figure 2.11 shows such an example. After some time the surface radius starts to diverge (note the logarithmic scaling) while $E = \kappa w \rho_s/2$ does not approach zero (hence close to the divergence $\rho_s \propto 1/w$). The proper times on both sides seem to be diverging, which is in agreement with the results from subsection 2.2.6. Note that the inner and outer regions have a rather unusual equation of state with $\gamma_+ = 0.7$ and $\gamma_- = 1.9$ — in this case it is really this choice of the equations of state which makes the energy-momentum tensor contribution $[T_{\mu\nu}u^\mu n^\nu]$ positive for small values of E (see subsection 2.2.6).

Figure 2.12 shows the evolution of the same initial situation, but with different equations of state (dust on both sides). This seems to change the sign of the energy-momentum contribution $[T_{\mu\nu}u^\mu n^\nu]$ for small values of E, which now approaches zero within a finite time. In fact, it can be verified that close to the singular point t_0 we have as expected $E \propto \sqrt[3]{t_0 - t}$. As predicted in subsection 2.2.6 the lapse functions appear to be integrable and the proper times do not diverge. The junction surface seems to reach the speed of light within a finite time.

Clearly, our formalism breaks down at this point. However, one could argue that after reaching the speed of light within a finite time, one should expect the junction to turn spacelike.

- *Closed inside, inflating open outside* Such a case with dust-equations of state on each side is illustrated in figure 2.13. Similarly to the last case E seems to reach zero within a finite coordinate and proper time.

  Figure 2.14 shows a similar situation, but with a radiation equation of state for the inside region ($\gamma_- = 4/3$). This is the example given in the previous subsection.

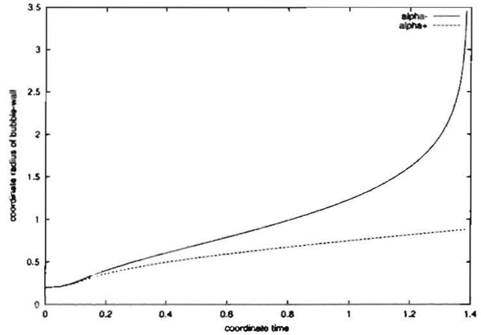
- *Flat inside with vacuum and inflating flat outside with radiation* This case is illustrated in figure 2.15. Note that here the coordinate radii and the proper speeds are plotted with respect to the proper time. Also in this case we encounter a non-integrable divergency in the lapse function. Note that in the outside region the bubble first grows and then shrinks for some time, before it enters a phase of indefinite expansion.

For the junction between two closed FLRW geometries figure 2.16 shows the evolution in the $E - \sqrt{a_+/a_-}$ plane for different initial surface-energy densities.

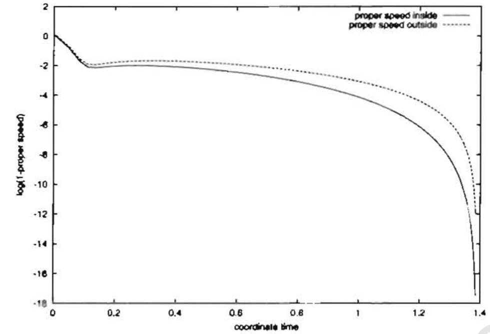
2.7 Conclusion

We developed a formalism for the treatment of timelike junctions between spherically symmetric solutions of the Einstein-field equation, which is based on the Lanczos equation and the Israel junction conditions. We introduce new coordinates such that two conditions are satisfied: Firstly, all coordinates are continuous at the junction surface, and secondly, the junction surface becomes a surface of constant 'radial' coordinate. In this approach the actual movement of the junction surface is absorbed into the metric, of which the transverse components are discontinuous at the junction surface.

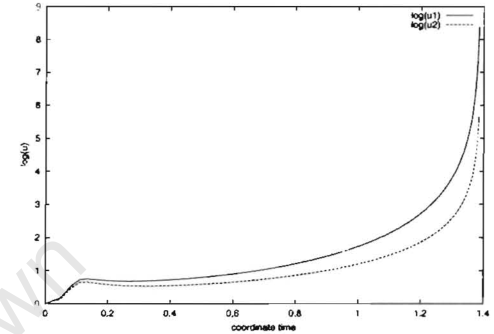
We evaluate the junction conditions and re-discover with (2.34) and (2.35) well-known relations between the extrinsic curvatures, the surface layer energy density, and the rate of change of the surface radius of the junction surface. It should be pointed out that these results follow without using the time-component of the Lanczos equation. As it was shown in subsection 2.4, for



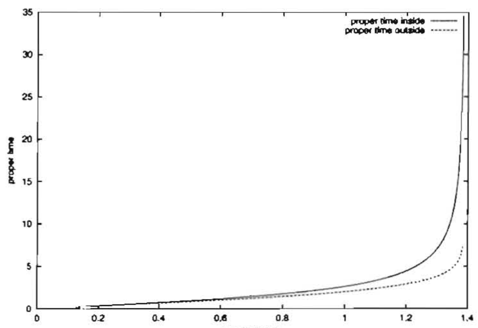
(a) Coordinate radius of the junction surface on each side



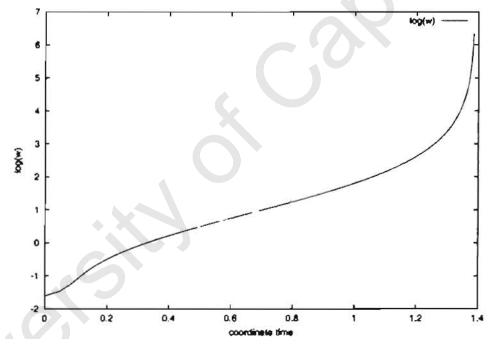
(b) Proper speed of the junction surface on each side.



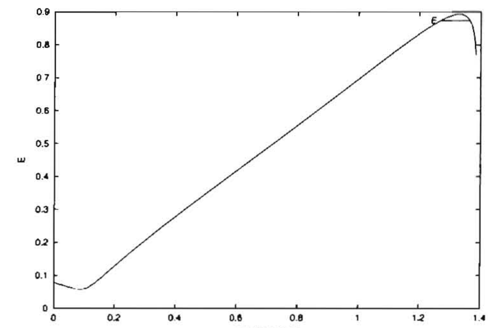
(c) Lapse function $N = u$ on each side



(d) Proper times on both sides

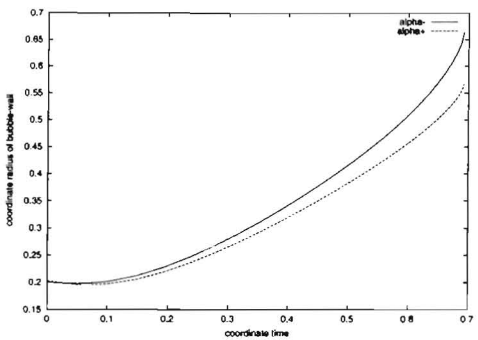


(e) Surface radius (w).

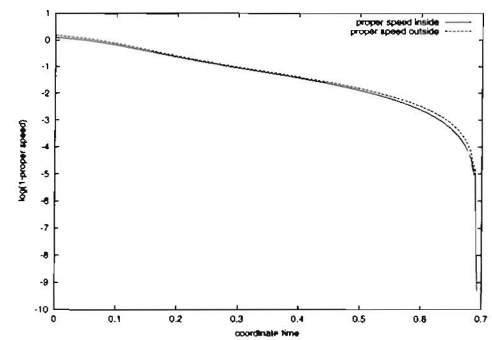


(f) E, which is related to the surface-matter energy density ρ_s by $E = \kappa w \rho_s/2$.

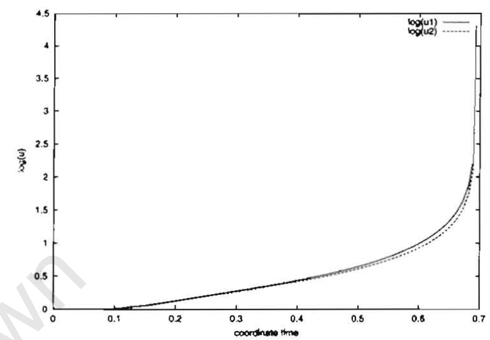
Figure 2.11: Evolution of a junction between an open (inside) and an inflating closed (outside) geometry. (Parameters:$\gamma_s = 1, \zeta_+ = +1, \zeta_- = -1, \chi_+ = 5, \chi_- = 2, \gamma_+ = 0.7, \gamma_- = 1.9, \Lambda_+ = 2, \Lambda_- = 0$; Initial values: $\rho_s = 0.031, \alpha_+ = 0.20272, \alpha_- = 0.2, l_+ = 1, l_- = 1, \rho_+ = 0.597, \rho_- = 0.239$.) Note that here u and w are plotted logarithmically.



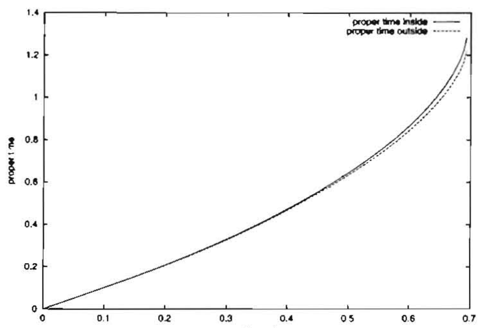
(a) Coordinate radius of the junction surface on each side



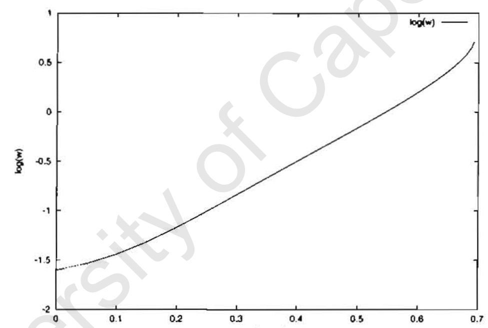
(b) Proper speed of the junction surface on each side.



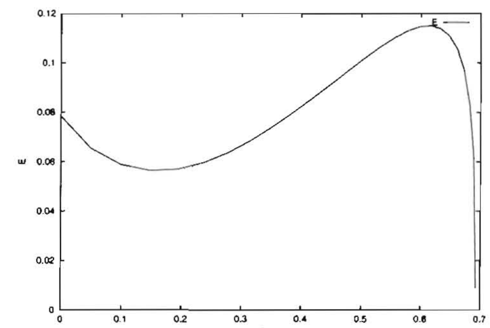
(c) Lapse function $N = u$ on each side



(d) Proper times on both sides

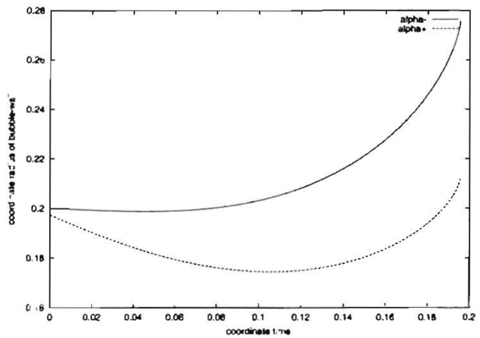


(e) Surface radius (w).

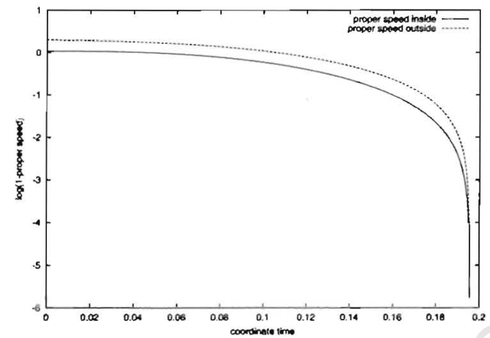


(f) E, which is related to the surface-matter energy density ρ_s by $E = \kappa w \rho_s / 2$.

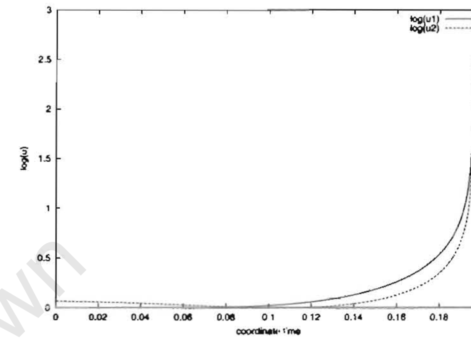
Figure 2.12: Evolution of a junction between an open (inside) and an inflating closed (outside) geometry. The lapse functions are integrable and the speed of light is reached within a finite time. (Parameters:$\gamma_s = 1, \zeta_+ = +1, \zeta_- = -1, \chi_+ = 5, \chi_- = 2, \gamma_+ = 1, \gamma_- = 1, \Lambda_+ = 2, \Lambda_- = 0$; Initial values: $\rho_s = 0.031, \alpha_+ = 0.20272, \alpha_- = 0.2, l_+ = 1, l_- = 1, \rho_+ = 0.597, \rho_- = 0.239$.) Note that here u and w are plotted logarithmically.



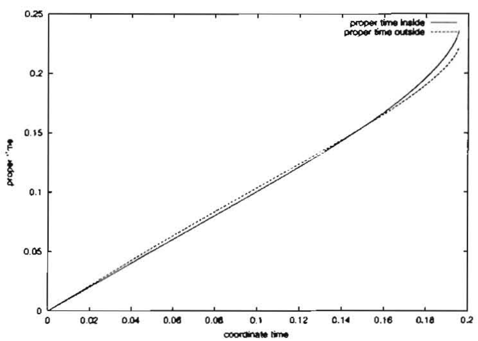
(a) Coordinate radius of the junction surface on each side



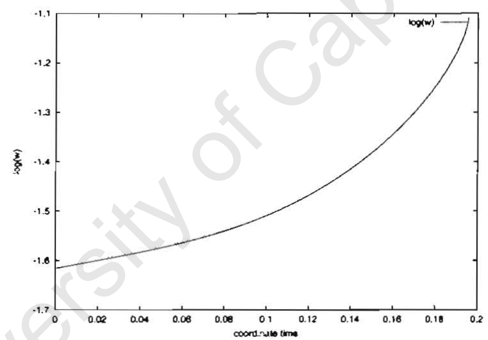
(b) Proper speed of the junction surface on each side.



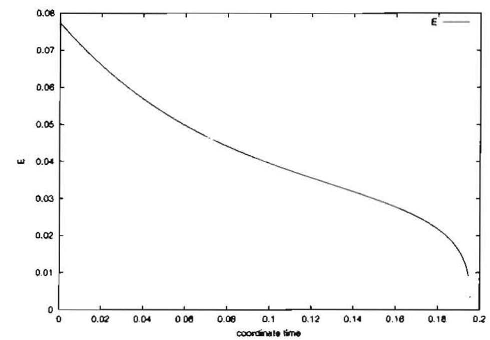
(c) Lapse function $N = u$ on each side



(d) Proper times on both sides

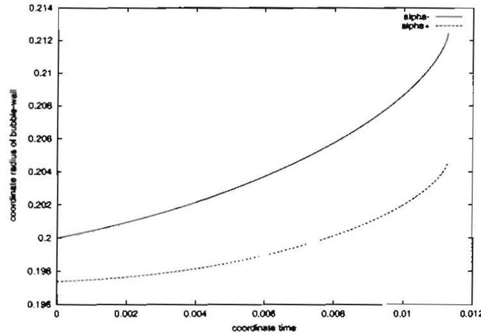


(e) Surface radius (w).

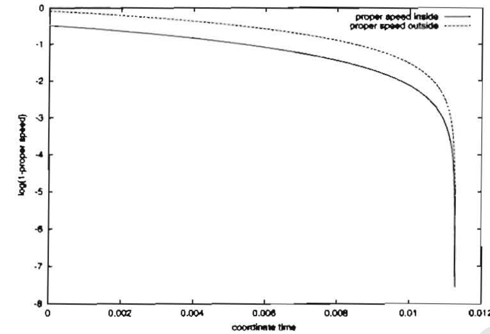


(f) E, which is related to the surface-matter energy density ρ_s by $E = \kappa w \rho_s / 2$.

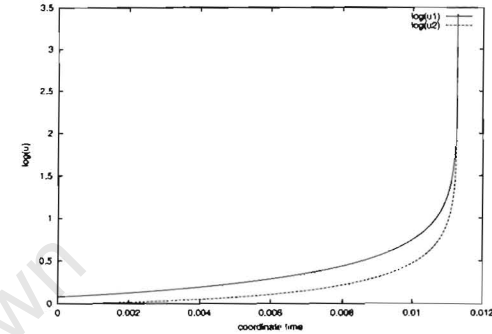
Figure 2.13: Evolution of a junction between a closed (inside) and an inflating open (outside) geometry. (Parameters: $\gamma_s = 1, \zeta_+ = -1, \zeta_- = +1, \chi_+ = 5, \chi_- = 2, \gamma_+ = 1, \gamma_- = 1, \Lambda_+ = 2, \Lambda_- = 0$; Initial values: $\rho_s = 0.031, \alpha_+ = 0.19739, \alpha_- = 0.2, l_+ = 1, l_- = 1, \rho_+ = 0.597, \rho_- = 0.239$.) Note that here u and w are plotted logarithmically.



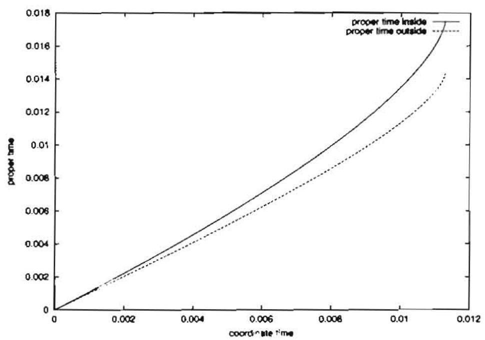
(a) Coordinate radius of the junction surface on each side



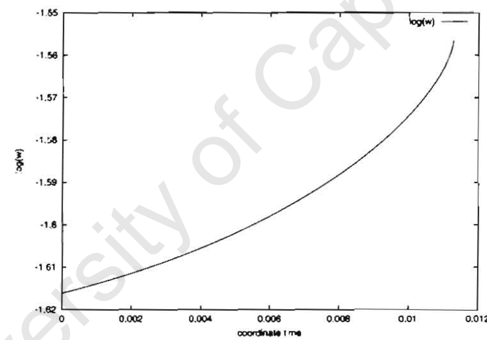
(b) Proper speed of the junction surface on each side.



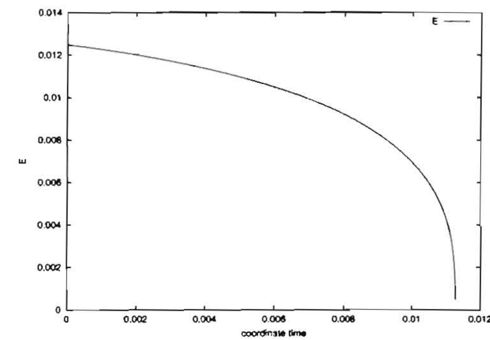
(c) Lapse function $N = u$ on each side



(d) Proper times on both sides

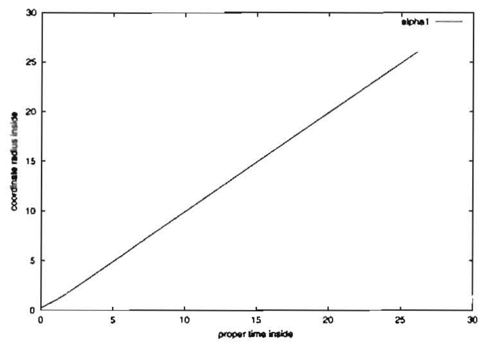


(e) Surface radius (w).

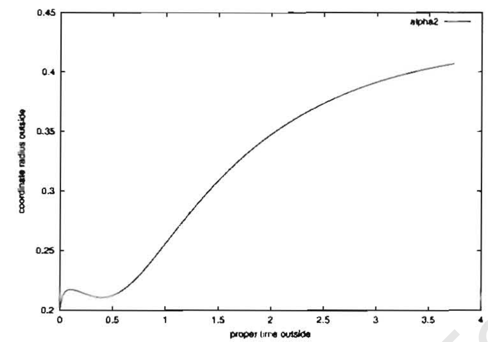


(f) E, which is related to the surface-matter energy density ρ_s by $E = \kappa w \rho_s / 2$.

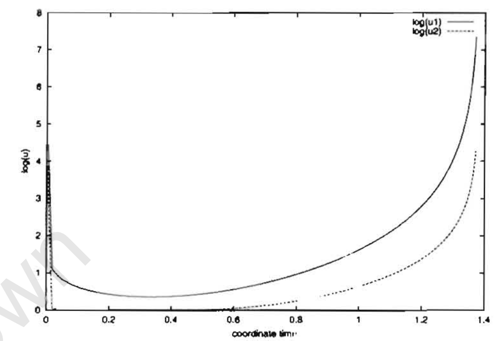
Figure 2.14: Evolution of a junction between a closed (inside) and an inflating open (outside) geometry. The inside has a radiation equation of state, while the outside is an inflating dust model. (Parameters: $\gamma_s = 1, \zeta_+ = -1, \zeta_- = +1, \chi_+ = 0, \chi_- = 1, \gamma_+ = 1, \gamma_- = 4/3, \Lambda_+ = 5, \Lambda_- = 0$; Initial values: $\rho_s = 0.005, \alpha_+ = 0.19739, \alpha_- = 0.2, l_+ = 1, l_- = 1, \rho_+ = 0, \rho_- = 0.119.$) Note that here u and w are plotted logarithmically.



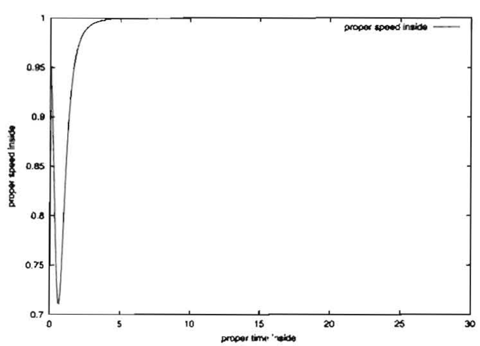
(a) Coordinate radius of the junction surface inside.



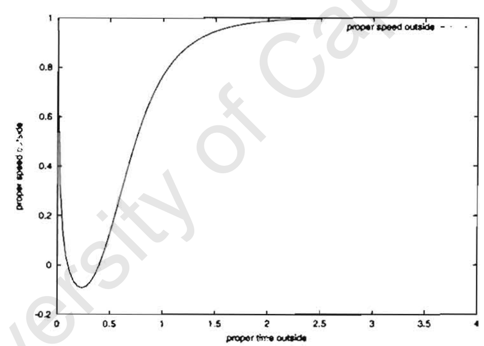
(b) Coordinate radius of the junction surface outside.



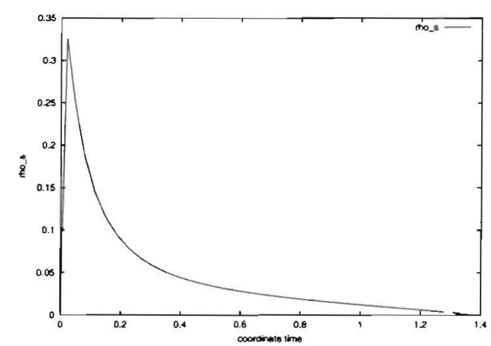
(c) Lapse function $N = u$ on each side



(d) Proper speed inside.



(e) Proper speed outside.



(f) Surface-matter energy density ρ_s.

Figure 2.15: Evolution of a junction between two flat geometries. The outside has a radiation equation of state and is inflating, while the inside is vacuum model. (Parameters: $\gamma_s = 1, \zeta_+ = +0, \zeta_- = +0, \chi_+ = 200, \chi_- = 0, \gamma_+ = 1.33, \gamma_- = 1, \Lambda_+ = 2, \Lambda_- = 0$; Initial values: $\rho_s = 0.005, \alpha_+ = 0.2, \alpha_- = 0.2, l_+ = 1, l_- = 1, \rho_+ = 23.9, \rho_- = 0.$) Note that here u is plotted logarithmically.

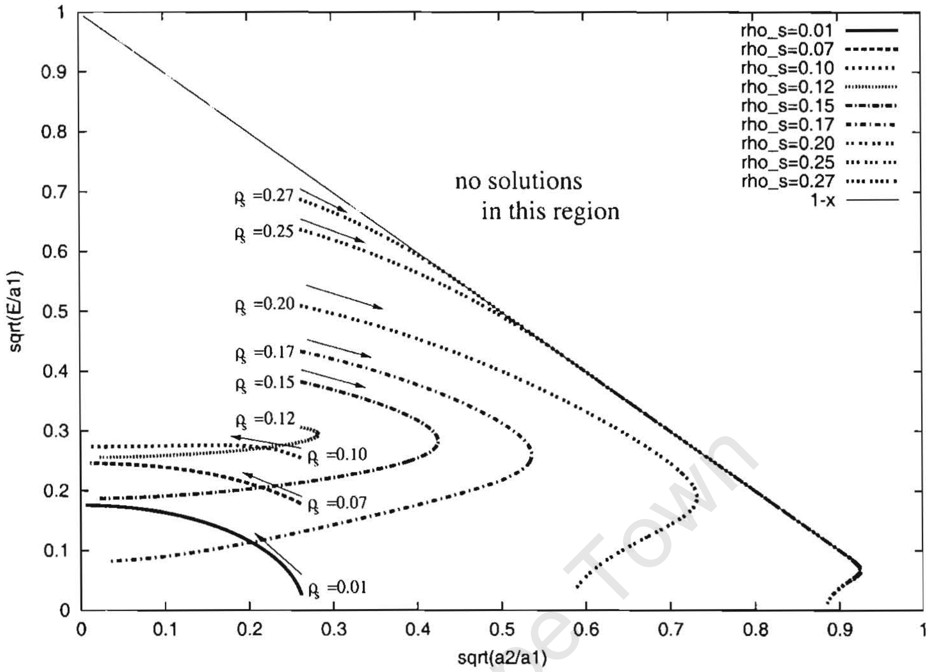


Figure 2.16: Evolution of junctions between two closed geometries in the $|E|/\sqrt{a_-}$ – $\sqrt{a_+/a_-}$ plane. The different trajectories correspond to different initial-values of the surface-energy density ρ_s. All other initial parameters are constant.

all spherically symmetric cases this remaining equation is in fact an identity. This was known for special cases, but it appears to be a new result in this general form.

The behaviour for small values of $E = \kappa w \rho_s/2$ has been investigated. It was shown that for certain cases E is driven to zero within a finite coordinate and proper time. At such a point our formalism breaks down. Nevertheless, we want to speculate here that in such cases the junction really turns spacelike. This can be seen as an inadequacy of the thin wall formulation in such situations — a causal propagation of a discontinuity should not exceed the speed of light. We suggest that in such cases the spatial extent of the transition region is not negligible.

The developed formalism gives us two sources for constraints on possible junctions. Firstly the time derivative of the surface radius is given by the quadratic equation (2.36). Demanding that real solutions to this equation must exist directly restricts the possible values of the surface energy density for a particular junction (see figure 2.5). Secondly, in our approach physical solutions must have a lapse function which is greater than or equal to unity. The resulting restrictions depend on the metric components on each side of the junction — they either determine the sign of the derivative of the proper surface radius, or they restrict the possible surface-energy densities (see figure 2.6). For the latter case the allowed ranges for $E = \kappa w \rho_s/2$ have been given explicitly.

For the special case of junctions between FLRW models with γ-equation of state it was shown that alone on geometrical grounds there can be no comoving junction surface — whether with or without surface layer.

A particularly interesting model, the junction between an empty, open, inflating FLRW region outside and a radiation dominated closed FLRW model inside, has been investigated in more detail. The inside region re-collapses after some finite proper time and hence the junction surface has to be terminated. Besides a disappearance or a detachment of the closed inner region we suggest that the junction can turn spacelike — an effective disappearance of the outer region.

This and other examples have been integrated numerically. It was observed that many models seem to reach the speed of light within a finite *proper* time, in accordance with the predictions from section 2.2.6. Since a spacelike junction violates causality, we suggest that a breakdown in the thin-wall approximation must have occurred.

Our results show that the thin-wall treatment of timelike junctions (without the presence of scalar fields) is on a *mathematical* sound level. Nevertheless, in many cases the junction surface reaches a singular point within a finite proper time. We believe that in these cases the thin-wall is not a *physically* acceptable approximation.

References

[1] Spergel D N *et al* 2003 *Astrophys. J.* (*Preprint* astro-ph/0302209)

[2] Guth A H 1981 *Phys. Rev.* D **23** 347

[3] Linde A D 1983 *Phys. Lett.* B **129** 177

[4] Linde A D 1990 *Particle Physics and Inflationary Cosmology* (Chur:Harwood Academic Publishers)

[5] Lanczos C 1922 *Phys. Zeits.* **23** 539; and 1924 *Ann. der Phys.* **74** 518

[6] Dautcourt R 1964 *Math. Nachr.* **27** 277

[7] Israel W 1966 *Nuovo Cimento* B **44** 4349

[8] Blau S K, Guendelman E I and Guth A H 1987 *Phys. Rev.* D **35** 1747

[9] Berezin V A, Kuzmin V A and Tkachev I I 1987 *Phys. Rev.* D **36** 2919

[10] Lake K 1987 in *Vth Brazilian School of Cosmology and Gravitation*, Ed. Novello, World Scientific, Singapore, pp1-82

[11] Hájíček P and Bičák J 1997 *Phys. Rev.* D **56** (*Preprint* gr-qc/9706022)

[12] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation*, W. H. Freeman and Company, San Francisco

[13] Sakai N and Maeda K 1994 *Phys. Rev.* D **50** 5425-5428 (*Preprint* gr-qc/9311024)

[14] Coleman S and De Luccia F 1980 *Phys. Rev.* D **21** 3305

[15] Jensen L G and Steinhardt P J 1984 *Nucl. Phys.* B **237** 176-188

[16] Lyth D H and Stewart E D 1994 *Preprint* hep-ph/9408324

[17] Amendola L *et al* 1996 *Preprint* astro-ph/9610038

Chapter 3

Defining Multiverses

3.1 Introduction

The idea of a multiverse has been proposed as the only scientifically based way of avoiding the fine-tuning required to set up the conditions for our seemingly very unlikely universe to exist. Stephen Weinberg (2000a,b), for example, uses it to explain the value of the cosmological constant, which he relates to anthropic issues. Martin Rees (2001a,b,c) employs it to explain the whole set of anthropic coincidences (Barrow and Tipler 1986), that is, to explain why our universe is a congenial home for life. These and similar proposals have been triggered by the dawning awareness among many researchers that there may be many other existing universes besides ours. This possibility has received strong stimulation from proposals like Andrei Linde's chaotic inflationary scenario (Linde 1983, 1990), in which the origin of our own observable universe region naturally involves the origin of many other similar expanding universe regions.

There is however a vagueness about the proposed nature of multiverses. They might occur in various ways, discussed by Leslie (1996), Weinberg (2000), Tegmark(2003), Gardner (2003). They might originate naturally in different times and places[1] through meta-cosmic processes like chaotic inflation (Linde 1983, 1990), or in accord with Lee Smolin's (1999) cosmic Darwinian vision. In the latter case, an ensemble of expanding universe regions grow from each other following gravitational collapse and re-expansion, where natural selection of universes through optimization of black hole production leads to bio-friendly universe regions. This is an intriguing idea, but with many uncertain steps – in particular no proof has been given of the last step, that the physics that maximizes black hole production also favours life. They might be associated with the multi-universe Everett-Wheeler-type interpretation of quantum mechanics Deutsch(1998) (but see discussion below), or perhaps multiverses can be truly disjoint collections of universes (see Sciama 1993, Lewis 2000, Rees 2001a,b,c, Tegmark 1998, 2003).

Some refer to the separate expanding universe regions in chaotic inflation as 'universes', even though they have a common causal origin and are all part of the same single space-time. In our view (as 'uni' means 'one') *the Universe*

[1] c.f. Ellis 1979.

is by definition the one unique connected [2] existing space-time of which our observed expanding cosmological domain is a part. We will refer to situations such as in chaotic inflation as a *Multi-Domain Universe*, as opposed to a completely causally disconnected *Multiverse*. Throughout this chapter, when our discussion pertains equally well to disjoint collections of universes (multiverses in the strict sense) and to the different domains of a Multi-Domain Universe, we shall for simplicity simply use the word "*ensemble*". When an ensemble of universes are all sub-regions of a larger connected space-time – the "Universe as a whole" – we have the multi-domain situation, which should be described as such. Then we could reserve "multiverse" for the collection of genuinely disconnected "universes" – those which are not locally causally related.

The multi-universe interpretation of quantum-mechanics takes a somehow special position: The different 'universes' have a common causal origin and well-defined probability distributions. However, these universes are only macroscopically causally disconnected, which results from decoherence rather than real causal disconnectedness. In fact, it is the interaction between the different 'universes' which explains the quantum effects on a microscopic level. Such a concept still faces the profound difficulties associated with realized infinities, which will be discussed below. However, with the current lack of a consistent theory of quantum-gravity it is not clear how to describe the variations in geometry across the ensemble in a multi-universe interpretation of quantum-mechanics. For this reason we want to exclude this kind of multiverse interpretation from the following discussion.

So far, none of these ideas have been developed to the point of actually describing such ensembles of universes in detail, nor has it been demonstrated that a generic well-defined ensemble will admit life. Some writers tend to imply that there is only one possible multiverse (characterized by "all that can exist does exist") (Lewis 2000, see also Gardner 2003). This vague prescription actually allows a vast variety of different realizations with differing properties, leading to major problems in the definition of the ensembles and in averaging, due to the lack of a well-defined measure and the infinite character of the ensemble itself. Furthermore it is not at all clear that we shall ever be able to accurately delineate the class of all possible universes.

The aim of this chapter is to help clarify what is involved in a full description of an ensemble of universes. The first contribution is clarifying what is required in order to describe the space of possible universes, where much hinges on what we regard as 'possible'. However that is only part of what is needed. It is crucial to recognize that anthropic arguments for *existence* based on ensembles of universes with specific properties require an *actually existing ensemble* with all the required properties. For purposes of providing an explanation of existence, simply having a conceptually possible ensemble is not adequate – one needs a link to objects or things that actually exist, or to mechanisms that make them exist.

The second contribution of this chapter is to show how an actually existing ensemble may be described in terms of a space of possible universes, by

[2] "Connected" implies "Locally causally connected", that is all universe domains are connected by C^0 timelike lines which allow any number of reversals in their direction of time, as in Feynman's approach to electrodynamics. Thus it is a union of regions that are causally connected to each other, and transcends particle and event horizons; for examples all points in de Sitter space time are connected to each other by such lines.

defining a distribution function (discrete or continuous) on the space of possible universes. This characterizes which of the theoretically possible universes have been actualized in the ensemble - it identifies those that have actually come into existence. This leads us to our third point: the problems arising when it is claimed that there is an actually existing ensemble containing an infinite number of universes or of expanding universe regimes. Actually existing infinities are very problematic.

There are fundamental issues that arise in considering ensembles of actually existing universes: what would explain the existence of an ensemble, and its specific properties? Why should there be this particular ensemble, rather than some other one? Why should there be any regularity at all in its properties? The fourth point we make is that if all the universes in an ensemble show regularities of structure then that implies some common generating mechanism. Some such structuring is necessary if we are to be able to describe a multiverse with specified properties - a coherent description is only possible through the existence of such regularities. Hence a multiverse consisting of completely causally disconnected universes is a problematic concept.

The issue of testability is a further important consideration: Is there any conceivable direct or indirect way of testing for existence of an ensemble to which our universe belongs? Our fifth point is that there is no way we can test any mechanism proposed to impose such regularities: they will of necessity always remain speculative. The sixth point (and see Leslie 1989, Gardner 2003) is to argue that existence of multiverses or ensembles is in principle untestable by any direct observations, and the same applies to any hypothesized properties we may suppose for them. However, certain observations would be able to disprove existence of some multi-domain ensembles. It is only in that sense that the idea is a testable proposition.

It is clear that in dealing with multiverses one inevitably runs up against philosophical and metaphysical issues, for example concerning the ability to make scientific conclusions in the absence of observational evidence, and in pursuing the issue of realized infinities.

3.2 Describing Ensembles: Possibility

To characterize an ensemble of existing universes, we first need to develop adequate methods for describing the class of all possible universes. This requires us to specify, at least in principle, all the ways in which universes can be different from one another, in terms of their physics, chemistry, biology, etc.

3.2.1 The Set of Possible Universes

The basis for describing ensembles or multiverses 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 characterized 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 parametrization 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, characterized 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 normalized 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 characterization of this space, and its relationship to $\mathcal{S}$, will depend on the matter description used and its behaviour. The overall characterization 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 Rindler1989).

2. $\mathcal{M}_{\mathrm{almost-FLRW}}$ – the subset of all perturbed FLRW model universes. These need to be characterized 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, 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? "All that can happen happens" must imply all possibilities, as characterized by our description in terms of families of parameters: all allowed values 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, along with many others we cannot even conceive of, but which can nevertheless in principle also be described by such parameters. An interesting related point has been pointed to us by Jean-Phillipe Uzan: it may be that 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.

From these considerations we have the first key issue:

Issue 1: What determines $\mathcal{M}$? Where does this structure come from? What is the meta-cause that delimits this set of possibilities? Why is there a uniform structure across all universes m in $\mathcal{M}$?

The meta-question is whether any of these questions can be answered scientifically. We return to that at the end.

3.2.2 Adequately Specifying Possible Anthropic Universes

When defining any ensemble of universes, possible or realized, 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 categorize 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 characterization 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 $p_1(i)$; 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, baryons, 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. barotropic equations of state and the potential function for scalar fields). These are characterizations 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 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 characterized by $p_3(i)$.

5. $p_5(i)$ are geometrical parameters characterising the space-time 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 characterized in the periodic table), $p_6(2)$ the number of different states of matter possible (solid, liquid, gas, plasma for example), and $p_6(3)$ the number of different types of molecular bonding. 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

characterize 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 characterize 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 characterized 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 characterized by particular class parameters.

3.2.3 Describing the Geometry of Possible Universes

Cosmological models are characterized by a preferred timelike vector field u : $u^a u_a = -1$, usually the fluid flow vector (Ellis 1971a), but sometimes chosen for other reasons, e.g., to fit local symmetries. 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 characterize 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 ($a, b, c \dots$ 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

$$\mathcal{P}_{\text{frame}} \equiv \{\mathcal{X}, \mathcal{E}\}. \tag{3.1}$$

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:

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

(because the tetrad is orthonormal) and $e^b{}_j(x^l)$ 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

$$ds^2 = -\left(e^0_i dx^i\right)^2 + \left(e^1_i dx^i\right)^2 + \left(e^2_i dx^i\right)^2 + \left(e^3_i dx^i\right)^2 \tag{3.2}$$

The basic geometric quantities used to determine the space-time 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

$$\mathcal{P}_{\text{geometry}} \equiv \{\dot{u}_\alpha, \theta, \sigma_{\alpha\beta}, \omega_{\alpha\beta}, \Omega_\gamma, a^\alpha, n_{\alpha\beta}\}, \tag{3.3}$$

characterized 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 space-time fields, although in particular high-symmetry cases they may be independent of many or of all the coordinates. 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). Except in the special cases of isotropic space-times and locally rotationally symmetric space-times (see Ellis 1967, van Elst and Ellis 1996), 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 space-time functions in the combined set $(\mathcal{E}, \mathcal{P}_{\text{geometry}})$ with some chosen specification of coordinates $\mathcal{X}$, with the metric then determined by (3.2). For detailed dynamical studies it is often useful to rescale the variables in terms of the expansion (see Wainwright and Ellis 1995, 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 space-times. 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.

3.2.4 Describing the Physics of Possible Universes

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

$$\mathcal{P}_{\text{matter}} \equiv \{\mu, q_\alpha, p, \pi_{ab}, \Phi_A\} \tag{3.4}$$

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 $(A = 1..A_{\max})$ 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 characterized 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). 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

$$\begin{aligned} \mathcal{P}_{\text{eos}} \equiv \; & \{q_\alpha = q_\alpha(\mu, \Phi_A), \quad p = p(\mu, \Phi_A), \\ & \pi_{ab} = \pi_{ab}(\mu, \Phi_A), \; \dot{\Phi}_A = \dot{\Phi}_A(\Phi_A)\}. \end{aligned} \tag{3.5}$$

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 (2002), Horwood et al (2002) for the case of Bianchi models (characterized 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 realized as branes in a higher dimensional space-time, others will not.

3.2.5 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 (Bennett et al 2002) gives for the expected number of planets with intelligent life in any particular universe m in an ensemble,

$$N_{\text{life}}(m) = N_g * N_S * \Pi * F, \tag{3.6}$$

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

$$\Pi = f_S * f_p * n_e \tag{3.7}$$

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

$$F = f_l * f_i. \tag{3.8}$$

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 characterized 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 depends on the physics parameters $p_1(i)$ in a crucial way that is not fully understood. It will be impossible at any stage to characterize 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 realized simultaneously. When we look at these combinations, they will span a very small subset of the whole parameter space (Davies 2003, Tegmark 2003).

If we wish to deal with specifically human life, we need to make the space $\mathcal{M}$ large enough to deal with all relevant parameters for this case, where free will arises. This raises substantial extra complications.

3.2.6 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. This means we can provide necessary sets of parameter values but cannot guarantee completeness or independence.

3.3 The Set of Realized Universes

We have now characterized the set of possible universes. But in any given ensemble, they may not all be realized, and some may be realized many times.

The purpose of this section is to set up a formalism making clear which of the *possible* universes (characterized above) occur in a specific *realized* ensemble.

3.3.1 A distribution function describing an ensemble of realized universes

In order to select from $\mathcal{M}$ a set of realized universes we need to define on $\mathcal{M}$ a distribution function $f(m)$ specifying how many times each type of possible universe m in $\mathcal{M}$ is realized. The function $f(m)$ expresses the contingency in any actualization – the fact that not every possible universe has to be realized, and that any actual universe does not have to be realized as a matter of necessity. Things could have been different! Thus, $f(m)$ describes the *ensemble of universes* or *multiverse* envisaged as being realized out of the set of possibilities. If these realizations were determined by the laws of necessity alone, they would simply be the set of possibilities described by $\mathcal{M}$. In general they include only a subset of possible universes, and multiple realizations of some of them. This is the way in which chance or contingency is realized in the ensemble[3].

The class of models considered is determined by all the parameters held constant ('class parameters'). Considering the varying parameters for the class ('member parameters'), some will take only discrete values, but for each one allowed to take continuous values we need a volume element of the possibility space M characterized by parameter increments $\mathrm{d}p_j(i)$ in all such varying parameters $p_j(i)$. The volume element will be given by a product

$$\pi = \Pi_{i,j}\, m_{ij}(m)\, \mathrm{d}p_j(i) \tag{3.9}$$

where the product $\Pi_{i,j}$ runs over all continuously varying member parameters i, j in the possibility space, and the m_{ij} weight the contributions of the different parameter increments relative to each other. These weights depend on the parameters $p_j(i)$ characterising the universe m. The number of universes corresponding to the set of parameter increments $\mathrm{d}p_j(i)$ will be $\mathrm{d}N$ given by

$$\mathrm{d}N = f(m)\pi \tag{3.10}$$

for continuous parameters; for discrete parameters, we add in the contribution from all allowed parameter values. The total number of universes in the ensemble will be given by

$$N = \int f(m)\pi \tag{3.11}$$

(which will often diverge), where the integral ranges over all allowed values of the member parameters and we take it to include all relevant discrete summations. The probable value of any specific quality $p(m)$ defined on the set of universes will be given by

$$P = \int p(m)f(m)\pi. \tag{3.12}$$

[3]It has been suggested to us that in mathematics terms it does not make sense to distinguish identical copies of the same object: they should be identified with each other because they are essentially the same. But we are here dealing with physics rather than mathematics, and with real existence rather than possible existence, and then multiple copies must be allowed (for example all electrons are identical to each other; physics would be very different if there were only one electron in existence).

Such integrals over the space of possibilities give numbers, averages, and probabilities.

Hence, a (realized) ensemble E of universes is described by a possibility space $\mathcal{M}$, a measure π on $\mathcal{M}$, and a distribution function $f(m)$ on $\mathcal{M}$:

$$E = \{\mathcal{M}, \pi, f(m)\}. \tag{3.13}$$

The distribution function $f(m)$ might be discrete (e. g., there are 3 copies of universe m_1 and 4 copies of universe m_2, with no copies of any other possible universe), or continuous (e.g., characterized by a given distribution of densities Ω_i). In many cases a distribution function will exclude many possible universes from the realization it specifies.

Now it is conceivable that all possibilities are realized – that all universes in $\mathcal{M}$ exist at least once. This would mean that the distribution function

$$f(m) \neq 0 \text{ for all } m \in \mathcal{M}.$$

But there are an infinite number of distribution functions which would fulfil this condition, and so a really existing 'ensemble of all possible universes' is not unique. In such ensembles, all possible values of each distinguishing parameter would be predicted to exist in different members of the multiverse in all possible combinations with all other parameters at least once, but they may occur many times. One of the problems is that this often means that the integrals associated with such distribution functions would diverge, preventing the calculation of probabilities from such models (see our treatment of the FLRW case below).

From this consideration we have the second key issue:

Issue 2: What determines $f(m)$? What is the meta-cause that delimits the set of realizations out of the set of possibilities?

The answer to this question has to be different from the answer to *Issue 1*, precisely because here we are describing the contingency of selection of a subset of possibilities from the set of all possibilities, determination of the latter being what is considered in *Issue 1*. Again, the meta-question is whether this can be answered scientifically.

3.3.2 Measures and Probabilities

It is clear that $f(m)$ will enable us to derive numbers and probabilities relative to the realization it defines only if we also have determined a unique measure π on the ensemble, characterized by a specific choice of the weights $m_{ij}(m)$ in (3.9), where these weights will depend on the $p_j(i)$. There are three issues here.

First, what may seem a "natural" measure for $\mathcal{M}$ in one set of coordinates will not be natural in another set of coordinates. Hence the concept of a measure is not unique, as is illustrated below in the FLRW case. This is aggravated by the fact that the parameter space will often contain completely different kinds of quantities (density parameters and the values of the gravitational constant and the cosmological constant, for example), and assigning the weights entails somehow assigning a relative weighting between these quite different kinds of quantities.

Second, it is possible that we might be able to assign probabilities $\chi(m)$ to points of $\mathcal{M}$ from some kind of physical argument, and then predict $f(m)$ from these, following the usual line of argument for determining entropy in a gas. However, we then have to determine some reason why $\chi(m)$ is what it is and how it then leads to $f(m)$. In the entropy case, we assume equal probability in each phase space volume; why should that hold for an ensemble of universes? Realising such probabilities seems to imply a causal mechanism relating the created members of the multiverse to one another so they are not in fact causally disjoint, otherwise, there is no reason why any probability law (Gaussian normal, for example) should be obeyed. We will return to this point later.

Finally, the relevant integrals may diverge. In that case, assigning mean values or averages for physical quantities in an ensemble of universes is problematic (see Chapter 1 and references therein).

3.3.3 The Anthropic subset

The expression (3.6) can be used in conjunction with the distribution function $f(m)$ of galaxies to determine the expected number of civilizations arising in the whole ensemble:

$$N_{\text{life}}(E) = \int f(m) * N_g * N_S * f_S * f_p * n_e * f_l * f_i * \pi \qquad (3.14)$$

(which is a particular case of (3.12) based on (3.6)). An anthropic ensemble is one for which $N_{\text{life}}(E) > 0$. If the distribution function derives from a probability function, we may combine the probability functions to get an overall anthropic probability function - for an example see Weinberg (2000a,b) discussed below, where it is assumed that the probability for galaxy formation is the only relevant parameter for the existence of life. This is equivalent to assuming that $N_S * f_S * f_p * n_e * f_l * f_i > 0$.

This assumption might be acceptable in our physically realized Universe, but there is no reason to believe it would hold generally in an ensemble because these parameters will depend on other ensemble parameters, which will vary.

3.3.4 Problems With Infinity

When speaking of multiverses or ensembles of universes – possible or realized – the issue of infinity often crops up. Researchers often envision an *infinite* set of universes, in which all possibilities are *realized.*

The question of infinities has a long history, and is still the focal point of a substantial amount of philosophical debate. Aristotle was the first one who explicitly distinguished between a potential and an actual infinite. While the potential infinite can be given an axiomatic and algebraic formulation ('every element is followed by more elements') this was not so clear for the actual infinite. Since Aristotle the possible existence of actual infinities has been subject to a heated debate.

At this point it is useful to distinguish between a mathematical or platonic and a physical or realized infinite. It is well-known that the proof of Cantor's theorem requires the actuality of infinite sets on a mathematical level and hence the whole axiomatic set theory is based on the mathematical actuality of infinite sets.

As is pointed out by Harré (see Robinson and Harré (1964)) "for mathematicians the only fruitful way to preceed is to choose 'actual' rather than 'potential' infinities". However, these actual infinities in mathematics can be thought of as infinities of possibilities (see Gijsbers or Thompson (1999)) in contrast to actual physical (or really existing) infinities.

The existence of actual physical infinites is much more controversial and one unsettled question intimately related is whether past or future time can be infinite. While Smith (1987) argues for a possible infinite past many philosophers hold the view that the past should be finite (see Whitrow (1966), Craig(1979), Huby(1971), Conway(1974)).

The question which primarily concerns us here is whether there can really be an infinite set of really existing universes? We suggest that, on the basis of well-known philosophical arguments, the answer is No. The common perception that this is possible arises from not taking seriously enough the difficulties associated with this profoundly difficult concept. Because we can assign a symbol to represent 'infinity' and can manipulate that symbol according to specified rules, we assume corresponding entities can exist in practice. This is highly questionable.

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-recognized 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). Others (see Spitzer 2000 and Stoeger 2003 and references therein) have further pointed out that realized infinite sets are not constructible – there is no procedure one can in principle implement to complete such a set – they are simply incompleteable. But, if that is the case, then "infinity" cannot be arrived at, or realized, in a real physical setting. On the contrary, the concept itself implies its inability to be realized! This is why for example a realized 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 realized set, or actualizing it[4].

Thus, it is important to recognize 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 recognized, that there has been a breakdown

[4]When contemplating mathematical concepts, it is debatable as to whether a procedure for construction is needed. But we are talking physics, and the issue is precisely whether the concept is realizable.

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 traverzing 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 realized 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 realized 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). We argue that the theoretically possible infinite space sections of many cosmologies at a given time are simply unattainable in practice - they are a theoretical idea that cannot be realized. Future infinite time also is never realized: 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 is simply not constructible in terms of events or instants of time, besides being conceptually indefinite.[5]

We emphasize that the problem with infinity is not primarily physical, in the usual sense – it is primarily a conceptual or philosophical problem with the idea of "realized infinity". Infinity as it is mathematically conceived and used is not the sort of property that can be physically realized, like a definite number can. As emphasized above, it means "indefinitely large", or "continues without limit". The same problem of a realized infinity may be true in terms of the supposed ensembles of universes. It is difficult enough conceiving of an ensemble of many 'really existing' universes that are totally causally disjoint from our own, and how that could come into being, particularly given two important features. Firstly, specifying the geometry of a generic universe requires an infinite amount of information because the quantities in $\mathcal{P}_{\text{geometry}}$ are fields on space-time, in general requiring specification at each point (or equivalently, an infinite number of Fourier coefficients) - they will almost always not be algorithmically compressible. This greatly aggravates all the problems regarding infinity and the ensemble. Only in highly symmetric cases, like the FLRW solutions, does this data reduce to a finite number of parameters. One can suggest that a statistical description would suffice, where a finite set of numbers describe the statistics of the solution, rather than giving a full description. Whether this suffices to adequately describe an ensemble where 'all that can happen, happens' is a moot point. We suggest not, for the simple reason that there is no guarantee that all possible models will obey any known statistical description. That assumption is a major restriction on what is assumed to be possible.

Secondly, many universes in the ensemble may themselves have infinite spa-

[5]One way out would be, as quite a bit of work in quantum cosmology seems to indicate, to have time originating or emerging from the quantum-gravity dominated primordial substrate only "later". In other words, there would have been a "time" or an epoch before time as such emerged. Past time would then be finite, as seems to be demanded by philosophical arguments, and yet the timeless primordial state could have lasted "forever", whatever that would mean. This possibility avoids the problem of constructibility.

tial extent and contain an infinite amount of matter, with the paradoxical conclusions that entails (Ellis and Brundrit 1979). To conceive of physical creation of an infinite set of universes (most requiring an infinite amount of information for their prescription, and many of which will themselves be spatially infinite) is at least an order of magnitude more difficult than specifying an existent infinitude of finitely specifiable objects.

The phrase 'everything that can exist, exists' implies such an infinitude, but glosses over all the profound difficulties implied. One should note here particularly that problems arise in this context in terms of the continuum assigned by classical theories to physical quantities and indeed to space-time itself. Suppose for example that we identify corresponding times in the models in an ensemble and then assume that *all* values of the density parameter occur at each spatial point at that time. Because of the real number continuum, this is an uncountably infinite set of models – and genuine existence of such an uncountable infinitude is highly problematic. But on the other hand, if the set of realized models is either finite or countably infinite, then almost all possible models are not realized – the ensemble represents a set of measure zero in the set of possible universes. Either way the situation is distinctly uncomfortable. However, we might try to argue around this by a discretization argument: maybe differences in some parameter of less than say 10^{-10} are unobservable, so we can replace the continuum version by a discretized one, and perhaps some such discretization is forced on us by quantum theory. If this is the intention, then that should be made explicit. That solves the 'ultraviolet divergence' associated with the small-scale continuum, but not the 'infrared divergence' associated with supposed infinite distances, infinite times, and infinite values of parameters describing cosmologies.

3.4 Ensembles of FLRW Universes

Having established the broad set of issues concerning multiverses that we believe need to be addressed, we shall for the remainder of this chapter limit ourselves to the FLRW sector $\mathcal{M}_{FLRW}$ of the ensemble of all possible universes $\mathcal{M}$ in order to illustrate these issues[6]. We assume the family considered is filled with matter components characterized by a γ-law equation of state, and mainly restrict our attention to their cosmological parameters, although full consideration of anthropic issues would be characterized by including all the other parameters. Our descriptive treatment will consider FLRW universe domains (whether a true multiverse or separate domains in a single space-time) as distinct but with common physical characteristics.

3.4.1 Properties of FLRW models

FLRW models are homogeneous and isotropic models described by the metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left(\frac{r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right), \tag{3.15}$$

[6]Many discussions implicitly suggest that this is the whole possibility space, as they only consider FLRW models as possibilities. However these clearly form a very small subspace of all geometrical possibilities.

where $d\Omega^2 = d\vartheta^2 + \sin^2(\vartheta)d\varphi^2$ denotes the line element on the two-dimensional unit sphere, $a(t)$ is the scale-factor, and

$$k = \begin{cases} 1 & \text{for closed models} \\ 0 & \text{for flat models} \\ -1 & \text{for open models} \end{cases}$$

is the normalized curvature. The FLRW model is completely determined by k and the scale-factor $a(t)$, which incorporates the time-evolution and is obtained from the Einstein-Field equations together with the matter description.[7]

Assuming gravity is described by the Einstein field equations, the evolution of FLRW models is described by the Friedmann equation

$$H^2(\Omega - 1) = \frac{k}{a^2}, \tag{3.16}$$

where $H \equiv \dot{a}/a$ (a dot denotes differentiation with respect to proper time) is the Hubble parameter and Ω the density parameter. We restrict our discussion to models with only a cosmological constant Λ and one matter component which obeys a γ-law equation of state, i.e., its pressure p and density ρ are related by $p = (\gamma - 1)\rho$, where γ is constant. This specification of parameters $p_3(i)$ includes in particular the case of dust ($\gamma = 1$) and radiation ($\gamma = 4/3$). The total density parameter is

$$\Omega = \Omega_m + \Omega_\Lambda, \tag{3.17}$$

where the matter density parameter is $\Omega_m \equiv \frac{\kappa\rho}{3H^2}$ and the vacuum-energy density parameter is $\Omega_\Lambda \equiv \frac{\Lambda}{3H^2}$ (representing a cosmological constant). These form the parameters $p_4(i)$.

The second time derivative of the scale factor is determined by the Raychaudhuri equation

$$2q = (3\gamma - 2)\Omega_m - \Omega_\Lambda, \tag{3.18}$$

where $q \equiv -\frac{\ddot{a}}{aH^2}$ is the dimensionless deceleration parameter. The matter evolution is given by the energy-conservation equation

$$\dot{\rho} = -3\gamma\rho H \tag{3.19}$$

or equivalently by

$$\dot{\Omega}_m = \Omega_m H(q + 1 - 3\gamma). \tag{3.20}$$

Besides the normalized curvature k there are two constants of motion, $\chi \equiv \kappa\rho a^{3\gamma}/3$ and the cosmological constant Λ. Given these parameters, the dynamical evolution is determined from the implied initial conditions: $\{a(t_0), \Omega_{i0}, \gamma_i, k\} \Rightarrow a(t)$.

3.4.2 Parametrising FLRW models

In order to define a FLRW ensemble we need a set of independent parameters which uniquely identify all possible models. We want to consider all possible FLRW models with the same physical laws as in our universe, but possibly

[7]The way the tetrad description given above relates to FLRW universes is described in detail in Ellis and MacCallum (1968); the standard coordinates given here are more convenient if one discusses only the FLRW models.

different coupling constants. There is then one set of parameters $p_2(i)$ which defines the gravitational "physics" of the model in terms of the coupling constants – for simplicity let us only consider the gravitational constant G here – and further sets $p_5(i)$, $p_4(i)$ which identify the geometry and matter content of the actual model, and which are related to each other via the Einstein field equations.

Among the various options there are two particularly useful parametrizations. Ehlers and Rindler (1989) developed a parametrization in terms of the observable density parameters (they also include a radiation component) and the Hubble parameter. With $\Omega_k \equiv \frac{k}{H^2 a^2}$ the Friedmann equation becomes

$$\Omega_m + \Omega_\Lambda - 1 = \Omega_k. \tag{3.21}$$

The curvature parameter Ω_k determines $k = \mathrm{sign}(\Omega_k)$. For $k \neq 0$ the scale-factor, and hence the metric (3.15), is determined by

$$a^2(t) = \frac{k}{H^2\Omega_k} = \frac{k}{H^2(\Omega_m + \Omega_\Lambda - 1)},$$

while for $k = 0$ its value is unimportant because of scale-invariance in that case. Hence any *state* is completely described by Ω_m, Ω_Λ, and H.

In order to parametrize the models rather than the states, we need to select one particular time t_0 for each model at which we take the above parameters as representative parameters $\Omega_{m0}, \Omega_{\Lambda 0}$, and H_0 for this model.[8] We note that this time t_0 can be model dependent because not all models will reach the age t_0.

All big-bang FLRW models start with an infinite positive Hubble parameter whose absolute value reaches or approaches asymptotically a minimum value $H_{\min}$. Hence we could define the time t_0 as the time when the model first takes a certain value $H_0(p_I) > H_{\min}(p_I)$, where p_I represents the model parameters. One particular choice of $H_0(p_I)$ is given by

$$H_0(p_I) \equiv \exp(H_{\min}(p_I))$$

On the other hand, by setting $H_0(p_I)$ = constant and excluding all models which never reach this value one finds easily a parametrization of all models which reach this Hubble value during their evolution.

While above choice of parameters give a convenient parametrization in terms of observables which covers closed, flat, and open models, it is disturbing that for each model an arbitrary time has to be chosen. This also leads to a technical difficulty, because the parameters $\{H_0, \Omega_{m0}, \Omega_{\Lambda 0}\}$ are subject to the constraint $H_0 = H_0(p_I)$.

For these reasons it is often convenient to use a set of parameters which are comoving in the state-space, i.e., parameters which are constants of motion. As mentioned above, for open and closed models such a set is given by the matter constant χ, the cosmological constant Λ, and the normalized curvature constant k. For flat models one can rescale the scale factor, which allows us to set $\chi = 1$.

These parameters are related to the observational quantities by (for $k = \pm 1$)

$$\chi = \frac{\Omega_m H^2}{(H^2|1 - \Omega_m - \Omega_\Lambda|)^{3\gamma/2}} \quad \text{and} \quad \Lambda = 3H^2\Omega_\Lambda.$$

[8]Hence, in general different models correspond to the same values of Ω_m and Ω_Λ, depending on the value of H.

The evolution of these models through state space is illustrated here in terms of two different parametrizations of the state space, see Figures 1a and 1b. For a detailed investigation of these evolutions for models with non-interacting matter and radiation, see Ehlers and Rindler (1989).

3.4.3 The possibility space

The structures defined so far are the uniform structures across the class of models in this possibility space, characterized both by laws of physics (in particular General Relativity) and by a restricted class of geometries. It is clear that universes in a multiverse should be able to differ in at least some properties from each other. We have just characterized the geometrical possibilities we are considering. The next question is, which physical laws and parameters can vary within the ensemble, and which values can they take? For this simplified discussion let us just assume that only the gravitational constant G and the cosmological constant Λ (which also qualifies as a model parameter) are variables, with the ranges[9] $G \in [0, \infty)$ and $\Lambda \in (-\infty, \infty)$. However, if we consider "all that is possible " within this restricted class of FLRW models, maybe we should consider $G \in (-\infty, \infty)$. There is still considerable uncertainty as to the nature of an ensemble even within this restricted context. Whatever is chosen here defines the set of possibilities that can arise.

3.4.4 The measure

For a complete probabilistic description of an ensemble we need not only a distribution function P, but also a measure π for the parameter space (see Section 3.1). The information entropy

$$S \equiv -\int \mathrm{d}x P(x) \log\left(\frac{P(x)}{\mu(x)}\right) \tag{3.22}$$

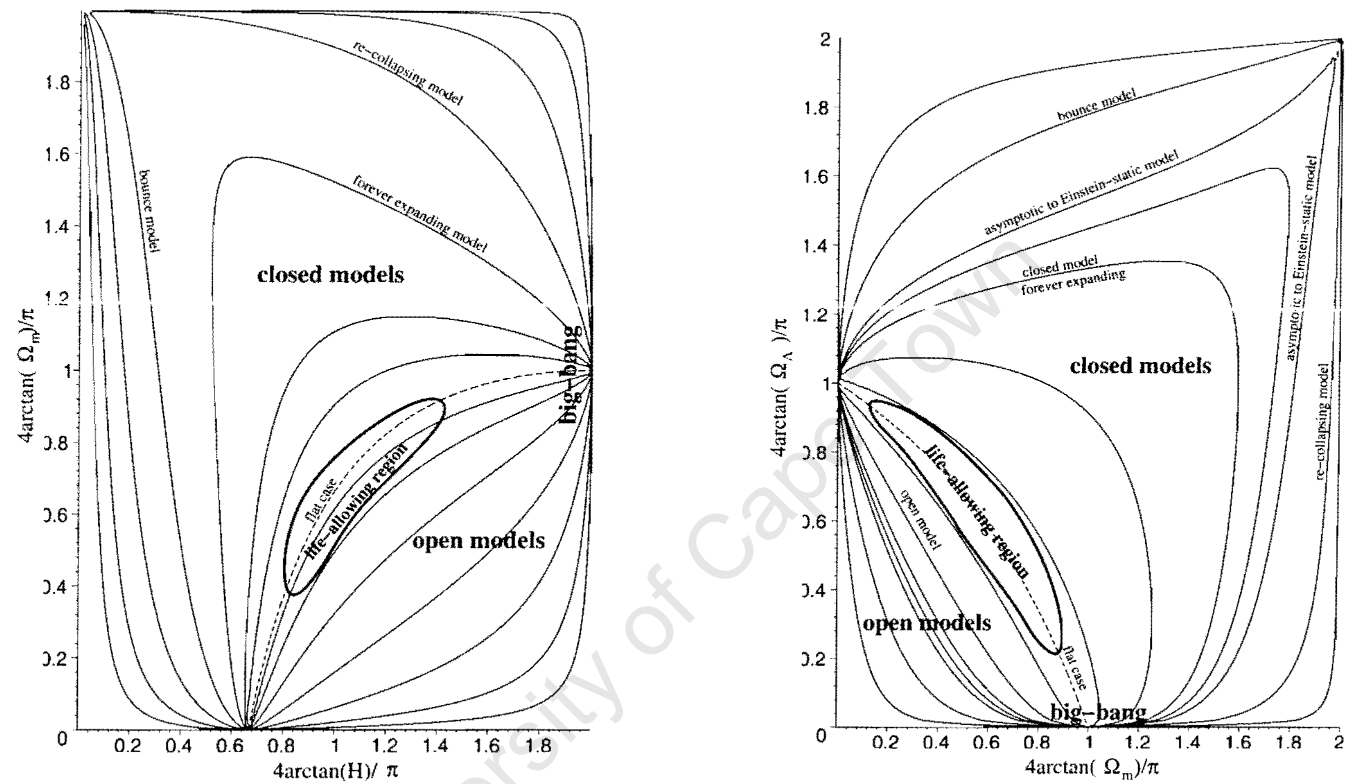
is then maximized for the probability distribution equal to this measure, representing the state of minimal knowledge.

Without knowledge of the creation mechanism it is impossible to determine this measure with certainty. Nevertheless, we might ask what our best guess for such a measure should be in a state of minimal information, where only a certain set of independent parameters, describing the ensemble, and their ranges are known.

The only known method for constructing such a measure is Jaynes' principle. Its application to FLRW models with γ-law equation of state has been discussed in Chapter 1.

There are two important points to note. Firstly the measure is derived from the chosen set of parameters. Generally a different choice of parameters yields a different minimum-information measure, predicting another maximum-entropy distribution function. Let us consider the example of an ensemble of dust-FLRW models. The different open and closed models are most conveniently parametrized by the constants of motion, which are given by the cosmological

[9] It is worth noting that when $G = 0$, we *do* obtain FLRW solutions to the Einstein-Field equations: the Milne universe which is effectively empty – no gravity effective.



(a) $\Omega_m - H$

(b) $\Omega_\Lambda - \Omega_m$- plane

Figure 3.1: Phase-space diagrams showing the evolution of FLRW-models (as given by the Friedmann equation) in the $\Omega_m - H$ and $\Omega_\Lambda - \Omega_m$ -plane for $\Lambda = 1$. Models can evolve in both directions along the lines. Models which start with a big-bang originate on the right in figure (a) (corresponding to $H = \infty$). Models which reach $H = 0$ (in a finite time) reverse their direction of evolution, i.e., they follow the same line back-wards. These are the re-collapsing models. The limiting case to the forever-expanding models is given by the model which approaches asymptotically the Einstein-static universe.

constant Λ and $\chi \equiv a\rho^{3\gamma}$, where ρ is the energy density. It was shown in Chapter 1 that this leads to the minimum-information measure

$$\pi \propto \frac{\mathrm{d}\Lambda \mathrm{d}\chi}{\sqrt{\chi}}. \tag{3.23}$$

Equation (1.18) expresses this measure in $\Omega_{m0} - \Omega_{\Lambda 0}$-parametrization for the subset of all open and closed dust ($\gamma = 1$) big-bang models which reach a certain Hubble parameter H_0 at a time t_0 during their evolution, i.e. in our notation here

$$\pi \propto \sqrt{\frac{\Omega_{m0}}{|\Omega_0 - 1|^{3/2}}} \left| \frac{1}{\Omega_{m0}} - \frac{3/2}{\Omega_0 - 1} \right| \mathrm{d}\Omega_{m0} \mathrm{d}\Omega_{\Lambda 0},$$

with $\Omega_{\Lambda 0} \leq 1 + \Omega_{m0}/2$. On the other hand, as mentioned above, there is a convenient parametrization for this particular subset of models (Ehlers and Rindler 1989) in terms of the observables Ω_{m0} and $\Omega_{\Lambda 0}$ (in Ehlers and Rindler (1989) an additional radiation component was also included). Using this parametrization yields the minimum information measure

$$\pi \propto \frac{1}{\sqrt{\Omega_{m0}}} \mathrm{d}\Omega_{m0} \mathrm{d}\Omega_{\Lambda 0},$$

which is clearly different from the above result.

Secondly, the measure is in general non-normalisable and hence there is no normalisable maximum-entropy distribution. Without additional information we are not able to calculate certain probabilities. Since it seems questionable whether there will ever be additional information about the ensemble of universes available, one has to accept that certain questions will have no well defined probabilities.

It should be mentioned that we encounter similar problems when we want to find a probability measure for physical parameters like the gravitational constant G. Let us assume that G can take any non-zero positive value. Jaynes' principle then suggests the probability measure $\pi_G \propto \frac{\mathrm{d}G}{G}$. On the other hand, if we decide to use [10] $m = \sinh(G)$ as our parameter, then we find the different measure

$$\pi = \frac{\mathrm{d}m}{m} = \frac{\cosh(G)}{\sinh(G)} \mathrm{d}G.$$

3.4.5 Distribution Functions on $\mathcal{M}_{\mathrm{FLRW}}$

Now, having properly parametrized $\mathcal{M}_{\mathrm{FLRW}}$ and defined a measure on it, we can represent particular multiverses by giving distribution functions over the parameter-space (as discussed in Section 3.4). Given a distribution function f it determines the number of universes in a small parameter-interval by

$$\mathrm{d}N = f(p_I)\pi,$$

which is invariant under a change of parametrization. Hence it is the combination of measure and distribution function which is of importance.

[10] However, it should be noted that any power ($m = G^n, m \in {}^{+}, n \neq 0$) and logarithmic relationship ($m = \ln(G), m \in$) leads to the same measure $\mathrm{d}\mu_G$.

While distribution functions can be parametrized by any set of coordinates over the possibility space, we need different distribution functions for different possibility spaces. For example, if universes in a multiverse *must* have a common value for the gravitational constant G then distribution functions must not depend on G.

It is clear that a particular distribution function can be expressed in any set of coordinates. Obviously there is a vast set of possible distribution functions. We want to examine some particular examples.

Firstly, one could have a distribution function which is constant over the parameter-space. The actual ensemble then really depends on the measure and is the maximum-entropy distribution (which maximizes (3.22)[11]). If we choose the observational quantities $H, \Omega_m, \Omega_\Lambda$ to represent the model and allow for different values of the gravitational coupling constant G this would be

$$f(H, \Omega_m, \Omega_\Lambda, G) = \text{const.}$$

for all allowed values of the stated parameters. On the other hand, if we choose the constants of motion as coordinates in possibility space

$$f(k, \chi, \Lambda, G) = \text{const.}$$

The probability $P_{\mathcal{A}}$ to find a universe in a certain parameter-region $\mathcal{A}$ is given by

$$P_{\mathcal{A}} = \frac{\int_{\mathcal{A}} f(k, \chi, \Lambda, G)\pi.}{\int f(k, \chi, \Lambda, G)\pi.},$$

where the integral in the denominator extends over the whole possibility space. For many distribution functions, like for the above constant distribution function together with (3.23), this expression is not well defined. Let us assume that the measure is non-integrable, i.e., non-normalisable, and that we have a constant distribution function If the set $\mathcal{A}$ does not include any point of non-integrability of the measure then $P_{\mathcal{A}} = 0$, if it includes all points of non-integrability then $P_{\mathcal{A}} = 1$, but if it includes only some of the non-integrabilities then $P_{\mathcal{A}}$ is not well defined.

Of course, the above expression might be integrable for all sets $\mathcal{A}$ given a good distribution function. For instance the distribution function

$$f(k, \chi, \Lambda) \propto \exp(-\chi - \Lambda^2)$$

together with the measure (3.23) is integrable everywhere. On the other hand

$$f(k, \chi, \Lambda) \propto \frac{\exp(-\Lambda^2)}{\sqrt{\chi + 1}}$$

diverges for $\chi \to \infty$. A distribution function can also introduce an additional divergence, for instance

$$f(k, \chi, \Lambda) \propto \frac{\exp(-\Lambda^2)}{\sqrt{\chi}}$$

is non-integrable at $\chi = 0$ and $\chi \to \infty$.

[11]In (3.22) $P(x)dx$ corresponds to $f(p_I)\pi$.

A multiverse might contain a finite or infinite countable number of universes. In these cases the distribution function contains Dirac δ -functions, e.g.,

$$f(k,\chi,\Lambda) = \begin{cases} \sqrt{\chi}5\delta^2(\chi-3,\Lambda-.5) & \text{for } k=1 \\ 2\delta(\Lambda-.7) & \text{for } k=0 \\ \sqrt{\chi}3\delta^2(\chi-.1,\Lambda-2) & \text{for } k=-1 \end{cases}$$

which represents a multiverse which contains 5 copies of closed FLRW models with $\chi = 3$ and $\Lambda = .5$, etc. The distribution function

$$f(k,\chi,\Lambda) = \begin{cases} \sqrt{\chi}\sum_{i=1}^{\infty}\delta(\chi-1/i,\Lambda-1) & \text{for } k=1 \\ 0 & \text{for } k\neq 1 \end{cases}$$

represents an ensemble with a countably infinite number of universes – all are closed with $\Lambda = 1$ and one for each $\chi = 1/i$. Similarly one could imagine an ensemble of 10^7 copies of our universe, which would be represented by the distribution function

$$f(k,\chi,\Lambda) = 10^7\sqrt{\chi}\delta_k^{k_0}\delta(\chi-\chi_0)\delta(\Lambda-\Lambda_0),$$

where k_0, χ_0, Λ_0 represent the parameter values for our universe. This is very unlikely in terms of a generating mechanism, but for ensembles without generating mechanisms it is as likely as any other possibility. If a multiverse is "tested" by its prediction that our universe is a likely member, then such an ensemble should be the most satisfying one – but then we might just as well be happy with one copy, i.e., just our universe.

Similar distribution functions determine the distribution of physical parameters like the gravitational constant G. For example with $G \in \mathbf{R}$ the minimum information measure is $d\mu_G = dG$ and

$$f(G) = \exp(-G^2)$$

gives a Gaussian distribution around $G = 0$. If, on the other hand $G \in (0,\infty)$ the measure is dG/G and $f(g) = G\exp(-G)$ would be an example of a distribution function.

One can imagine various types of distributions, e. g., a Gaussian distribution in G or in H_0, or in the other parameters. But, in order to establish these in a non-arbitrary way, we need a theory of how this particular ensemble is selected for from all the other possible ones.

3.4.6 The problem of infinities again

Even within the restricted set of FLRW models, one of the most profound issues is the problem of realized infinities: if all that is possible in this restricted subset happens, we have multiple infinities of realized universes in the ensemble. First, there are an infinite number of possible spatial topologies in the negative curvature case, so an infinite number of ways that universes which are locally equivalent can differ globally. Second, even though the geometry is so simple, the uncountable continuum of numbers plays a devastating role locally: is it really conceivable that FLRW universes actually occur with *all* values independently of both the cosmological constant and the gravitational constant, and also all values of the Hubble constant at the instant when the density parameter takes the value 0.97? This gives 3 separate uncountably infinite aspects of the ensemble of universes that is supposed to exist. The problem would be allayed if space-time is quantized at the Planck level, as suggested for example by loop quantum gravity. In that case one can argue that all physical quantities also are quantized, and the uncountable infinities of the real line get transmuted into finite numbers in any finite interval – a much better situation. We believe that this is a physically reasonable assumption to make, thus softening a major problem for many ensemble proposals. But the intervals are still infinite for many parameters in the possibility space. Reducing the uncountably infinite to countably infinite does not in the end resolve the problem of infinities in these ensembles. It is still an extraordinarily extravagant proposal.

3.4.7 The anthropic subset

We can identify those FLRW universes in which the emergence and sustenance of life is possible at a broad level[12] – the necessary cosmological conditions have been fulfilled allowing existence of galaxies, stars, and planets if the universe is perturbed, so allowing a non-zero factor $\Pi = P_{\rm gal} * R * f_S * f_p * n_e$ as discussed above. These are indicated in the figures above (anthropic universes are those intersecting the regions labelled "life allowing"). The fraction of these that will actually be life-bearing depends on the fulfilment of a large number of other conditions represented by the factor $F = f_l * f_i$, which will also vary across a generic ensemble, and the above assumes this factor is non-zero.

3.5 On the origin of ensembles

Ensembles have been envisaged both as resulting from a single causal process, and as simply consisting of discrete entities. We discuss these two cases in turn, and then show that they are ultimately not distinguishable from each other.

3.5.1 Processes Naturally Producing Ensembles

Over the past 15 or 20 years, many researchers investigating the very early universe have proposed processes at or near the Planck era which would generate a really existing ensemble of expanding universe domains, one of which is our own observable universe. In fact, their work has provided both the context and stimulus for our discussions in this chapter. Each of these processes essentially selects a really existing ensemble through a generating process from a set of possible universes, and often lead to proposals for a natural definition of a probability distribution on the space of possible universes. Here we briefly describe some of these proposals, and comment on how they fit within the framework we have been discussing.

Andrei Linde's (1983, 1990) chaotic inflationary proposal (see also Linde (2003) and references therein) is one of the best known scenarios of this type. The scalar field (inflaton) in these scenarios drives inflation and leads to the generation of a large number of causally disconnected regions of the Universe.

[12]More accurately, perturbations of these models can allow life – the exact FLRW models themselves cannot do so.

This process is capable of generating a really existing ensemble of expanding FLRW-like regions, one of which may be our own observable universe region, situated in a much larger universe that is inhomogeneous on the largest scales. No FLRW approximation is possible globally; rather there are many FLRW-like sub-domains of a single fractal universe. These domains can be very different from one another, and can be modelled locally by FLRW cosmologies with different parameters.

Linde and others have applied a stochastic approach to inflation (Starobinsky 1986, Linde, *et al.* 1994, Vilenkin 1995, Garriga and Vilenkin 2001, Linde 2003), through which probability distributions can be derived from inflaton potentials along with the usual cosmological equations (the Friedmann equation and the Klein-Gordon equation for the inflaton) and the slow-roll approximation for the inflationary era. A detailed example of this approach, in which specific probability distributions are derived from a Langevin-type equation describing the stochastic behaviour of the inflaton over horizon-sized regions before inflation begins, is given in Linde and Mezhlumian (2003) and in Linde *et al.* (1994). The probability distributions determined in this way generally are functions of the inflaton potential.

This kind of scenario suggests how overarching physics, or a "law of laws" (represented by the inflaton field and its potential), can lead to a really existing ensemble of many very different FLRW-like regions of a larger Universe. However these proposals rely on extrapolations of presently known physics to realms far beyond where its reliability is assured. They also employ inflaton potentials which as yet have no connection to the particle physics we know at lower energies. And these proposals are not directly observationally testable – we have no astronomical evidence the supposed other FLRW-like regions exist. Thus they remain theoretically based proposals rather than established fact. There additionally remains the difficult problem of infinities: eternal inflation with its continual reproduction of different inflating domains of the Universe is claimed to lead to an infinite number of universes of each particular type (Linde, private communication). How can one deal with these infinities in terms of distribution functions and an adequate measure? As we have pointed out above, there is a philosophical problem surrounding a realized infinite set of any kind.

Finally, from the point of view of the ensemble of all possible universes often invoked in discussions of multiverses, all possible inflaton potentials should be considered, as well as all solutions to all those potentials. They should all exist in such a multiverse, which will include chaotic inflationary models which are stationary as well as those which are non-stationary. Many of these potentials may yield ensembles which are uninteresting as far as the emergence of life is concerned, but some will be bio-friendly. The price of this process for creating anthropically favourable universe regions is the multiplication of realized infinities, most of which will be uncountable (for example the parameters in any particular form of inflaton potential will take all possible values in an interval of real numbers).

3.5.2 Probability distributions for the cosmological constant

Weinberg (2000) and Garriga and Vilenkin (2001) derive a probability distribution for the cosmological constant in the context of an ensemble of regions generated in the same inflationary sequence via the action of a given inflaton potential where the cosmological constant is given by the potential energy of a scalar-field. In multi-domain universes, where spatial variations in a scalar-field cause different regions to inflate at different rates, the cosmological constant should be distributed according to some probability distribution $P(\rho_\Lambda)$. During inflation the scalar field undergoes randomization by quantum fluctuations, such that later on its values in different regions are distributed according to "the length" in field space (Garriga and Vilenkin 2002). This leads to a probability distribution (or distribution function – the probability distribution is just the normalized distribution function) of values of the vacuum-energy density ρ_Λ in these regions given by

$$P(\rho_\Lambda)\mathrm{d}\rho_\Lambda \propto \frac{\mathrm{d}\rho_\Lambda}{\mid V'(\phi) \mid},$$

where $V(\phi)$ is the inflaton potential, and the prime signifies differentiation with respect to the inflaton ϕ.

It has been suggested (Vilenkin 1995, Weinberg 1997) that the way the probability distribution for existence of galaxies depends on the cosmological constant can be approximated by

$$P_{gal} = N(\rho_\Lambda)P(\rho_\Lambda) \tag{3.24}$$

where $N(\rho_\Lambda)$ is the fraction of baryons that form galaxies. The requirement of structure formation as a pre-requisite for life places strong anthropic constraints on the domains in which observers could exist; these constraints must be satisfied in the really existing universe.

Let us first note that galaxy formation is only possible for a narrow range around $\rho_\Lambda = 0$ (Weinberg 2000). It has been shown that anthropic restrictions demand $\rho_\Lambda \lesssim 10^{-28}\frac{\mathrm{g}}{\mathrm{cm}^3}$ (Barrow and Tipler 1986, Kallosh and Linde 2002, Garriga and Vilenkin 2002). Consequently the anthropic selection factor $N(\rho_\Lambda)$ is sharply peaked and vanishes for $|\rho_\Lambda| > \rho_{\Lambda\,\max}$ for some $\rho_{\Lambda\,\max}$, which is of the same order of magnitude as the observed cosmological constant. In scalar-field models $P(\rho_\Lambda)$ is in direct relation to the *a priori* distribution of the scalar-field fluctuations and it has been argued (Weinberg 2000) that for a wide class of potentials the variations of $P(\rho_\Lambda)$ over the anthropically allowed range (where $N(\rho_\Lambda) \neq 0$) should be negligible. Nevertheless, as has been shown in (Vilenkin-Garriga 2002) this is not always the case, in particular for power-law potentials $V(\phi) = \phi^n$ with $n > 1$ one finds an integrable divergence at $\rho_\Lambda = 0$.

It is clear that a similar relation to (3.24) should hold for multiverses in the wider sense. Nevertheless, one could imagine multiverses containing universes with and without scalar-field, or with different potentials. Hence we cannot link the distribution of the cosmological constant to that of the scalar-field in a unique way, and there is a vast choice for possible *a priori* probability distributions for the cosmological constant. Let us assume that the cosmological constant is a remnant of some underlying (unknown) theory and as such might be restricted to some domain of values. Depending on this domain one finds different possible minimum information measures, which result from Jaynes' principle. If the domain is given by all real numbers then the (non-normalisable) measure will be constant. If on the other hand the domain is given by all positive real numbers then the minimum-information measure gives an (non-normalisable) *a priori* probability distribution proportional to $1/\rho_\Lambda$ (Kirchner and Ellis 2003). In this case the divergence is located inside the anthropically allowed region and is non-integrable. For this case the expectation value vanishes, i.e.,

$$\bar{\rho}_\Lambda = \frac{\int_0^{\rho_{\Lambda\,\max}} \rho_\Lambda \frac{1}{\rho_\Lambda} \mathrm{d}\Lambda}{\int_0^{\rho_{\Lambda\,\max}} \frac{1}{\rho_\Lambda} \mathrm{d}\Lambda} = 0$$

and we fail to explain the observed non-zero value of the cosmological constant.

An interesting alternative is given by allowing the cosmological constant to take values in the domain $R^+ \cup \{0\}$ (e.g., if the cosmological constant prediction is given by a quadratic term). The minimum information-measure is then proportional to $1/\sqrt{\rho_\Lambda}$. Again there is a divergence in the anthropically allowed region, but this time it is integrable. The expectation value becomes

$$\bar{\rho}_\Lambda = \frac{\int_0^{\rho_{\Lambda\,\max}} \rho_\Lambda \frac{1}{\sqrt{\rho_\Lambda}} \mathrm{d}\Lambda}{\int_0^{\rho_{\Lambda\,\max}} \frac{1}{\sqrt{\rho_\Lambda}} \mathrm{d}\Lambda} = \frac{1}{3}\rho_{\Lambda\,\max}.$$

3.5.3 The existence of regularities

Consider now a genuine multiverse. Why should there be any regularity at all in the properties of universes in such an ensemble, where the universes are completely disconnected from each other? If there are such regularities and specific resulting properties, this suggests a mechanism creating that family of universes, and hence a causal link to a higher domain which is the seat of processes leading to these regularities. This in turn means that the individual universes making up the ensemble are not actually independent of each other. They are, instead, products of a single process, as in the case of chaotic inflation. A common generating mechanism is clearly a causal connection, even if not situated in a single connected space-time – and some such mechanism is needed if all the universes in an ensemble have the same class of properties, for example being governed by the same physical laws or meta-laws.

The point then is that, as emphasized when we considered how one can describe ensembles, any multiverse with regular properties that we can characterize systematically is necessarily of this kind. If it did not have regularities of properties across the class of universes included in the ensemble, we could not even describe it, much less calculate any properties or even characterize a distribution function.

Thus in the end the idea of a completely disconnected multiverse with regular properties but without a common causal mechanism of some kind is not viable. There must necessarily be some pre-realization causal mechanism at work determining the properties of the universes in the ensemble. What are claimed to be totally disjoint universes must in some sense indeed be causally connected together, albeit in some pre-physics or meta-physical domain that is causally effective in determining the common properties of the multiverse.

Related to this is the issue that we have emphasized above, namely where does the possibility space come from and where does the distribution function come from that characterizes realized models? As emphasized above, we have to assume that some relevant meta-laws pre-exist. We now see that we need to explain also what particular meta-laws pre-exist. If we are to examine 'all that might be, exists', then we need to look at the ensemble of all such meta-laws and a distribution function on this set. We seem to face an infinite regress as we follow this logic to its conclusion, and it is not clear how to end it except by arbitrarily calling a stop to this process. But then we have not looked at all conceivable possibilities.

3.6 Testability and Existence

Finally, the issue of evidence and testing has already been briefly mentioned. This is at the heart of whether an ensemble or multiverse proposal should be regarded as physics or as metaphysics.

3.6.1 Evidence and existence

Given all the possibilities discussed here, which specific kind of ensemble is claimed to exist? Given a specific such claim, how can one show that this is the particular ensemble that exists rather than all the other possibilities?

There is no direct evidence of existence of the claimed other universe regions, nor can there be any, for they lie beyond the visual horizon; most will even be beyond the particle horizon, so there is no causal connection with them; and in the case of a true multiverse, there is not even any possibility of any indirect causal connection of any kind - the universes are then completely disjoint and nothing that happens in any one of them is linked to what happens in any other one.

What weight does a claim of such existence carry, in this context when no direct observational evidence can ever be available? The point is that there is not just an issue of showing a multiverse exists - if this is a scientific proposition one needs to be able to show which specific multiverse exists; but there is no observational way to do this. Indeed if you can't show which particular one exists, it is doubtful you have shown any one exists. What does a claim for such existence mean in this context?

3.6.2 Observations and Physics

The one way one might make a reasonable claim for existence of a multiverse would be if one could show its existence was a more or less inevitable consequence of well-established physical laws and processes. Indeed, this is essentially the claim that is made in the case of chaotic inflation. However the problem is that the proposed underlying physics has not been tested, and indeed may be untestable. There is no evidence that the postulated physics is true in this universe, much less in some pre-existing metaspace that might generate a multiverse. Thus belief in the validity of the claimed physics that could lead to such consequences is just that, a belief - it is based on unproved extrapolation of established physics to vastly beyond where it has been tested. The issue is not just that the inflaton is not identified and its potential untested by any observational means - it is also that, for example, we are assuming quantum field theory remains valid far beyond the domain where it has been tested, and we have faith in that extreme extrapolation despite all the unsolved problems at the foundation of quantum theory, the divergences of quantum field theory, and the failure of that theory to provide a satisfactory resolution of the cosmological constant problem.

3.6.3 Observations and probabilities

The 'doomsday argument' has led to a substantial literature on relating existence of universe models to evidence, based on analysis of probabilities, often using a model of choosing a ball randomly from an urn, and of associated selection effects (see e.g. Bostrom 2002). However usually these models either in effect assume an ensemble exists, or else are content to deal with potentially existing ensembles rather than actually existing ones (see e.g. Olum 2002). That does not deal with the case at hand. One would have to extend those arguments to trying to decide, on the basis of a single ball drawn from the urn, as to whether there was one ball in the urn or an infinite number. It is not clear to us that the statistical arguments used in those papers lead to a useful conclusion in this singular case, which is the case of interest for the argument in this chapter.

In any case, in the end those papers all deal just with observational probabilities, which are never conclusive. Indeed the whole reason for the anthropic literature is precisely the fact that biophilic universes are clearly highly improbable within the set of all possible universes (see e.g. the use of Anthropic arguments as regards the value of Λ referred to in Section 5.2). We are working in a context where large improbabilities are the order of the day. Indeed that is why multiverse concepts were introduced in the first place - to try to introduce some form of scientific explanation into a context where the probabilities of existence of specific universe models preferred by observation are known to be very small.

3.6.4 Observations and disproof

Despite the gloomy prognosis given above, there are some specific cases where the existence of a chaotic inflation (multi-domain) type scenario can be disproved. These are when we live in a 'small universe' where we have already seen right round the universe (Ellis and Schreiber 1986, Lachieze-Ray and Luminet 1995) for then the universe closes up on itself in a single FLRW-like domain and so no further such domains that are causally connected to us in a single connected space-time can exist.

This 'small universe' situation is observationally testable, and indeed it has been suggested that the CBR power spectrum might already be giving us evidence that this is indeed so, because of its lack of power on the largest angular scales (Luminet et al, 2003). This proposal can be tested in the future by searching for identical circles in the CMB sky.

Such tests make it possible to disprove the usual chaotic inflationary scenario, but not a true multiverse proposal, for that cannot be shown to be false by any observation. Neither can it be shown to be true.

Taking an inductive approach to science, we need a method of confirmation in order to be able to update the likelihood of the theory. True multiverse theories can be confirmed only by their predictions about our observable universe — a very small basis for a scientific inquiry. It appears more appropriate to take Karl Popper's view in which it is falsifiability which distinguishes science from metaphysics.

3.7 Conclusion

The introduction of the multiverse or ensemble idea is a fundamental change in the nature of cosmology, because it aims to challenge one of the most basic aspects of standard cosmology, namely the uniqueness of the universe (see Ellis 1991, 1999 and references therein). However previous discussions have not made clear what is required in order to define a multiverse, although some specific physical calculations have been given based on restricted low-dimensional multiverses. The aim of this chapter is to make clear what is needed in order to properly define a multiverse, and then examine some of the consequences that flow from this.

Our fundamental starting point is the recognition that there is an important distinction to be made between possible universes and realized universes, and our main conclusion is that a really existing ensemble or multiverse is not *a priori* unique, nor uniquely defined. It must somehow be selected for. We have pointed out a clear distinction between an ensemble of possible universes $\mathcal{M}$, and an ensemble of really existing universes, which is envisioned as generated by the given primordial process or action of an overarching cosmic principle. These effectively select a really existing multiverse from the possibilities in $\mathcal{M}$, and, as such, effectively define a distribution function over $\mathcal{M}$. Thus, there is a definite causal connection, or "law of laws", relating all the universes in these multiverses. It is this really existing ensemble of universes, *not* the ensemble of all possible universes, which provides the basis for anthropic arguments. Anthropic universes lie in a small subset of $\mathcal{M}$, whose characteristics we understand to some extent. It is very likely that the simultaneous realization of *all* the conditions for life will pick out only a very small sector of the parameter space of all possibilities: anthropic universes are fine-tuned.

The fine-tuning problem is very controversial. Two counter-attacks maintain that there is no fine-tuning problem, so it is not necessary to construct solutions to it by employing the multiverse idea. The first promotes the view that whatever happens will always be unlikely (any hand of cards is as unlikely as any other). Thus, since it is just an example of chance, there is nothing special about a universe that admits life. The counter response is that the existence of life is quite unlike anything else in the physical world – its coming into being is not just like choosing one out of numerous essentially identical hands of cards. It is like being transformed into an entirely different higher level game, and so does indeed require explanation. The second counter-attack argues that inflation explains the current state of the universe, making its apparently unlikely state probable. However, this move is only partially successful, since very anisotropic or inhomogeneous models may never inflate. The counter response is that this does not matter: however small the chances are, if it works just once then that is sufficient to give a model close enough to the standard FLRW cosmological models to be friendly to life. But this does not account for the rest of the coincidences enabling life, involving particle masses and the values of the fundamental constants. Perhaps progress in quantum cosmology will in the future lead to some unique theory of creation and existence that will guide the discussion. At present, uniqueness eludes us.

Among those universes in which the necessary cosmic conditions for life have been fulfilled is the subset of almost-FLRW universes which are possible models of our own observable universe, given the precision of the observational data we have at present. It is, however, abundantly clear that "really existing multiverses" which can be defined as candidates for the one to which our universe belongs are *not* unique, and neither their properties nor their existence is directly testable. The only way in which arguments for the existence of one particular kind of multiverse would be scientifically acceptable is if, for instance, there would emerge evidence (either direct or indirect) for the existence of specific inflaton potential which would generate one particular kind of ensemble of expanding universe domains.

Despite these problems, the idea of a multiverse is probably here to stay with us - it is an important concept that needs exploration and elucidation. Does the idea that 'all that can exist, exists' in the ensemble context provide an explanation for the anthropic puzzles? Yes it does do so. The issue of fine tuning is the statement that the biophilic set of universes is a very small subset of the set of possible universes; but if all that can exist exists then there are universe models occupying this biophilic subspace. However there are the following problems: (i) the issue of realized infinities discussed above, (ii) the problem of our inability to describe such ensembles because we don't know what all the possibilities are, so our solution is in terms of a category we cannot fully describe, and (iii) the multiverse idea is not testable or provable in the usual scientific sense; existence of the hypothesized ensemble remains a matter of faith rather than proof. Furthermore in the end, it simply represents a regress of causation. Ultimate questions remain: Why this multiverse with these properties rather than others? What endows these with existence and with this particular type of overall order? What are the ultimate boundaries of possibility – what makes something possible, even though it may never be realized? In our view these questions - Issues 1 and 2 discussed above - cannot be answered scientifically because of the lack of any possibility of verification of any proposed underlying theory. They will of necessity have to be argued in philosophical terms.

The concept of a multiverse raises many fascinating issues that have not yet been adequately explored. The discussion given here on how they can be described will be useful in furthering this endeavour.

References

[1] Aguirre A 2001 The Cold Big-Bang Cosmology as a Counter- example to Several Anthropic Arguments *Phys.Rev.* **D64** 083508 (*Archive* astro-ph/0106143).

[2] Barrow J D and Tipler F J 1986 *The Cosmic Anthropological Principle* (Oxford: Oxford University Press)

[3] Bennett J, Shostak S and Jakowsky B 2002 *Life in the Universe* (Addison Wesley).

[4] Bostrom N 2002 *Anthropic Bias* (Routledge)

[5] Conway D A 1974 "Possibility and Infinite Time: A Logical Paradox in St Thomas' Third Way" *International Philosophical Quarterly* **14**.

[6] Craig W L 1979 "The Kalam Cosmological Argument" (New York: Harper Row).

[7] Davies P 2003 "Multiverse or Design? Reflections on a 'Third Way"', unpublished manuscript and talk given at the Centre for Theology and the Natural Sciences, Berkeley, California, March 22, 2003.

[8] Deutsch D 1998 *The Fabric of reality: The science of parallel universes - and its implications* (Penguin).

[9] Drake F and Shostak S *Astrobiology: The Search for Life in Space* (Cambridge: Cambridge University Press)

[10] Ehlers J and Rindler W 1989 A phase space representation of Friedmann-Lemaître universes containing both dust and radiation and the inevitability of a big bang *Mon. Not. Roy. Ast. Soc.* **238** 503-521

[11] Ellis G F R 1967 The dynamics of pressure-free matter in general relativity *Journ. Math. Phys.* **8** 1171-1194

[12] Ellis G F R 1971 "Relativistic Cosmology" in *General Relativity and Cosmology, Proc Int School of Physics "Enrico Fermi"* (Varenna), Course XLVII. ed R K Sachs (Academic Press) 104-179

[13] Ellis G F R 1979 The homogeneity of the universe *Gen Rel Grav* **11**, 281-289.

[14] Ellis G F R 1991 Major Themes in the relation between Philosophy and Cosmology *Mem. Ital. Ast. Soc.* **62** 553-605

[15] Ellis G F R 1999 "Before the Beginning: Emerging Questions and Uncertainties" in *Toward a New Millenium in Galaxy Morphology* ed D Block, I Puerari, A Stockton and D Ferreira (Kluwer, Dordrecht), *Astrophysics and Space Science* **269-279** 693-720

[16] Ellis G F R and Brundrit G B 1979 Life in the infinite universe *Qu. Journ. Roy. Ast. Soc.* **20** 37-41

[17] Ellis G F R and Bruni M 1989 A covariant and gauge-free approach to density fluctuations in cosmology *Phys. Rev.* D **40** 1804-1818

[18] Ellis G F R and MacCallum M A H 1969 A class of homogeneous cosmological models *Comm. Math. Phys.* **12** 108-141

[19] Ellis G F R and Schrieber G 1986 *Phys. Lett.* A **115** 97

[20] 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 (*Preprint* gr-qc/9812046)

[21] van Elst H and Ellis G F R 1996 The covariant approach to LRS perfect fluid spacetime geometries *Class. Quant. Grav.* **13** 1099-1127

[22] 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)

[23] Gardner M 2003 *Are Universes Thicker than Blackberries?* (New York: Norton).

[24] Garriga J Vilenkin A 2000 On likely values of the cosmological constant *Phys. Rev.* D **61** 083502 (*Preprint* astro-ph/9908115)

[25] Garriga J and Vilenkin A 2001 Many worlds in one *Phys. Rev.* D **64** 043511 (*Preprint* gr-qc/0102010 and gr-qc/0102090)

[26] Garriga J and Vilenkin A 2002 *Phys. Rev.* D **67** 043503 (*Preprint* astro-ph/0210358)

[27] Gijsbers V "Infinity and the Mind", paper prepared for CS Logica 2.

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

[29] 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

[30] 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 (*Preprint* gr-qc/0210031)

[31] Huby P M 1971 "Kant or Cantor: That the Universe, if real, must be finite in both spaqce and time" *Philosophy* **46** pp 121-132.

[32] Kallosh R, Linde A D 2002 M-theory, cosmological constant and anthropic principle *Phys. Rev.* D **67** 023510 (*Preprint* hep-th/0208157)

[33] Kirchner U and Ellis G F R 2003 A probability measure for FLRW models *Class. Quantum Grav.* **20** 1199-1213

[34] Lachieze-ray M, and Luminet J P 1995 *Phys. Rep.* **254** 135 (*Preprint* gr-qc/9605010)

[35] Leslie J 1996 *Universes* (Routledge).

[36] Lewis D K (2000) *On the Plurality of Worlds* (Oxford: Blackwell).

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

[38] Linde A D 1983 *Phys. Lett.* B **129** 177

[39] Linde A D 1990 *Particle Physics and Inflationary Cosmology* (Chur:Harwood Academic Publishers)

[40] Linde A D 2003 "Inflation, Quantum Cosmology and the Anthropic Principle", to appear in *Science and Ultimate Reality: From Quantum to Cosmos*, honouring John Wheeler's 90th birthday, ed J D Barrow, P C W Davies and C L Harper (Cambridge:Cambridge University Press) (*Preprint* hep-th/0211048)

[41] Linde A D, Linde D A, and Mezhlumian A 1994 *Phys. Rev.* D**49** 1783

[42] Linde A D, and Mezhlumian A 1993 Stationary Universe *Phys. Lett.* B **307** 25-33 (*Preprint* gr-qc/9304015)

[43] Luminet J-P, Weeks J R, Riazuelo A, Lehoucq R, and Uzan J-P 2003 "Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background" to appear in *Nature*

[44] Olum K 2002 The doomsday argument and the number of possible observers *Phil. Q.* **52** 164 (*Preprint* gr-qc/0009081)

[45] Rees M J 2001a *Just Six Numbers: The Deep Forces that Shape the Universe* (Basic Books)

[46] Rees M J 2001b *Our Cosmic Habitat* (Princeton: Princeton University Press).

[47] Rees M J 2001c 'Concluding Perspective' *Archive* astro-ph/0101268.

[48] Robinson G and Harré R 1964 'Infinity' *Proceedings of the Aristotelian Society* **Supp Vol 38** pp 39-68; reprinted in *Infinity* (Dartmouth) ed A W Moore.

[49] Sciama D 1993 *Is the universe unique?* in *Die Kosmologie der Gegenwart* ed G Borner and J Ehlers (Serie Piper)

[50] Smith Q 1987 "Infinity and the Past" *Philosophy of Science* **54** 3 pp 63-74.

[51] Smolin L 1999 *The Life of the Universe* (Oxford:Oxford University Press)

[52] Spitzer R J 2000 Definitions of Real Time and Ultimate Reality *Ultimate Reality and Meaning* **23** (No. 3) 260-276

[53] Starobinsky A A 1986 "Stochastic De Sitter (Inflationary) Stage in the Early Universe" in *Current Topics in Field Theory, Quantum Gravity and Strings* ed H J de Vega and N Sanchez (Heidelberg:Springer) pp 107

[54] Stoeger W R 2003 "What is 'the Universe' that Cosmology Studies", to be published in the Ian Barbour Festschrift

[55] Tegmark M 1998 Is the Theory of Everything Merely the Ultimate Ensemble Theory? *Annal. Phys.* **270** 1-51 (*Preprint* gr-qc/9704009)

[56] Tegmark M 2003 Parallel Universes *Scientific American* (May 2003), 41-51 (*Preprint* astro-ph/0302131)

[57] Thompson I 1999 *Philosphy of Nature and Quantum Reality* (http://www.ph.surrey.ac.uk/~phslit/papers/pnb/node2.html).

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

[59] Vilenkin A 1995 *Phys. Rev. Lett.* **74** 846

[60] Wainwright J and Ellis G F R (Eds) 1996 *The dynamical systems approach to cosmology* (Cambridge:Cambridge University Press)

[61] Weinberg S 1997 in *Critical Dialogs in Cosmology* ed N Turok (Singapore:World Scientific)

[62] Weinberg S 2000 The Cosmological Constant Problems *Preprint* astro-ph/0005265
Weinberg S 2000 A Priori Probability Distribution of the Cosmological Constant *Phys. Rev.* D **61** 103505 (*Preprint* astro-ph/0002387)

[63] Whitrow G T 1966 "Time and the Universe" in *The Voices of Time* ed J T Fraser (George Braziller).

Conclusion

In this thesis I investigated several different aspects of ensembles of universes and multi-domain universes.

The total density parameter Ω_0 of our universe appears to be very close to unity. Going backwards in time it must have been even closer to unity. It is now commonly argued that 'this should be an extremely unlikely case' — the well-known Flatness Problem, which has served as a reason for the need of inflation and motivated the idea of really existing ensembles.

In chapter 1 I suggested new minimum-information probability measures for the space of FLRW models. In contrast to previous investigations I start with a measure over the constants of motion $\chi = \kappa\rho l^{3\gamma}/3$ and the cosmological constant Λ. The measure is then transformed into $\Omega_{\Lambda 0} - \Omega_{\rho 0}$-parametrization. We find a non-integrable divergency around $\Omega_0 = 1$. Since there are no restrictions on the possible values of the cosmological constant the measure gives a flat probability distribution for Λ.

This very interesting and important result has far-reaching consequences for the flatness problem. For any fixed value of Λ the measures predict Ω_0 infinitesimally close to unity — the flatness problem now appears to be rather a parametrization problem. The cosmological density parameters are not constants of motion and (except for $\Omega = 0$) all curves tend towards $\Omega = 1$ in the past. If almost all models originate from $\Omega_0 = 1$, how can we expect a flat distribution around $\Omega_0 = 1$?

The suggested measures express this formally by showing that (for a fixed cosmological constant) at any moment in time still almost all models are infinitesimally close to the flat model.

In chapter 2 spherically symmetric junction surfaces between solutions of the Einstein field equations, an often used model for multi-domain universes, were investigated. The presented approach differs from previous investigations in the choice of coordinates and variables. By introducing new re-scaled time and radial coordinates we construct a coordinate system in which the junction surface is at a fixed radial coordinate, while all coordinates are continuous at the junction surface. The motion of the junction surface is now absorbed into the time dependence of the metric components. It was shown that this approach leads to the same expression for the extrinsic curvature of the junction surface as in the conventional approach (with different coordinate charts on each side).

The junction conditions are of two kinds. Firstly, the curves on the junction surface must measure the same length 'on both sides of the junction', i.e., the tangential metric components must be continuous at the junction surface. Secondly, the Lanczos equation, which relates the jump in extrinsic curvature to the surface-matter content, must be satisfied. In the spherically symmetric case this tensor equation has two independent components. As was well-known, the angular component together with the matching of the tangential metric components relates the surface-matter content to the junction surface motion. The equations take the form of a coupled system of first order differential equations. On the other hand, the time component of the Lanczos equation contains second order derivatives of the junction surface position. It has been known that for certain special cases this equation is identically satisfied. For more general cases (in particular the matching of FLRW sections) it had been suggested that this equation acts as an equation of state for the surface matter, i.e., determining the surface-pressure.

It was shown in section 2.4 that this equation is in fact an identity for all junctions between spherically symmetric sections satisfying the Einstein field equation. Consequently the system needs to be supplemented by a surface equation of state.

In general relativity the energy-density is generally (except for scalar-field cases) positive. Since the junction surface is just an idealised transition region between two models, one would expect that surface-energies should be positive too. This effectively restricts possible junctions. The presented formalism allows the study of restrictions on the possible values of the surface energy-density and a detailed study of possible cases was given. This included analytic expressions for the allowed ranges of the surface-energy density.

A short discussion of timelike junctions without surface-layer was given and it was shown that they can only exist if the metric satisfies some extraordinary constraints. One could speculate that such junctions have to expand or contract at the speed of light. However, these singular cases were excluded in the formalism presented here and should be studied separately.

Of particular cosmological interest are models in which two FLRW sections are joined along a timelike junction surface. These cases were studied in detail in section 2.5 and numerous numerical examples were given. It was shown that in many cases the proper time along the junction is finite. In some of these cases the speed of light was approached asymptotically and for a comoving observer the junction appears to exist forever. In other cases the junction seems to be terminated within some finite proper time on each side. It was speculated that in such cases the junction turns spacelike.

The more general concept of a 'multiverse' was discussed in chapter 3. It was emphasized that there is an important distinction between ensembles of merely possible and actually realized universes — they are linked by a selection and creation process and only the latter allows to explain the fine-tuning of cosmological parameters by anthropic selection.

It was argued that the often used phrase 'all that can exist exists' does not identify a unique ensemble. Furthermore, it is impossible for us to judge what is possible and what not, since we would need to understand the creation mechanism. And what about an ensemble containing universes created by different creation mechanisms? In the end the only way we can handle an ensemble is by extrapolating observed properties of our universe (or universe region) over the whole ensemble. One such example is an ensemble of almost-FLRW models, which was discussed to some detail in subsection 3.4.

When dealing with multiverses the issue of infinities is unavoidable. Is it possible to have infinitely many universes of infinite spatial extend? Even worse, if everything possible should be realized and there are continuous parameters one would need an uncountable 'number' of universes! Is this really conceivable?

Notwithstanding all the ambiguities and uncertainties surrounding the multiverse concept, the intuitive explanations which a multiverse model offers for the appearence of life will always attract interest. However, whether we live in a multiverse or not will remain an unanswered question – probably forever.

Acknowledgements

I would like to thank my supervisor G F R Ellis for the extraordinary support during the work on this thesis. Without his insight and intuition the presented work would have not been possible.

I am particularly indebted to William R Stoeger for his support.

For helpful comments on and discussions about chapter 2 I would like to thank C W Hellaby.

The work with J-P Uzan was a particularly motivating experience and I want to thank him for his generosity and support.

Appendix A

WMAP data and the curvature of space

Inter alia, the high precision WMAP data on Cosmic Background Radiation marginally indicate that the universe has positively curved (and hence spherical) spatial sections. Here we take this data seriously and consider some of the consequences for the background dynamics. In particular, we show that this implies a limit to the number of e-foldings that could have taken place in the inflationary epoch; however this limit is consistent with some inflationary models that solve all the usual cosmological problems and are consistent with standard structure formation theory.

The Wilkinson Microwave Anisotropy Probe (WMAP) has recently provided high resolution Cosmic Microwave Background data [1, 2] of importance for cosmology - indeed they constitute a significant contribution towards the goal of developing precision cosmology. Among the interesting conclusions that have been reached from this data are constraints on the total density parameter of the universe, Ω. The WMAP results interestingly indicate that while the universe is close to being flat ($\Omega \simeq 1$), a closed universe is marginally preferred: $\Omega > 1$ [1]. In particular, with a prior on the Hubble constant, one gets that $\Omega = 1.03 \pm 0.05$ at 95% confidence level, while combining WMAP data with type Ia supernovae leads to $\Omega = 1.02 \pm 0.04$ or to $\Omega = 1.02 \pm 0.02$ respectively without and with a prior on the Hubble parameter. The latter may be regarded as the present best estimate of this parameter. Note also that this tendency to point toward a closed universe is not strictly speaking either new or restricted to the WMAP data. For instance with a prior on the nature of the initial conditions, the Hubble parameter and the age of the universe, analysis of the DASI, BOOMERanG, MAXIMA and DMR data [3, 4] lead to $\Omega = 0.99 \pm 0.12$ but to $\Omega = 1.04 \pm 0.05$ if one takes into account only the DASI, BOOMERanG and CBI data, both at 1σ-level. The Archeops balloon experiment [5] points toward $\Omega = 1.00^{+0.03}_{-0.02}$ with data from the HST and ST a prior on the Hubble constant but to values in the range $\Omega = 1.16^{+0.24}_{-0.20}$ from combined CBR data alone and to $\Omega = 1.04^{+0.10}_{-0.12}$ from combined CBR and supernova data. The improved precision from WMAP is clear.

The confirmation of the existence of Doppler peaks and their respective locations tends to confirm the inflationary paradigm [6], as does existence of an almost scale invariant power spectrum. Furthermore a nearly flat universe, argued to be predicted generically by most inflationary models [7], is usually seen as a further evidence. Since

$$\delta(t) \equiv \Omega(t) - 1 = \frac{K}{a^2H^2} \tag{A.1}$$

where $a(t)$ is the scale function and $H(t) = \dot{a}/a$ the Hubble parameter, the WMAP constraints on the curvature of space imply that the curvature density parameter, $\Omega_K(t) \equiv K/a^2H^2$, has been smaller than a few percent from the time the largest observable scale, $k_{\rm max} \sim a_0H_0$, crossed the horizon during inflation up to today. It is then a good approximation to neglect the curvature from the time $k_{\rm max}$ crossed the horizon to today when performing structure formation calculations, mainly because $\Omega_{k_{\rm max}} = \Omega_{\rm today}$ (see Fig. 1). The constraints on the curvature of the universe thus confirm validity of calculations underlying the tests of inflation performed up to now that assume a flat universe and focus on the properties of the perturbations (spectral indices, tensor modes, statistics, etc.).

Nevertheless, when $K \neq 0$ (which is highly favoured over the case $K = 0$ by probability considerations: the latter are of measure zero in the space of Robertson-Walker universes in most measures) and matter is described by a γ-equation of state, Ω_K is driven toward 0 exponentially during inflation. The curvature term therefore was necessarily dominant at early enough times, if inflation undergoes a sufficient number of e-foldings, mainly because the curvature term behaves as a^{-2} and will tend to dominate over any kind of matter having an equation of state stiffer than $-1/3$, and thus over a slow rolling scalar field or a cosmological constant, in the early inflationary era.

We argue that the study of dynamics of the background for such models can give interesting insight into the inflationary era and its dynamics prior to $k_{\rm max}$ horizon exit. Working backward in time from the present, we aim to emphasize the constraints that arise on the inflationary models from the WMAP data. In doing so, we take seriously the indication that a value of Ω slightly larger than 1 is favoured. During a slow-roll inflationary phase, the universe is described by an almost de Sitter spacetime with the $K = +1$ scale factor

$$a(t) \propto \cosh Ht \tag{A.2}$$

instead of the exponential growth obtained in the flat case. It follows that the number of e-foldings between the onset and the end of inflation cannot be arbitrarily large, when we take present day cosmological parameters into account. The curvature enhances the effect of inflation in its early stage so that there exists a turn-around point in the past. It follows [8, 9] that the number of e-folds during inflation is bounded by a maximum possible number of e-folds, $N_{\rm max}$, related to the present day density parameter Ω.

Assuming that the inflation is driven simply by a cosmological constant, the scale factor of the universe in this epoch is given by

$$a = a_{\rm min}\cosh(\lambda t) \tag{A.3}$$

with $\lambda \equiv \sqrt{\Lambda/3}$ and $t = 0$ corresponds to the minimum expansion of the universe. It follows that $\delta = \Omega - 1 = 1/\sinh^2(\lambda t)$. The time t_* at which

$k_{\max} = a_0 H_0$ crosses the horizon is thus obtained to be

$$\lambda t_* = \mathrm{Arcsinh}(1/\sqrt{\delta_0}). \tag{A.4}$$

The number of e-folds between $t = 0$ and t_* is simply

$$\Delta N_{\max} = \ln \frac{a_{k_{\max}}}{a_{\min}} \tag{A.5}$$

that is

$$\Delta N_{\max} = \frac{1}{2} \ln \left(1 + \frac{1}{\delta_0} \right). \tag{A.6}$$

It typically ranges from 2.3 to 1.5 when δ_0 varies from 0.01 to 0.05 (see Fig. 2).

This argument also holds for negatively curved universes, in which case, the scale factor behaves as $a \propto \sinh \lambda t$. In that case, there will not be any maximal number of e-foldings but the universe starts off from a singularity. This will enhance the Transplanckian problem [10].

Another way to look at the existence of a maximum number of e-folds in the $K = +1$ case, is to relate the number of e-foldings during inflation to the later history of the universe [8, 9]. In particular, it is fruitful to estimate the number of e-folds allowed before the largest observable scales exit the horizon during inflation. During inflation, the comoving Hubble length aH is decreasing, while it is increasing at all further times. A given comoving scale, k, crosses the horizon when $k = a(t_k)H(t_k) \equiv a_k H_k$. The number of e-folds, $N(k)$, between when this scale crosses the horizon and the end of inflation is obtained to be

$$N(k) = \ln \left(\frac{a_{\mathrm{end}} H_{\mathrm{end}}}{a_k H_k} \right). \tag{A.7}$$

The determination of $N(k)$ requires that we compute

$$\mathcal{R} \equiv \frac{a_0}{a_{\mathrm{end}}} \tag{A.8}$$

which requires a complete history of the evolution of the universe from the end of inflation to today. If one assumes that the universe is matter dominated from the end of (slow roll) inflation to reheating, radiation dominated from reheating to equality and matter dominated up to now [1], then [11, 12]

$$\begin{aligned} N(k) \quad \simeq \quad & 62 - \ln \left(\frac{k}{a_0 H_0} \right) - \ln \left(\frac{10^{16}\,\mathrm{GeV}}{V_k^{1/4}} \right) \\ & + \frac{1}{4} \ln \frac{V_k}{V_{\mathrm{end}}} - \frac{1}{3} \ln \left(\frac{V_{\mathrm{end}}^{1/4}}{\rho_{\mathrm{reh}}^{1/4}} \right) - \ln h, \end{aligned} \tag{A.9}$$

h being the Hubble parameter in unites of 100 km.s^{-1}/Mpc. For our purpose, it will be more convenient to relate $N(k)$ to $\mathcal{R}$ which can easily be found to be given by

$$-\ln \mathcal{R} \simeq -66 - \frac{1}{3} \ln \left(\frac{V_{\mathrm{end}}^{1/4}}{\rho_{\mathrm{reh}}^{1/4}} \right) + \ln \left(\frac{10^{16}\,\mathrm{GeV}}{V_{\mathrm{end}}^{1/4}} \right). \tag{A.10}$$

[1]The results will be only marginally affected by a late-time cosmological constant as indicated by the supernova data.

It follows that

$$N(k) \simeq 128 - \ln\mathcal{R} - \ln\left(\frac{k}{a_0 H_0}\right) -2\ln\left(\frac{10^{16}\,\mathrm{GeV}}{V_k^{1/4}}\right) - \ln h. \tag{A.11}$$

The cosmic microwave background roughly probe scales from 10 Mpc (i.e. of the order of the thickness of the last scattering surface) to 10^4 Mpc (i.e. of the order of the size of the observable universe) while galaxy surveys can probe the range 1 Mpc to 100 Mpc.

Concerning the characteristic scales involved in the problem, we can assume that $V_k \sim V_{\mathrm{end}}$ as long as slow rolling holds. The reheating temperature can be argued to be larger than $\rho_{\mathrm{reh}}^{1/4} > 10^{10}$ GeV to avoid the gravitino problem [13] and may be pushed in the extreme case to 10^3 GeV, i.e. just before the electroweak transition so, that baryogenesis can take place. The amplitude of the cosmological fluctuations (typically of order 2×10^{-5} on Hubble scales) roughly implies that $V_{\mathrm{end}}^{1/4}$ is smaller than a few times 10^{16} GeV and, for the same reason as above, has to be larger than 10^{10} GeV in the extreme case.

This implies that the number of e-foldings has approximately to lie between 50 and 70, which is the order of magnitude also required to solve the horizon and flatness problem when $K = 0$ [7], but it can be lowered to 25 in the extreme case of thermal inflation [14], which we are not considering here.

In the following, we restrict our analysis to the case where reheating takes place just at the end of inflation, i.e. $\rho_{\mathrm{reh}} = V_{\mathrm{end}}$. It has been shown [8, 9] that, in that case, the existence of a positive curvature limits the number of e-folds to

$$N_{\max}(\mathcal{R}, \delta_0) = \ln\sqrt{\frac{\alpha}{\delta_0}} \tag{A.12}$$

with

$$\alpha = (1 + \delta_0 - \Omega_{\mathrm{rad0}})\mathcal{R} + \Omega_{\mathrm{rad0}}\mathcal{R}^2 \tag{A.13}$$

where $\delta_0 \equiv \Omega_0 - 1$. $\Omega_{\mathrm{rad0}} \simeq 4.17 \times 10^{-5} h^{-2}$ is the radiation density parameter today.

This allows us to estimate the number of e-foldings allowed prior to the time $k_{\max}$ crossed the horizon,

$$\Delta N = N_{\max}[\mathcal{R}(V_{\mathrm{end}}^{1/4}), \delta_0] - N(k_{\max}, V_{\mathrm{end}}^{1/4}) \tag{A.14}$$

where we will let δ_0 vary in the range $0.01 - 0.05$ (see Fig. 2). It can be checked, as expected from Eq. (A.12) that ΔN depends slightly on $V_{\mathrm{end}}^{1/4}$.

We see that the allowed number of e-foldings are compatible with the requirements of structure formation. 'Flatness' is of course solved to the accuracy represented by δ_0. The scenarios sketched here do not lead to as small a value of δ_0 as is often supposed - and thereby is compatible with the best-fit WMAP data. However, for the mean value of the curvature, one obtains that

$$\delta_0 = 0.02 \Longleftrightarrow \Delta N_{\max} = 1.97. \tag{A.15}$$

With standard parameter values this gives the maximum possible number of efoldings as about$62 + 1.97 \simeq 64$ - compatible with the estimates given above, but not nearly as large as suggested in some models of inflation.

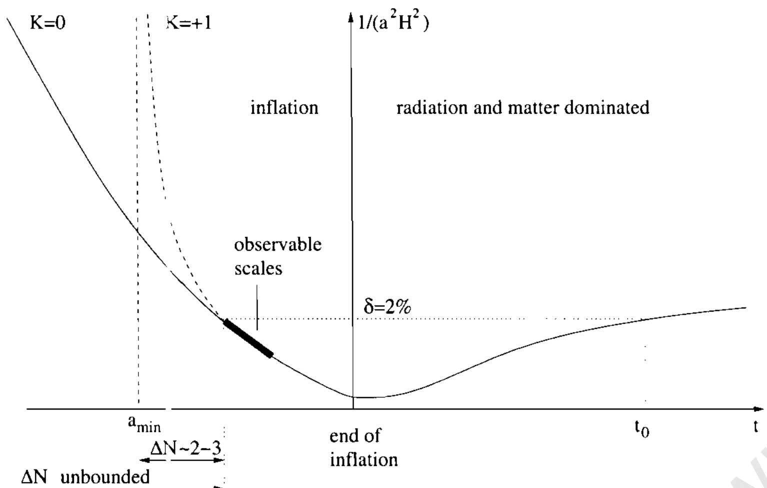


Figure A.1: $\Omega - 1$ as a function of time (schematically). The solid line represents the case of a flat universe ($K = 0$)with exponential inflation while the dashed line depicts the case of a closed universe ($K = +1$). Before the end of inflation aH scales either as $e^{\lambda t}$ ($K = 0$) or $1/\tanh\lambda t$ ($K = +1$) while it scales as $\alpha t^{\alpha-1}$ during matter and radiation dominated eras. The largest observed scale is depicted as well, and one can easily see that $\Omega - 1$ will be smaller than a few percent between the time it crosses the horizon during inflation to today when it reenters the horizon.

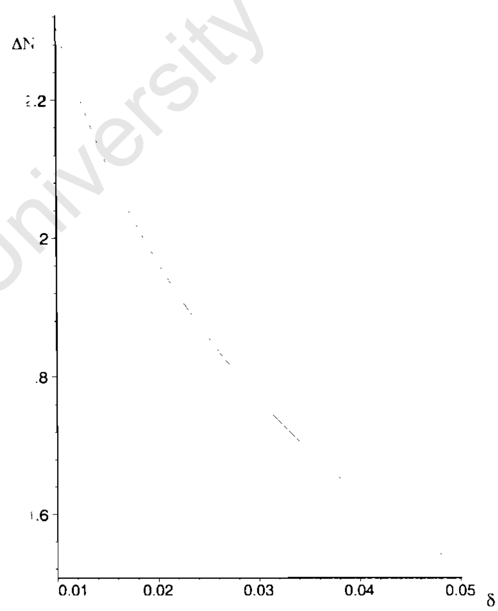


Figure A.2: The maximum number of e-folds allowed before the largest observable scale, $k_{\max}$ exits the horizon during inflation, as a function of δ_0.

In view of the bounds on $N_{\max}$, there are also bounds on the ability of inflation to solve the horizon problem when one considers inflationary dynamics with $K = +1$. Roughly speaking, the horizon problem is only solved during the inflationary era itself if the start of the inflationary era is close enough to the throat at $t = 0$ in the expansion (A.3). If inflation starts sufficiently far from that throat, then it does not matter how many e-foldings take place, they cannot provide causal connectivity on large enough scale to solve the horizon problem; rather that connectivity has to be set up in the pre-inflationary era (and is evidenced by a smooth enough small patch that is then expanded to large scales by inflation). However inflation certainly can start near enough the throat to solve the horizon problem during the inflationary era and still lead to $\delta_0 = 0.02$. Indeed this will necessarily be true for models with positive curvature that start slowly from an pre-inflationary Einstein static era [15], either in the form of an emergent or awakening universe. In effect, these models start precisely at the throat in the de Sitter model.

In this note we have argued that the curvature of space may play an important role in the dynamics of inflation in its early stage. We have argued that a maximum number of 1.5 – 2.5 e-folds can take place before the largest scale observed in the universe crossed the horizon during slow-roll inflation. However despite the fact that the curvature term will be smaller than a few percent during all times ranging from the horizon exit of the largest observable mode to its reentry today, a positive curvature term as small as the one indicated by current data will dramatically change the early phase of inflation and hence lead to the limits on the number of e-folds allowed as described above. These restrict the families of inflationary universes compatible with the current data, but do not exclude inflationary models even though the value of δ_0 indicated by WMAP is larger than supposed in most inflationary scenarios.

Note a possible way around the argument. In the case inflation is driven by a scalar field, the existence of a maximal number of e-folds may be avoided if it exits the slow-roll regime. It was however shown that the dynamics of closed universes filled with a scalar field is chaotic [16, 17, 18] and can lead the existence of singularity free universes [19] with periodic [20] or aperiodic [16] trajectories. In that case, again, a positive curvature affects the dynamics of the early phase of inflation and the slow-roll regime has to be left before $N_{\max}$ is reached, which is also at odds with the standard picture.

Let us finally consider further features of spherical universes that may turn to out be useful in understanding the WMAP data. Surprisingly, the WMAP angular correlation function seems to lack signal strength on scales larger than 60 degrees [1, 2]. This may indicate a possible discreteness of the initial power spectrum, as expected e.g. from non-standard topology of spherical space sections. First, the spectrum of the Laplacian in spherical spaces is always discrete [21, 22]. As emphasized in Ref. [23], a non trivial topology is most likely to be detectable in the case of spherical spaces, resulting in reduced power at large angular scales, and examples of the resulting observational effects are discussed in Ref. [24].

The same conclusion cannot be reached for negatively curved universe. It is unlikely that we could detect a compact hyperbolic space because, for an observer at a generic point, the topology scale is comparable to the curvature radius or longer. The possibility of a detection would require the observer to sit near a short closed geodesic [25].

In conclusion, if the indication that Ω_0 differs from unity at a level of a few percent is confirmed then it may imply important results concerning the dynamics of the early stage of our universe. The $K = +1$ background de Sitter model (3) differs fundamentally from the scale-free de Sitter model in the $K = 0$ frame - not least in terms of being geodesically complete. The WMAP observation suggest the former may be the appropriate model to use in investigating early inflationary dynamics.

References

[1] Bennett C L *et al* 2003 *Preprint* astro-ph/0302207

[2] Spergel D N *et al* 2003 *Preprint* astro-ph/0302209

[3] Sievers J L *et al* *Preprint* astro-ph/0205387

[4] Netterfield C B *et al* 2002 *Astrophys. J.* **571** 604

[5] Benoît A *et al* 2002 *Preprint* astro-ph/0210306
Benoît A *et al* 2002 *Preprint* astro-ph/0210305

[6] Peiris H V *et al* 2003 *Preprint* astro-ph/0302225
Barger V, Lee H S, and Marfatia D 2003 *Preprint* hep-ph/0302150

[7] Guth A H 1981 *Phys. Rev.* D **23** 347
Linde A D 1982 *Phys. Lett.* B **108** 389
Hawking S W 1982 *Phys. Lett.* B **115** 295
Starobinski A A 1982 *Phys. Lett.* B **117** 175
Albrecht A and Steinhardt P J 1982 *Phys. Rev. Lett.* **48** 1220

[8] Ellis G F R *et al* 2002 *Gen. Rel. Grav.* **34** 1445

[9] Ellis G F R *et al* 2002 *Gen. Rel. Grav.* **34** 1461

[10] Martin J and Brandenberger R 2001 *Phys. Rev.* D **65** 123501
Brandenberger R and Martin J 2001 *Mod. Phys. Lett.* A **16** 999
Lemoine M *et al* 2002 *Phys. Rev.* D **65** 023510

[11] Lyth D and Liddle A 1993 *Phys. Rept.* **231** 1

[12] Lidsey J E *et al* 1997 *Rev. Mod. Phys.* **69** 373

[13] Sarkar S 1996 *Rep. Prog. Phys.* **59** 1493

[14] Lyth D H and Stewart E D 1996 *Phys. Rev.* D **53** 1784
ibid. Phys. Rev. Lett. **75** 201

[15] Ellis G F R and Maartens R 2002 *Preprint* gr-qc/0211082

[16] Page D N 1984 Class. Quantum Grav. **1** 417

[17] Toporensky A V 2000 *Preprint* gr-qc/0011113

[18] Cornish N and Shellard E P S 1998 *Phys. Rev. Lett.* **81** 3571

[19] Parker L and Fulling S A 1973 *Phys. Rev.* D **7** 2357
Starobinsky A A 1978 *Sov. Astron. Lett.* **4** 82

[20] Hawking S W 1984 in *Relativity, Group and Topology II* ed B S de Witt and R Stora (Amsterdam: North Holland).

[21] Lehoucq R, Uzan J-P and Weeks J *Kodai Math. J.* (to appear) *Preprint* math.SP/0202072
Lehoucq R *et al* 2002 *Class. Quantum Grav.* **19** 4683

[22] Riazuelo *et al* 2002 *Preprint* astro-ph/0212223

[23] Weeks J, Lehoucq R and Uzan J-P *Class. Quantum Grav.* (to appear) (*Preprint* astro-ph/0209389)

[24] Riazuelo *et al* "Cosmic microwave background maps of multi-connected spherical spaces" (in preparation).

[25] Weeks J 2003 *Int. J. Mod. Phys.* A (submitted)
Gomero G I, Rebouças M J and Tavakol R 2001 *Class. Quantum Grav.* **18** 4461
Gomero G I, Rebouças M J, and Tavakol R 2002 *Int. J. Mod. Phys.* A **17** 4261