Renormalization Group, coarse-graining and critical phenomena in cosmology

Thesis presented by
Camilla M. Piotrkowska
for the degree of
"Doctor Philosophiae"
Supervisor: Prof. D.W.Sciama
SISSA/ISAS
Astrophysics sector
Academic year 1993/94

Acknowledgements

I would like to take this opportunity to express my gratitude to all the people with whom I collaborated and who got me to the point, where I could commence my research at SISSA.

It should go without saying that I owe a great deal to Dennis Sciama. I want to thank him for the way in which he so ungrudgingly gives his time and wisdom. I very much enjoyed many stimulating discussions with him. He also made possible my participation at a few conferences during my studies at SISSA.

I would like to express my deepest gratitude to Mauro Carfora and acknowledge our collaboration. I would like to thank Mauro also for a careful reading of a part of the manuscript and for much useful criticism and invaluable advice, encouragement and teaching, which enabled me to arrive at a better understanding of many mathematical subtleties behind the problem treated in this thesis. I shall always count among my greatest of good fortunes the possibility of our joint research. Extra special thanks to Mauro for the drawings included in chapter three of this thesis. Also I want to express my thanks to Annalisa Marzuoli and Mauro for the warm hospitality during my visit to Pavia which I thoroughly enjoyed.

I also owe a great debt to Roustam Zalaletdinov whose unique blend of physical intuition and drive for theory gave me a sense of what a large, unifying framework for cosmology might look like and helped to articulate the importance of asking the right questions. I thank him for his generous advice going beyond explanations asked, for carefully reading a part of the manuscript and suggesting revisions, as well as for our collaboration.

Extra special thanks to George Ellis for suggesting to me the research subject on the averaging problem which, somewhat unexpectedly, turned out to be so extremely exciting, keeping me on my toes and providing, in the end, more fun than I could ever have hoped for. More than anyone else, George has made me understand that doing research is a tough game though it has some rewards too. Therefore I want to thank him for his friendly support, unstinting encouragement, our collaboration and many illuminating and enjoyable discussions on cosmology and philosophy.

I want also to thank Marco Bruni for interesting comments and discussions. I enjoyed a great deal a collaboration with him on perturbations.

I am also very grateful to Juan Pérez-Mercader for enlightening discussions with him which provided me with a new impetus for the research into critical phenomena in General Relativity and a deeper understanding of various aspects of Renormalization Group. I want to thank him for the enjoyable collaboration and warm hospitality extended, while I was visiting L.A.E.F.F. in Madrid, where some of key aspects of our collaboration were initiated.

Thanks are also due to Andrzej Krasiński who let me have his review work prior to a publication and to Roustam Zalaletdinov (once more) for his help in collecting the bibliography.

I benefited from interactions with many people during my studies at SISSA, but it is not possible to name them all here. In particular though, I want to thank Masahiro Morikawa, Toshifumi Futamase, Malcolm MacCallum, Robin Tucker, Giampiero Esposito, Lee Smolin, Antonio Lanza, Y'aqub Anini, Jorge Zanelli, Colin McIntosh, Paul Haines, Hugo Morales and Noboru Takeuchi. I thank also Ewa Szuszkiewicz for her comments upon reading parts of the manuscript and correcting the English.

To Jorge Devoto I owe a debt of thanks for his invaluable help and support.

And special thank you to all the people who spent hours discussing my thesis with me, reading papers, for sharing with me their opinions, never-ending encouragement and helping me discover the true depth of things...

For the financial support I thank Ministero Italiano per l'Università e la Ricerca Scientifica e Tecnologica.

Cap. LXXVII , Donde se cuentan mil zarandajas, tan impertinentes, como necesarias al verdadero entendimiento de esta grande Historia.

Cervantes, Don Quixote.

Abstract

The central problem studied in this thesis is, broadly speaking, the issue of coarse-graining in GR approximations, and *the effect of averaging on the field equations.* The important observation made is that there are some *smoothing* procedures implicit in the standard, homogeneous and isotropic Friedman-Lemaître-Robertson-Walker cosmological models. The point is that if such effects are not allowed for, we may actually be using the wrong field equations in cosmology. There has been recently an increased effort in this direction with some interesting results, as for example that the coarse-graining effects could be non-negligible in the context of affecting the age of the universe.

The central idea explored at length in this thesis (see chapter three) is the possibility of applying the Renormalization Group (RG) concepts in gravitation to tackle the averaging problem. Research presented in this thesis produced results as follows: In section 3.6 an explicit smoothing-out procedure for inhomogeneous cosmologies is introduced. This approach is implemented in a "$3+1$" formalism and invokes the coarse-graining arguments, provided and supported by the real-space RG methods in an analogy with lattice models of Statistical Mechanics. One of the results obtained is a re-interpretation of the Ricci-Hamilton flow in terms of a RG flow, thereby providing the Ricci-Hamilton flow with a physical meaning and showing how the averaging problem is rooted in the geometry. Moreover, block variables are introduced and the recursion relations written down explicitly, making thus possible a study of the system's critical behaviour. The *criticality* is discussed and it is argued that it may be related to the formation of sheet-like structures in the universe. Moreover, the explicit expression for the renormalized Hubble constant is proposed. A discussion follows of the consequences of this approach in cosmology and astrophysics.

Finally in chapter four the evolution of perturbations is studied in the dust-radiation FLRW universe model. This is done in the framework of dynamical systems theory which seems well suited to this purpose. The evolution of density perturbations is presented in the form of phase diagrams. Some scales are discovered that are over-damped.

A number of ideas of relevance for future research are summarized in chapter five.

Table of Contents

1 Introduction

1.1 Theoretical foundations of Cosmology

The aim of cosmology is an investigation of the structure of space-time in the largest possible scale. One can think of two different directions of doing this [133]:

– "top-down" way, whereby one assumes *a priori* some postulates about the large scale of the universe and tries to deduce local physics (Milne, Bondi and Gold); and

– "bottom-up" way, whereby one extrapolates the accepted local physics as far as possible, in order to guess the global structure of space-time. This is indeed the present paradigm and it also involves some, more or less, aprioristic assumptions about the Universe as a whole, or cosmology itself.

In order that cosmology can be concerned with the global structure of space-time one has to make some non-empirical assumptions. In fact, it turns out that non-local assumptions are involved in every empirical prediction of standard physics. From the point of view of field theory, local extrapolations into the future can be valid only under the so-called no-interference assumption (see e.g. [91]), that there are no signals generated by a distant event which at the next moment could affect the system under investigation. If one wants to interpret modern cosmological observations a theory of space-time is needed. The conclusions one arrives at about the properties of the universe, essentially depend on an integration of some differential equations over a "large scale" domain of the space-time manifold. Usually, as a rule it is the simplest cosmological models that are implicitly used in practice (highly symmetrical), which obscure the above fact.

General Relativity (GR henceforth) is the best classical theory of gravitation we have.

The field equations of GR are correct to a high order of accuracy for the solar system (see e.g. [252]) and relatively small binary systems (PSR 1913+16, 4U1820−30 and PSR 0655+64) [137, 260, 15, 65], i.e. to the distance scales less than $50\,AU$. Nevertheless, GR is applied to cosmic structures such as galaxies, clusters of galaxies, superclusters and ultimately the whole universe, which are typically $10^6 \div 10^{15}$ times larger than the distances over which the theory has been tested with a high accuracy. In the theoretical studies of gravitational lensing [32] GR is assumed to be the correct theory of gravity, but doing so GR is not being tested explicitly. In fact, some authors questioned the validity of GR over large distances in the context of dark matter problem [223][1]. As far as cosmology is concerned, the most significant result obtained with the help of classical tests is the support given to the standard model against the steady-state cosmology.

The whole subject of experimental gravitation is rather subtle. This is so finally due to the fact that GR is a general covariant theory, and due to this the very concept of observables is quite involved (see e.g. [24] for still excellent review). The strategy that has to be employed in all the measurements is to use concrete physical objects as clocks and spatial references, and these objects cannot be taken as independent from the dynamics of the system. The importance of the Lorentz metric is not to be underestimated as it is its existence that makes it possible to perform space and time measurements in different reference frames. The geometry of space-time determines the laws of measurement. The Lorentz metric further allows us to construct relativistic mechanics and relativistic electrodynamics.

In GR the space-time manifold, formally determined by the matter content and boundary conditions, with its various geometric structures has a very special status as a "primitive element" of the theory. It has to be a differentiable, at least class C^4 [64], Hausdorff and connected manifold, if it is to model a physical world. Locally the manifold is (pseudo-)euclidean - concept not directly related to the curvature - in the sense that every point of it lies in a neighbourhood which can be coordinated. Moreover different points of it have coordinates (four real numbers) related in a continuous way. At every point of the

[1] One can have doubts about the validity of GR when spatial separations are of the order of l_{Planck}, but that would be relevant to the very early stages of the standard model with which we are not concerned in the present thesis.

manifold there is a rich algebraic structure (vectors, tensors, etc.) representing various physical fields, which otherwise is introduced quite arbitrarily. This structure is built using the local differentiable properties of the manifold. Apart from the differential one there is also a geometric structure. The pseudo-riemannian manifold is in effect taken to model the space-time in the relativistic theory of gravitation. It is only through the Einstein equations that the physical fields acquire a non-local character.

The space-time evolves during the evolution of the universe and the mathematical structure of GR space-time at different times should reflect the changes. Possibly we should even allow for a breakdown of the smoothness of the manifold. In the early universe close to the initial singularity, each physical property implied global consequences, the distinction between local and non-local ceases to be clear-cut. The singularity itself (initial) is related to the space-time being geodesically incomplete. The singularity theorems [131] show that geodesic incompleteness of the considered space-time cannot coexist with some of its properties, like e.g. compactness, non-existence of closed time-like curves, existence of Cauchy hypersurface, which are non-local ones. Every physically acceptable model of the universe can and should contain singularities [131] (e.g. the Schwarzschild solution), they had better constitute a set of measure zero! (for a comprehensive treatment of the mathematical structure of space-time, see e.g. [95]). There are different kinds of singularities and some are "more global" then others.

The principle of equivalence identifies a riemannian space-time with the gravitational field. Einstein's equations provide a relation between the sources of the gravitational field and that field itself. They are the prescription of how matter and any energy-carrying field determines the geometry of space-time and, *a fortiori* the gravitational field. The gravitational field is represented by the space-time metric $g_{\alpha\beta}$ which becomes a dynamical field coupled to matter, through the field equations. In this sense the gravitational field is contained in the geometry, as knowledge of the metric is sufficient (for the pseudo-riemannian manifold) for the determination of all properties of our manifold. The information about the sources of the gravitational field is contained in the energy-momentum tensor $T_{\alpha\beta}$. In order to specify it we should know in advance the space-time structure of the manifold itself, which we cannot, until we have solved the Einstein equations[2]

$$E_{\alpha\beta} \equiv R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = T_{\alpha\beta}. \tag{1.1}$$

GR contains in fact no general recipe of how the energy-momentum tensor should be constructed, given a particular physical situation. Indeed, it can only be "constructed" on the basis of pre-relativistic physics with no account, consequently, of the geometrization view. The generally adopted procedure in GR is to resort to field theories, covariant in the sense of restricted relativity (special). A covariant form of the tensor is then brought up by taking ordinary derivatives into covariant derivatives. Notwithstanding, the requirement of covariance alone is not enough to construct $T_{\alpha\beta}$.

The energy-momentum tensor is postulated as a functional of the metric tensor and other relevant state variables. For example, in the case of a perfect-fluid, the energy-momentum functional has the form

$$T_{\alpha\beta} = (\mu + p) u_\alpha u_\beta + p g_{\alpha\beta}, \tag{1.2}$$

where, p is the scalar pressure, μ the (proper) energy density and the velocity field u_α of the fluid is normalized, $u_\alpha u^\alpha = -1$. Usually the barotropic equation of state $f(\mu, p) = 0$ is taken to hold.

By solving the Einstein equations we come up explicitly with the energy-momentum tensor and, obviously, the components of the metric tensor. The mathematical problem is complete if we take into account the constitutive equations from outside the theory itself, e.g. the equation of state mentioned above. However, with the Bianchi identities (contracted)

$$T^{\alpha\beta}{}_{;\beta} = 0, \tag{1.3}$$

we can reduce the number of independent equations from ten to six. The remaining four degrees of freedom correspond to the freedom we have in selecting a coordinate system in advance of solving Einstein's equations. (In some special situations the number of coordinate conditions may be more.) The components of the energy-momentum tensor are not independent of this choice, and it is these components that are identified with physically

[2]Natural units are chosen unless evident otherwise; the signature $(- + + +)$ is adopted throughout.

measurable quantities. Here we meet again very important problem of measurability. For example, Eddington [75] has emphasized the distinction, very important in cosmology, between the invariant and relative mass, whose (or rather its associated mass density) value determines in the GR models whether the universe is closed or open. Therefore, when we talk about mass density we have to be careful whether we mean an abstract invariant or a measurable, but coordinate dependent, relative density.

Any theory of gravitation must of course be germane to cosmology and the models of the universe provided by GR have to be consistent (in some sense) with the observational data. Now, one of the biggest difficulties of cosmology is that most of the observational data are theory-dependent, i.e., their meaning can be interpreted by assuming certain theoretical explanations. We interpret what we see in terms of the laws we know, extrapolating them far into the universe - this cannot be avoided. We encounter here a subtle matter of verifiability of the field equations. As far as a theory (any) is concerned, it is clear that evidence whose meaningful interpretation involves its assumptions cannot be used for its verification. Therefore, cosmological solutions should not be considered as large scale verifications of the theory of gravitation, due to the number of other assumptions invoked from outside the theory itself. This also can be looked at as one of the needs calling for a fitting (discussed at some depth later on) between the GR–models and the observations.

Most of the current cosmologies employ GR[3] and as a rule the relativistic cosmologies[4] rely on some form of Cosmological Principle[5], which is usually a smoothing-out hypothesis imposed *a priori* on the distribution of matter in the universe. A well known example is provided by the Friedmann-Lemaître-Robertson-Walker (FLRW henceforth) metric which is usually assumed to describe the real universe. Specifying a restricted family of geometries (e.g. FLRW) and physical behaviour serves to "reduce" the fully general equations of GR, too difficult to handle, into say, a manageable set of differential equations. In this

[3]We are not interested in any other theory of gravitation than GR in this thesis, neither models other than the standard one which by itself assumes the validity of GR.

[4]With a possible exception of the De Sitter model which does not describe the real universe anyway.

[5]A cosmological principle can also be looked at as the criterion of choice by which a certain solution to the field equations is asserted to be a model of the universe.

view, Cosmological Principle is just a working hypothesis leading to the simplest models that are yet acceptable and at the same time operationally functioning. It is fair to say that there is really no alternative since we do not know any solution of Einstein's equations capable of describing a clumpy universe. But one can ask questions about the limits of validity and meaning of FLRW models, since the universe around us is not any ideal, in the sense of being homogeneous and isotropic. *This question was our main objective for research pursued in this thesis.*

The above problem can be posed in terms of the so-called *fitting problem*, namely, the question of how best to fit a smooth FLRW universe to a lumpy reality. This way we recognize that the FLRW description can be valid only in some *averaged* sense. In fact the problem of construction of a physically reliable stress-energy tensor is closely related to the fitting problem. Looked at this way, there is a feedback from the observational cosmology side as our observations are ever more accurate and further reaching. In effect, the whole procedure of confronting the theory and observations, and fitting in particular, could proceed in iterative steps. But the point is still more complicated. Considering a particular solution with a non-vanishing energy-momentum tensor, we can always ask whether the energy-momentum tensor has been correctly constructed or not. It seems that this question on its own is extremely difficult to answer experimentally, if not in principle impossible. This is because, due to the principle of equivalence, GR is a theory of gravitational field and has been empirically confirmed only in this rôle (and as far as solutions of the empty field equations are concerned). Consequently it is not possible to circumscribe the domain of validity of GR from within, because of the non-geometrization of anything other than gravitation. The ideal solution would be to calculate $T_{\alpha\beta}$ once the field equations are solved, instead of postulating it before attempting to solve the equations. In other words, the Einstein field equations only determine the gravitational field corresponding to a given energy distribution and by doing so, they do not provide us with a theoretical description of non-gravitational fields. Certainly, even if the resulting geometrical structure were observed, the question whether the equations have been correctly applied or not, would still be an open question. This disadvantage was already recognised by Einstein himself [80]. He emphasized the non-relativistic nature of any assumed form of $T_{\alpha\beta}$ rather than the possible non-verifiability of the resulting theory. This was one of the reasons that in fact prompted him to work on his unified theory.

With these remarks in mind, it is a matter of great importance to make every effort in order that the foundations from which a cosmological model is obtained are as sound as possible and free from assumptions that may not be warranted. Let us stress at this point that we are obviously not going to be concerned with what might be regarded as a completion of GR in this respect, in the sense of meaningful incorporation of fields into a continuous picture of reality[6], but we will rather try to follow a way enabling to clarify the relation between the homogeneous and isotropic "background" models of the universe and the more detailed fine-grained ones and to extract such a background model.

Having chosen the model, we have in practice at our disposal a procedure for testing the cosmological models. Let us reiterate again the facts that must be kept in mind when we are discussing the universe at large. Cosmological information is obtained along the null geodesics of a pseudo-riemannian space-time, which represent light rays and the very understanding of what is happening is always, to a certain degree, theory-dependent in the sense that the interpretation of cosmological observations is impossible without assuming a working model of space-time. The choice for this model (or a class of models) is eventually made on the basis of postulates, principles, and even philosophical tastes.

The comparison is then made between the relations amongst observables in the model and these same relations implied by observations. This line of approach was paved by a beautiful paper of Kristian and Sachs [163]. Worth mentioning is Kristian's [162] attempt to measure the distortion of images of clusters due to conformal curvature of the universe. This effect is very important, since it is in principle capable of detecting departures from the FLRW geometry in the real universe (it is zero in conformally flat models). The idea is that in a galaxy cluster, the angular distribution of galaxies of each shape should be random, and the anisotropy in their observed images would be determined by the magnetic part of the Weyl tensor.

Further, the problem of the observational basis of cosmology was treated in a series of

[6]In any case, a description of the fields related to the weak and strong interactions would have to be left aside within this approach, since their current description involves ideas which cannot be meaningfully incorporated into the continuous picture.

papers by G. Ellis and collaborators (see [85] and references therein), who discussed in details such problems as: what quantities can be estimated and how from observations, how could observations imply that our universe is FLRW, the practical limitations of observations. The lack of well-defined criteria for acceptance or rejection of the FLRW models was emphasized. In particular, the near-isotropy of the cosmic microwave background (CMB) radiation does not prove the near-isotropy of the matter distribution nor that of the space-time. The general conclusion of Ellis' programme was that the ideal observations on our past light cone are not sufficient to uniquely construct its space-time geometry. In other words, without dynamical equations (field equations) we can only reach a conclusion of consistency of many different cosmological models with the observations. If we take into account the field equations the situation gets much more involved. However, it is possible to ascertain that the maximal observable data set, $D(w_o, z^*)$, is at the same time, the minimal one needed to uniquely determine the geometry of our light cone up to z^* (the redshift of the last scattering surface) [85]. In reality however observations can never be made precise and this aggravates even more the whole problem of fitting.

Clearly the standard cosmological model, the FLRW one, used to describe the real inhomogeneous universe is a very special one and very unlikely at the same time. Interesting, though not surprising, is the fact that the FLRW solutions make up in fact only a set of measure zero in the space of solutions of the Einstein field equations. To ask, how well and how bad [206] describe they the real universe is a sound attitude and should be a matter of temperate and balanced evaluation, and in particular, the question whether the FLRW line element is a good representation of the geometry of the universe should be asked and embarked to answer quantitatively. Only recently, as new observational data is coming to the fore, the importance of this issue has begun being realised. New more precise data are expected to make the issue of fitting more conspicuous and urgent to address in any cosmological considerations. Current observational information does not exclusively mandate the standard theory.

Interestingly, many cosmologists and physicists appreciated that the FLRW models are an oversimplification of Nature. Below, we quote several of them, but a great merit, it should be said, of the standard model is the absence of any *ad hoc* modification of prevalent theoretical ideas [207].

> ...the grounds on which homogeneity is generally assumed appearing to be those of convenience rather than generality. (...) We must categorically dissent from the extreme idea (...) that homogeneity is included in the definition of the universe (...) We hold that the assumption of spatial homogeneity is (...) a working hypothesis, valid so long as it does not conflict with observation or with theoretical probability, and justifiable during that time as a restriction on arbitrary speculation. (Dingle, 1933)

> The foregoing results demonstrate the lack of existence of any general kind of gravitational action which would necessarily lead to the disappearance of inhomogeneities in cosmological models. (...) it is at least evident from the results obtained, that we must proceed with caution in applying to the actual universe any *wide* extrapolations - either spatial or temporal - of results deduced from strictly homogeneous models. (Tolman, 1934)

> It is often claimed that the universe in the large must be isotropic or homogeneous. Certainly this view has immense aesthetic and philosophical appeal, but is it strongly supported by current observations? Unfortunately, it is not (...) observations neither confirm nor deny the "cosmological principle" that the universe is isotropic and homogeneous, or even homogeneous, and (...) measurements at the present time cannot prove, but can only disprove, that particular models represent the actual structure of the universe. (...) global theoretical models that are inhomogeneous should be looked for. (Kristian and Sachs, 1966)

> But this approximation[7] is a very crude one. The Einstein field equations are not linear, so that the disturbances in the field produced by the various stars cannot be just added, or averaged in any way. (Dirac, 1981)

[7] "This approximation" means replacing the real inhomogeneous distribution of matter in the universe by a smoothed-out one.

> Yet, it[8] seems based on absurdly simple assumptions. The universe is assumed to be spatially homogeneous and isotropic while at first sight it is remarkably non-homogeneous and anisotropic. True, one talks of a "large scale" in this connection but that large scale remains beautifully vague and undefined. (Raychaudhuri, 1988)

A more detailed discussion of the question of modeling of the universe will be given in the following sections – in fact, the large part of the burden of this thesis is devoted to it.

The question of choice of an appropriate cosmological model is an important one, but so is the problem of the global understanding of Einstein's equations which can be gained by studying the space of their solutions, the so-called *ensemble* of universes[9]. The field equations are defined on and at the same time determine the space-time manifold. This is what the construction of cosmological models refers to, and this is the point of view of observational cosmology. One simply assumes that each solution to the Einstein field equations describes a universe, and cosmology should but specify boundary or initial conditions relevant to the universe resulting from the astronomical observations.

The observations are never precise, consequently, a cosmological model can describe the real universe only within some limits of accuracy. It is usually assumed that there exists a one-to-one correspondence between the observations of a real system and a mathematical model. Assuming Einstein's field equations to be an appropriate mathematical model one is allowed to predict the evolution of the system. We are already aware that this picture is too idealistic. Even more, the model is usually assumed to have the property of *structural stability* [245], in the sense that the single exact solutions upon which the whole modeling rests are assumed to be in some sense representative, implying that our inherent inability to specify the model precisely would not have significant effects on the qualitative dynamical outcome. But as shown in [245] such a framework cannot be assumed *a priori*[10]

[8] I.e. the standard model.

[9] E.g., the Cauchy problem in a cosmological context naturally leads to the ensemble of possible universes: the set of all admissible kinds of initial data (on a space-like hypersurface) is equivalent to the set of universes evolving from it.

[10] This argument is essentially based on a number of theorems in dynamical systems theory which show

and in fact it might be that the appropriate theoretical framework turns out to be that of structural *fragility.*

Given these caveats the progress in cosmology can be considered amazing.

This thesis is organized as follows:

In chapter one, apart from rewieving the standard cosmological model, we introduce the problem of fitting and of averaging. A discussion of them follows in physical and cosmological contexts.

In chapter two, a classification and discussion of (some of) the approaches attempting the averaging of the small-scale (microscopic) Einstein equations to obtain the large scale cosmological field equations is given.

Chapter three is concerned with the Renormalization Group (RG) approach in gravity. The main ideas of RG theory and its application for studying critical phenomena in statistical mechanical models are discussed. Recently discovered numerically critical phenomena in GR gravitational collapse are reviewed in section 3.4.1. In section 3.4.2 some, rather radical, conclusions are offered pertaining to critical phenomena in gravitational collapse and the significance of the critical exponent close to 0.37. For example, in [60] a point is made of the fact that $0.37 \approx \frac{1}{e}$, while in our opinion it has no significance. Section 3.6 contains original material (work done in collaboration with M. Carfora), dealing explicitly with a smoothing procedure for inhomogeneous cosmologies based on RG approach [57]. This approach employs the arguments and methods paradigmatic of the real-space RG theory. The system's critical behaviour is studied and the significance of *criticality* in the universe is discussed. This point of view may have far-reaching impact on the study of structure formation and clustering in the universe, though we are well aware that a lot is to be done along this line of approach.

Chapter four contains original material (work done jointly with M. Bruni), a description of the evolution of perturbations in form of phase diagrams in the dust-radiation FLRW universe from the viewpoint of dynamical systems theory [47]. (As such this contribution is not directly related to the rest of the thesis.)

that structurally stable systems are not dense in the space of all systems.

Finally, chapter five contains conclusions and some indications for future research. The bibliography at the end is arranged in the alphabetical order.

1.2 Meaning of cosmological models

1.2.1 Standard cosmological solution

In this section a standard material is overwied and references may not be given explicitly each time; for pertinent references see e.g. [192] or [82].

The assumptions of the standard model are the validity of GR and Cosmological Principle. The classical term "cosmological model" usually means a geometrical description of the space-time structure and the smoothed-out matter and radiation content of an expanding universe, upon adopting GR as a fundamental theory of gravity. In cosmology, in order to handle the Einstein equations one introduces simplifications, better or less founded, dealing necessarily with models which by virtue of their building tend to eschew detailed realism.

To make the above assumptions more explicit we will spell out hypotheses that have been introduced, namely, the *isotropy hypothesis* - motivated by the fact that from our observation point space appears to be nearly isotropic with no indication of noticeable anisotropy. What is assumed is that on a sufficiently large averaging scale ($\sim 500\,Mpc$, say) there exists a mean motion of matter and radiation in the universe with respect to which all (averaged) observable properties are isotropic (i.e. independent of direction). Indeed, the counts of galaxies and radio-sources give near the same results for all parts of the sky; the cosmological expansion seems to follow the same pattern in various directions and the cosmic microwave background (CMB) radiation is very nearly isotropic. In this respect, (taken on its face value) there exist direct observational evidence favouring the isotropy of the universe, namely, the COBE data [233].

The *second hypothesis* is to assume the *homogeneity*, namely, all fundamental observers (following the mean motion) endowed with clocks experience the same history of the universe and the same observable properties, i.e. the universe is observer-homogeneous.

With these premises along, basically under the assumption about the symmetry of the space-like hypersurfaces, it is possible to arrive at a metric element known as the FLRW metric. In other words, a system of coordinates can always be found in which the line element can be written in one of its standard FLRW form

$$\begin{aligned} ds^2 &= -dt^2 + R^2(t)\frac{dr^2 + r^2 d\Omega^2}{(1 + kr^2/4)^2}, \quad \text{or} \\ ds^2 &= -dt^2 + R^2(t)\frac{dx^2 + dy^2 + dz^2}{(1 + \frac{1}{4}kr^2)^2}, \end{aligned} \tag{1.4}$$

where, $r^2 = x^2 + y^2 + z^2$; or

$$\begin{aligned} ds^2 &= -dt^2 + R^2(t)[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2] \\ ds^2 &= -dt^2 + R^2(t)[\frac{dr^2}{1 - kr^2} + \frac{r^2(dx^2 + dy^2)}{1 + \frac{1}{4}(x^2 + y^2)}], \end{aligned} \tag{1.5}$$

where the function $R(t)$ is unknown before solving the Einstein equations; k is an arbitrary constant and where r is a radial coordinate measured from the Earth and $d\Omega^2$ is the angular element corresponding to the angular coordinates θ and ϕ (i.e. $d\Omega^2 = d\theta^2 + sin^2\theta d\phi^2$); t denotes a cosmic time which coincides for each fundamental observer $\mathbf{r} = const.$ with the proper time, for neighboring observers the same values of t correspond to events simultaneous in the Einstein sense.

FLRW models have exact spherical symmetry about every point, i.e. space-time is spatially homogeneous and admits a 6-parameter group of isometries, whose orbits are space-like 3-surfaces of constant curvature (Minkowski and De Sitter space are examples of FLRW space-times with additional higher symmetry). The groups are direct products $B_3 \otimes O(3)$, where B_3 is one of the Bianchi groups [118].

Upon assuming the metric in the form of (1.5) it defines automatically a perfect fluid energy momentum tensor. $R(t)$ can be constrained only by an equation of state.

It is further taken that the universe is filled with a *perfect fluid.* It is justified assuming that one considers galaxies as the molecules of a gas that fills the space. Likewise, at the epoch when galaxies would not have existed and the universe would have been filled with a photon gas, it would behaved like a perfect fluid, as well. This particular form of the energy-momentum tensor is given by (1.2).

The necessary and sufficient conditions for a space-time to be FLRW are the following:

(1) the metric obeys Einstein equations with a perfect fluid source;

(2) the source velocity field has zero vorticity, shear and acceleration.

The necessity of these conditions can be ascertained by a direct computation. The sufficiency follows from the evolution equations for vorticity $\omega_{\alpha\beta}$, shear $\sigma_{\alpha\beta}$ and acceleration $\dot{u}^\alpha$ [82], which imply that a perfect fluid solution with $\sigma = \omega = 0 = \dot{u}^\alpha$ must be conformally flat. All such solutions were found by Stephani, in general they have $\dot{u}^\alpha \neq 0$, by requiring $\dot{u}^\alpha = 0$ we get FLRW models (e.g. [160]).

Another invariant definition of the FLRW space-times which makes no use of the field equations is the following:

(1′) the space-time admits a foliation into spacelike hypersurfaces of constant curvature;

(2′) the congruence orthogonal to the leaves of foliation are shear-free geodesics;

(3′) the expansion scalar of the geodesic congruence has its gradient tangent to the geodesics.

Another often met representation of FLRW metric is the following

$$ds^2 = -dt^2 + R^2(t)(dr^2 + f^2(r)d\Omega^2), \tag{1.6}$$

where

$$f(r) = \begin{cases} \sin r & \text{for } k > 0 \\ r & \text{for } k = 0 \\ \sinh r & \text{for } k < 0 \end{cases}$$

All three cases are covered by

$$ds^2 = -dt^2 + R^2(t)(dr^2 + k^{-1}\sin^2(k^{1/2}r)d\Omega^2). \tag{1.7}$$

The range of r is finite or infinite depending on the sign of k and coordinates used.

A few points are worth recalling, namely:

1. FLRW metric (or element) is a very particular solution of the Einstein equations.

2. For the time variable t fixed ($dt = 0$), the ds^2 is that of a 3-space with constant riemannian curvature at every point of the space and therefore the space is spherical, euclidean or hyperbolic, for $k = 1, 0$ or -1, respectively. When $k = 0$, the group B_3 can be of Bianchi type I or VII_o; for $k < 0$ (open FLRW models) its B_3 can be of Bianchi type V or VII_h and for $k > 0$ (closed FLRW models) its B_3 is of type IX.

3. The FLRW metric leads to homogeneous model universes, through the assumption of isotropy with respect to the Earth, which leads to a space of constant curvature and, consequently, to a space that is isotropic with respect to all other points. One should not forget, however, that the near-isotropy with respect to the Earth is valid within certain limits: redshifts up to 0.5 for galaxies, 5 for radio-sources, 7 for the CMB, and in any case all observations of this kind are limited by the cosmological horizon.

4. We are dealing with uniform model universes, that is to say for t fixed, the density μ and pressure p each have the same values at every point in the space.

5. The worldlines $r = const., \theta = const., \phi = const.$ are geodesics of the space-time. What this means is that the worldlines in question are possible solutions of the equations of GR and they thus represent possible motions, e.g. of field galaxies (after correcting for random motions) or galaxy clusters, which in this picture have fixed coordinates r, θ, ϕ, termed as comoving coordinates.

6. $R(t)$ appears as a scaling parameter for the universe, if it is an increasing function of t we are then led to an expansion of the universe. It gives the rate at which two points at fixed comoving coordinates increase their mutual physical distance as $R(t)$ increases.

The time t (cosmic time) is the proper time at every point fixed in the comoving coordinates, in particular, it would be shown by clocks in various galaxies synchronized by exchanging light signals.

The comoving coordinates r, θ, ϕ owe their properties to the fact that effectively the distribution of mass in the universe may be described by masses having fixed comoving coordinates[11].

To completely solve the cosmological problem, the two Einstein equations together with the equation of state for the perfect fluid $p = p(\mu)$, we have to determine the three functions: $R(t), \mu(t)$ and $p(t)$. If we consider only the case in which the cosmological constant $\Lambda = 0$, we have then the simplest family of relativistic models, the FLRW models. These models are necessarily an approximation to the real universe, due to various simplifying assumptions employed in their derivation.

[11] The comoving distance is not an observable, unlike e.g. the luminosity distance.

In the case of FLRW model the Einstein equations reduce to a system of two equations

$$\left(\frac{\dot{R}}{R}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\mu - \frac{k}{R^2}, \tag{1.8}$$

$$-\frac{\ddot{R}}{R} = \frac{4\pi G}{3}(\mu + 3p). \tag{1.9}$$

Heuristically, these equations can be seen as the equivalent of the energy conservation and the second dynamics law of classical (non-relativistic) mechanics. We see that two points at a distance $d \equiv R(t)r$ move apart with velocity $v = Hd$. The Hubble constant determinations at the present epoch give $H = 100\, h\, km\, s^{-1}\, Mpc^{-1}$ with $0.4 \leq h \leq 1$. The large uncertainty is mainly due to the discrepancies between different distance estimators.

A standard model - one of the homogeneous, isotropic FLRW model (and the present epoch (t_o)) can be determined by the measurement of two dimensionless numbers: the deceleration parameter $q_o \equiv -\ddot{R}_o/(R_o H^2)$, the density parameter $\Omega_o \equiv \mu_o/\mu_{crit}$, and a scale constant H_o ($\Lambda = 0$ is equivalent to $2q_o = \Omega_o$). The critical density $\mu_{crit} = 3H_o^2/(8\pi G) = 1.9 \times 10^{-29}\, h^{-2}\, g\, cm^{-3}$; present density values of μ, above, below or equal to μ_{crit} correspond to closed, open or flat geometries, respectively. Current limits on Ω_o are, $0.2 < \Omega_o \leq 1$. The limit of $\Omega_o = 1$ is usually preferred on the grounds of the standard inflationary universe scenario. Standard cosmology puts forward the equation $4\pi G\mu/3 = qH^2$. If and when the precise value of μ_o, q_o and H_o are available, the decision about the standard model will be easy to reach.

At present the cosmological observables include also: the age of the universe, composition of the universe, CMB radiation and other cosmological background radiation, the abundance of light elements, the baryon number of the universe (quantified as the baryon-to-photon ratio), and the distribution of galaxies and larger structures. Observations of the universe on scales similar to the size of a galaxy $\sim 10\, kpc$ display significant inhomogeneities, but the current interpretation is that below such scales non-gravitational forces are dynamically dominant. On the other hand, scales $\gg 10\, kpc$ are considered relevant to the large scale structure. In any case, non-negligible anisotropies were never observed at scales comparable to the horizon scale $d_H \sim (cH_o)^{-1} = 3000\, h^{-1}\, Mpc$. Therefore the cosmological framework of the hot Big Bang in a spatially homogeneous and isotropic universe

- the Standard Model - is taken to be experimentally vindicated, through the successful prediction of primeval element abundance and the observations of relic radiation in the form of CMB radiation [252][12].

However, from observational point of view the standard models have currently some problems. Some of them were listed in [23][13].

1. Dark matter not visible, but revealed through non-Keplerian motions of galaxies and stars in galaxies.

2. Streaming motions of galaxies with velocities $v \leq 5000\, km/s$ towards invisible attractors.

3. Sponge-like large scale structure and possibly spatial periodicity in the distribution of galaxies.

There seems to exist observational evidence in favour of larger and larger structures (e.g. [61] recently), nevertheless, [37] indicate that there is a tendency to homogeneity at large scales, although it is difficult to point out the scale at which homogeneity is reached, due to the small size of present redshift surveys. The striking feature of the (luminous) matter distribution is the existence of voids (in the scales probed so far, up to a few hundred Mpc) surrounded by sheet-like structures containing galaxies (e.g. [111]), with a typical size of voids $50 - 60\, h^{-1}\, Mpc$. Also, larger underdense regions of $\sim 130\, h^{-1}\, Mpc$ probably exist [42]. Interestingly, dynamical estimates of FLRW density parameter Ω_o give different results on different scales. Observations of galactic haloes on scales $10 - 30\, Mpc$ give $\Omega \simeq 0.2 \pm 0.1$ [220], whereas smoothing the observations over scales $\sim 100\, Mpc$ indicates the existence of a less clustered component with a contribution probably as high as 0.8 ± 0.2 [76]. Hopes have been raised that new inputs like the inflationary scenario and cosmic strings may solve the problem.

4. According to the present evidence quasars must contain large magnitude density concentrations (especially if they accrete black holes), but matter distribution at the last scattering should be nearly isotropic, as confirmed by the very high degree of isotropy of the CMB radiation. The question therefore is, how such density contrasts might have

[12] Doubts still persist, e.g. regarding the exact primordial abundance of helium.

[13] Several objections can be accommodated within either the Lemaître-Tolman [167, 248] or Szekeres [244] models with a non-simultaneous Big Bang.

evolved from such a homogeneous initial state in the time implied by the standard model.

Generally, the problem of structure formation in spatially homogeneous models is an important problem of principle. *"Statistical fluctuations in FRW models cannot collapse fast enough to form the observed galaxies. This suggests that there must be real inhomogeneities at all stages in the universe. Moreover, some perturbations of FRW models are decaying modes which would have been more important in the past."* (MacCallum, 1979). (Cf. point **3.**)

5. Astronomical observations contradicting the Hubble law [10].

The linear velocity distance relation was questioned in [198]. Their conclusion about the quadratic relation between velocity and distance did not gain an acceptance[14].

McCrea in 1939 [190] on the basis of Milne's result, that any cloud of particles with different velocities and initially being confined to a finite space volume will disperse and eventually obey the Hubble law in the first approximation, concluded that in order to infer matter distribution in the universe from the observed motion of galaxies, a higher approximation is needed than the Hubble law.

6. The quadrupole anisotropy in the CMB radiation on large angular scales $\simeq 80^o$ [99].

The quadrupole anisotropy could serve as a test of the anisotropic expansion of the universe at the last scattering epoch since it is due to the non-vanishing vorticity of the universe; its matter is rather subtle[15] and still disputable.

7. Anomalous redshifts and "quantization" of redshifts.

The so-called anomalous redshifts refer to objects which are presumably close together and show significantly different redshifts. There is also the curious quantization of redshifts, with Δz multiples of Δz_o corresponding to a velocity of $72\,km\,s^{-1}$ [11]. Both of these cast doubt on the idea of redshift being *only* due to universal expansion.

8. Still controversial interpretations on the observations of radio sources, like the angular size – flux density relations.

All the above facts cause problems with fitting, however not only the data but also

[14]It was shown later that even such a relation may be accommodated in the standard model [208].

[15]It might be seen as contradicting Mach's Principle; attempts are therefore advanced to attribute it rather to local causes, like the non-homogeneity in the mass distributions of galactic clusters.

the logic of the standard approach make the fitting problem particularly urgent (cf. next section).

1.2.2 Modeling of the real universe

A standard way to analyze a real system is to make a mathematical model, which can then be studied analytically or numerically. The relation between the real system and its model should be properly assessed, though quite often it remains obscure.

On local scales, like the scale of planetary systems, we can safely assume GR to hold, but there the space-time geometry determined by Einstein's equations should be very complex. In fact, a mathematical model of the matter distribution going down to these small scales is almost impossible to obtain, as there is no observational data on which it could be based. Therefore, when considering the kinematics and dynamics of the universe as a whole, one usually ignores the fine-graining due to the local inhomogeneities and deals with the simpler structure of space-time geometry which is more illuminating from the point of view of cosmology.

As we have already seen, the standard FLRW universe models are perfectly homogeneous and isotropic. Therefore one can question their applicability for modeling the universe accurately, as it is manifestly not a FLRW universe (on at least some scales). But, these standard models are usually taken to represent the real universe in some vague "average sense", or on some averaged scale. This implied averaging procedure should be of great importance in cosmology, especially in terms of interpreting the meaning of FLRW models. However, it hardly ever receives a due attention, even though it underlies the geometric and physical applications of FLRW metrics to describing the real universe, and even though it has been posed by a number of authors [229, 227, 83]. Shirokov and Fisher [229] seem to have been the first to consider it in 1962. Recently, this problem has been brought to general attention by G.Ellis.

Problems arise when we start to relate the realistic inhomogeneous universe models to these idealised, smoothed-out models. The very relation between them is not clear, in particular, it is not quite obvious how the galaxies or clusters of galaxies are related to the comoving coordinates of the averaged idealised models, nor how particular light rays correspond to the idealised geodesics of these models, etc. The standard approach in this respect is a theory of the perturbed FLRW models (see e.g. [18]) and their relation to observations of galaxies and background radiation [150, 218] which is still a matter of investigation (see e.g. [239, 46, 215]).

In observational cosmology it is standard to follow this scheme: firstly - to observe the distribution, masses and velocities relative to us of neighbouring galaxies; secondly - to calculate the averaged quantities assuming the Hubble law, i.e., the isotropy (on average) of the relative velocity field, and homogeneity of the distribution of galaxies; thirdly - to compare these mean properties with those of FLRW models, having the same density as that of the total mass of the galaxies uniformly distributed in the observed region. This procedure basically means that the FLRW models are assumed instead of deduced from observations. But in principle, what one could hope to get in this way, are the best-fit parameters of the FLRW models. Unfortunately, the discrepancy between the observational data and the properties of FLRW models is such that it is usually necessary to introduce an additional matter contents besides that estimated from the visible matter (dark matter), or even a cosmological constant (see e.g. [151] and [225]). However, in effect, even without the benefit of a sound theory as to how to obtain the best-fit of idealised model universes to reality, the observers are already "fitting" the observed velocity patterns to a hypothetical FLRW velocity field (e.g. in the efforts to determine the velocity of the local supercluster relative to a FLRW background).

An attempt to translate the mathematical prescriptions into practical observational procedure, unfortunately seems to be a very difficult issue with various problems, concerning in particular, the determination of the average density and velocity of matter at a given distance down the past null cone [81, 83]. Other alternative analyses of homogeneity, based on "almost Killing vectors" [189] or "observational homogeneity" [35], do not yet seem able to resolve these issues as well.

But a reliable description of the inhomogeneities in the expanding universe is wanting above all. Clumpiness obviously affects the analysis of observations in an inhomogeneous universe model, by affecting the dynamics of photons in the universe altering the shear and the convergence of null rays. This has its effect on the cosmological area-distance relation and in consequence on apparent sizes and luminosities (see e.g. [26]).

It is possible to take this effect into account, but the form of the focusing depends sensitively on how strongly the matter is clumped. A useful idealization is the Dyer-Roeder formula (see e.g. [226]) valid for universes "not too much filled with clumps", it represents the largest possible angular diameters distance (for a given redshift) for light bundles which have not yet passed through a caustic.

The effect of clumpiness on geodesics paths can change the relations adopted in standard model, but that strongly depends on the clustering of matter in the universe. The same applies to the issue of Ω_o. The underlying assumption is that there is a good limit to Ω_o if we take large enough averaging volumes, but it is not so if we live in a hierarchical universe [251], where there is no non-zero limit to Ω_o as we use larger and larger averaging volumes[16].

The transition from an inhomogeneous model to an averaged (or smoothed) standard model is of fundamental importance, also in the context of structure formation and the interpretation of distance measurements.

In practical terms the heuristical justification for using FLRW models asserts that for them to hold the matter inhomogeneities have to be averaged (or smoothed-out) and redistributed homogeneously (e.g. in the form of a perfect fluid). A series of related mathematical problems arise. First of all, let us notice that we are using continuous functions in modeling the universe (matter density, pressure or kinematical scalars of the velocity field), assuming that they represent "volume averages" of the corresponding fine-scale quantities. The results of such averaging in an inhomogeneous medium depend on the scale, but this scale (of averaging) was never explicitly agreed upon. Additional problem is that a volume average for tensors is a non-covariant quantity (unlike for scalars), so a more sophisticated definition is required. The basic and tacit assumption which underlies this whole procedure is that a smoothed-out universe and the actual, inhomogeneous one,

[16]Suggestions that the universe might have a fractal structure [36] might support this argument.

behave identically under their own gravitation. Or, to be more precise, almost identically on some scales of interest, e.g. the ones that are much greater than a characteristic scale of the local inhomogeneities and much smaller than a characteristic length of the universe model under study. This assumption is usually taken for granted, but does by no means need to be true. Indeed, the non-commutativity of averaging of the metric and calculating the Einstein tensor (highly non-linear in the metric) is a severe problem (this will be discussed in more detail in section 1.2.2.2).

1.2.2.1 The "fitting" problem in cosmology

The basic idea here is that we do not *a priori* assume that the universe is necessarily well described at all times by the FLRW model, but nevertheless decide to use such a model for, say, pragmatic reasons [92].

The problem is then how to determine a *best-fit* between a clumpy cosmological model $\mathcal{U}$, which is supposed to give a realistic representation of the universe including all inhomogeneities down to some specified length scale, and a smoothed-out, idealised FLRW model $\mathcal{U}'$. The focus in this approach is on the relation between the idealised model and more realistic descriptions of reality. Therefore, one should also be able to specify details of that fitting, including the issue of how good the fit is. Basically, one could aim at the repeated use of the smoothing procedure, i.e., to consider a best fit between any lumpy universe model and a model $\mathcal{U}''$ which gives an even better description of the real universe than $\mathcal{U}'$, by describing the inhomogeneities at an even more detailed level.

In principle, this process would allow one to determine the best description at any prescribed level of detail.

The above task can be approached in many ways, based on:

(i) the space of space-times,

(ii) initial data for space-times,

(iii) the "gauges" adopted in perturbation studies,

(iv) near equivalence,

(v) average behaviour,

(vi) null data,

(vii) normal coordinates.

The (i) approach gives a useful concept of the fitting idea, but it does not take into account the dynamics of GR. In this approach, the lumpy model $\mathcal{U}$ and the ideal one (FLRW) $\mathcal{U}'$ are represented by points P, P' in the space of space-times S, and the point P' is constrained to lie in the hypersurface of FLRW space-times. If a suitable distance function on S is known then, given $\mathcal{U}, \mathcal{U}'$ is chosen in such a way as to minimize the distance between P and P'. Unfortunately, no natural positive definite metric on S exists. Also to be able to distinguish whether two different points in S represent the same space-time or not, one would need to factor out the coordinate freedom (the diffeomorphism group) which seems also to be problematic, unless one could find the best fit in some specific and operationally defined coordinate system rather than the general one. Moreover, matter distributions would need to be fitted as well. Most likely this approach cannot be easily related to an observational procedure on the past null cone.

Instead, in the (ii) approach we consider the space S^* of initial data for space-times (the phase space of a cosmological model in GR), with initial data given on a spacelike hypersurface Σ : $(g_{ab}, K^{ab}, \mu, q^a)$, where, g_{ab} is the 3-metric on Σ, K^{ab} its second fundamental form, μ the matter energy density, q^a a 3-momentum relative to Σ. Given this, each point Q in S^* will correspond to a specification of all of these quantities at each point on the initial hypersurface Σ, chosen so as to satisfy the Einstein constraints on Σ.

The problem then is stated similarly as in the (i) approach. Thus the difficulties of (i) remain. One has to take into account, here as well, the fact that different data can represent the same cosmology $\mathcal{U}$ (e.g., choose two different spacelike slicing Σ of $\mathcal{U}$ to get two different 3-space metrics g_{ab}). However, involved here is rather the hamiltonian structure of GR, and roughly speaking, the symplectic form of that structure gives a way of comparing metrics on two different spacelike hypersurfaces, to see if they represent slicings of the same space-time. Therefore in principle, once the fit of 3-metrics is determined, the 4-dimensional fit would be as well, even if not explicitly known [113].

Clearly, the (ii) approach originates from the *ensemble* viewpoint. To understand the equivalence of models and the "distance" between them calls for the introduction of the appropriate topology and metric structure on the *ensemble* which is a difficult and, up to now, an open problem.

Usually, one starts with $Lor(\mathcal{M})$, the space of all Lorentz metrics on the manifold $\mathcal{M}$. This space admits infinitely many topologies, none of which is "natural". However, what one is really interested in is a subset of $Lor(\mathcal{M})$, namely, those metrics that are the solutions of Einstein's equations. Fortunately, this set usually is a smooth manifold, with a local representation in the space of four functions of three variables, and Einstein's equations acting as a hamiltonian system. Further, the space of linearized solutions of the Einstein equations is tangent to it (this is so only, if our set of metrics is a smooth manifold in the neighborhood of a given solution) and in this region Einstein's equations are stable with respect to the linearization[17]. This approach can be helpful for the global analysis of the solutions to Einstein's equations.

In fact the averaging procedure advocated in section 3.6 takes on just this viewpoint.

The point of view of approach (iii) is quite useful since when describing perturbations of FLRW universe we need to choose a gauge, which is essentially a question of fitting a FLRW universe $\mathcal{U}'$ to a lumpy one $\mathcal{U}$ [18, 237]. The choice of a gauge in a perturbed universe model $\mathcal{U}$ is equivalent to choosing a point identification between the FLRW model $\mathcal{U}'$ and $\mathcal{U}$. Given a choice of local coordinates in $\mathcal{U}'$ and in $\mathcal{U}$, the above correspondence can be expressed in terms of a relation between these coordinates. But instead this is usually taken to be the identity, i.e. the coordinates of the corresponding points are taken to be the same.

A specific gauge can be characterized in terms of a choice of "hypersurfaces of simultaneity in the physical space-time" [18], which is actually the choice of a mapping of the FLRW surfaces ($t = const.$) in $\mathcal{U}'$ into the lumpy model $\mathcal{U}$. Once a point correspondence between space-times is set up one can then look for the parameters H_o and q_o giving the best fit. Seen this way the problem of fitting is the problem of choice of gauge in disguise.

The classic equivalence problem of GR (see [77] for a review) is much of relevance for the fitting problem - (iv) approach. However, a direct approach by comparison of curvature

[17] In general, this is not true for space-times with Killing vectors.

invariants of two space-times is problematic because the metric tensors are indefinite. But to prove an equivalence of two space-times it is sufficient to have the equivalence of the curvature tensors and their covariant derivatives, evaluated in orthonormal frames [148]. Using a procedure of this kind would allow to determine if two cosmologies are almost equivalent. There is also here the question of how to choose which pairs of points p, p' to try to identify in the two cosmologies. Secondly, the idea of almost equivalence is more complex in the context of specific choices of tetrad canonically related to particular Petrov types. An extended algebraic classification of the curvature tensor and its derivatives should make possible determination of near-equivalence of space-times by extending the methods used to examine exact equivalence. Moreover, the problem would be simplified in the cosmological context. Thus this is a promising approach though its relation to possible observational procedure seems obscure.

The (v) approach, based on the concept that the smoothed-out model $\mathcal{U}'$ should accurately represent the average behaviour of the more realistic model $\mathcal{U}$, is going to be discussed at length in the next section. The important issue to address here is what is precisely meant by the "average" model which is also what most of this thesis is devoted to. This line of approach to fitting is also made use of in section 3.6.

The null data approach (vi) [92] is a specific observational prescription to best-fit null data to obtain optimal, i.e. best fit parameters describing FLRW universe, given an optimal correspondence between any FLRW model and lumpy universe. This approach is in practice similar to what is done at present by observers (as it extends the approach of Kristian and Sachs [163]), but it can further be related to the averaging approach and suggests moreover the nature of possible criteria of acceptable fit.

Finally the (vii) approach, closely related to the nature of local physics in a lumpy universe, is in fact a local almost equivalence approach. This is so since in the analytic case, one can examine the curvature near any space-time point by using normal coordinates about that point. Different space-times can be locally compared by writing each of them in such coordinate systems centered on points p, p', and directly comparing the metric components up to some required order of accuracy. As before, the choice of correspondence of points p, p' to make in the two space-times is a problem. This approach can nevertheless be related to the question of a best fit to astronomical observations by an appropriate adaptation of [163], whereas carrying out the analysis in a non-local way ends up in (vi) approach [92].

The implications of fitting can be analysed in terms of Traschen integral constraints [249]. In an almost homogeneous universe model with inhomogeneities due to local physical processes, the local energy and momentum conservation imply the existence of conserved quantities expressing the conservation of monopole and dipole terms. In such models, the Sachs–Wolfe effect is reduced with respect to the ones where the effect is ignored [249][18].

According to [250] this argument would not apply to matter perturbations created by quantum fluctuations in the inflationary epoch because of their non-local size, in terms of today's scales. Nevertheless, the constraints can be thought of as the fitting conditions required to be satisfied, if the chosen FLRW background model has the right monopole and dipole terms corresponding correctly to the real universe [89].

For example, the very definition of Ω_o refers always to an idealised background model which cannot be determined without simultaneously solving for its perturbations.

We can start with a uniform universe model A and model high density regions in it by adding some over-densities here and there, resulting in a non-uniform model B. Obviously, the average density Ω_B in B is greater than the background density Ω_A of A, thus using model B means also changing the background model to, now A', which has the density $\Omega_{A'} = \Omega_B$ and different dynamics from the initial model considered. If we use the same background density value (Ω_A) in the lumpy universe model, model B has to be replaced by model C, say, in which the high density regions are surrounded by void regions, in order that the average density $\Omega_C = \Omega_A$. This is basically Traschen's condition [89].

1.2.2.2 The "averaging" problem in cosmology

One important way of thinking of the use of a smoothed-out model is that it represents the average properties of an inhomogeneous model. If $\mathcal{U}^*$ is the smoothed-out model universe,

[18] Other dynamic relations are affected as well.

obtained form a clumpy one $\mathcal{U}$ by an averaging procedure, then it represents the nature of $\mathcal{U}$ when described over some averaging length scale L. The best-fit FLRW model $\mathcal{U}'$ should be the same as the averaged model $\mathcal{U}^*$, if one can indeed describe the large scale nature of $\mathcal{U}$ by a FLRW space-time [92][19].

In standard cosmology to describe the discrete matter distribution in the universe we use a continuously distributed stress tensor, most often the so-called perfect fluid form of it. But strictly speaking using the Einstein equations in this situation is not well-founded. What it means is that an effective averaging of real inhomogeneities has by this been carried out while, at the same time, the unchanged left hand side of Einstein's equations is tacitly assumed to describe the "averaged" gravitational field. However, we should bear in mind that the Einstein equations are highly non-linear, which is why any averaging process on them is far from trivial in general[20]. In other words, the averaging process may change their structure and consequently the geometric and physical meaning of the very gravitational field would be changed.

In particular, Ellis conjectured [83] that, upon smoothing-out the space-time geometry there would appear geometric correction terms in the sources to Einstein equations. They may have influence on the dynamics of the universe. In general, there would always be a non-zero backreaction of inhomogeneities on the dynamic behaviour at the smoothed-out scales affecting the expansion rate and the estimate of age for the universe.

To put it differently, if we calculate the Einstein tensor $\tilde{E}_{\mu\nu}$ from an "averaged" (whatever this means) metric $\bar{g}_{\mu\nu}$, it will not be equal to the Einstein tensor $\bar{E}_{\mu\nu}$ which was first calculated from the fine-scale metric $g_{\mu\nu}$ and then averaged. As a consequence, the Einstein equations seem not to hold on scales where averaging is required if they hold on, say, planetary scale. Most probably, the Einstein tensor $\tilde{E}_{\mu\nu}$ determined by the smoothed-out metric $\bar{g}_{\mu\nu}$ will be related to $E_{\mu\nu}$ by a map S'' distinct from the smoothing operator S' acting on the matter tensor $T_{\mu\nu}$. However in cosmology the following is assumed: one

[19] An intriguing possibility one can think of is that one could construct clumpy Small Universe [84] models which look like the perturbed FLRW models, but for which the smoothed-out (large scale) version is not a FLRW universe. They would appear approximately homogeneous, but their topology would be incompatible with the symmetry of exactly homogeneous universes.

[20] Unlike in electrodynamics, where the macroscopic Maxwell equations can be derived by averaging out the microscopic Maxwell-Lorentz equations over 4-regions in Minkowski space-time (see e.g. [120]).

calculates the Einstein tensor from a metric that is supposed to be already averaged and equates it to an energy momentum tensor – already averaged as well.

Let us now introduce a tensor $\Pi_{\mu\nu}$ to represent the difference between the Einstein tensor $\tilde{E}_{\mu\nu}$, obtained from the smoothed-out metric $\bar{g}_{\mu\nu}$, and the smoothed-out matter tensor $\bar{T}_{\mu\nu}$ [83]. The point is that the Einstein equations should in fact be "corrected", so that the difference $\Pi_{\mu\nu} \equiv \tilde{E}_{\mu\nu} - \bar{E}_{\mu\nu}$ is compensated, namely,

$$\tilde{E}_{\mu\nu} - \Pi_{\mu\nu} = \kappa \bar{T}_{\mu\nu}. \tag{1.10}$$

Writing the term $\Pi_{\mu\nu}$ on the right hand side of (1.10), we can interpret it as a correction to the source resulting from averaging out the small scale inhomogeneities of the gravitational field. This brings back the equations again to their familiar form, but with the effective source term,

$$\tilde{E}_{\mu\nu} = \kappa \bar{T}_{\mu\nu} + \Pi_{\mu\nu}. \tag{1.11}$$

This correction term is to be added to the field equations at scales other than the scale at which they are verified and which is therefore the scale on which they are supposed to hold. If we assume that the averaged metric is the FLRW metric and that $\bar{T}_{\mu\nu}$ has the perfect fluid form (homogeneous and isotropic), the correction $\Pi_{\mu\nu}$ will perturb the energy density and pressure of the source, invalidating the FLRW relation between the sign of spatial curvature, on the one hand, and the size and lifetime of the universe, on the other[21].

A correct and consistent treatment of this problem would require one to average the geometry *and* matter present, i.e. a microscopic matter distribution and the Einstein field operator, to determine both sides of the averaged Einstein field equations. Considering the above kind of averaging, one has to determine a relation between the manifold structures and corresponding points in the two models that we deal with at each step of averaging. Finally, this would allow us to explicitly determine the averaged field equations responsible for the large-scale, average dynamics and study the observational properties of the average universes, and the relation between more detailed and the averaged behaviour in them.

[21] It holds anyway only for dust without a cosmological constant.

In the weak-field or almost FLRW cases one can use direct methods to define the averaging (see chapter two for a review of some approaches); in the full theory the situation is more involved.

It is probably worth stressing at this point that speaking of averaging of the Einstein equations requires an utmost care and one has to clearly state what this means, since by themselves the Einstein equations (as any differential equations) *do not contain any in-built fundamental length*, so that they can be used *a priori* to describe cosmoses of any size. It should be emphasized that while doing this, these different metric and matter tensors used are intended to describe *the same* physical system, i.e. the same space-time, but at different scales of description enabling one to resolve less or more details[22].

[22]Posing the problem this way turns out to be very useful from the Renormalization Group approach viewpoint which we are going to propose in chapter three.

2 A survey of approaches for averaging

In this chapter we shall review some of the approaches for constructing the relativistic cosmological models via averaging out inhomogeneities of geometry and matter. A helpful reference on the subject is [161] listing many of the relevant papers.

We do not adopt a unique notational convention in this chapter. This was mainly dictated by the fact that there are various conventions adopted by various authors and for the reader's convenience (who might wish to consult the original papers) we decided to maintain the notation of each reviewed paper unchanged.

2.1 Averaging within approximation schemes

The definition of averaging can be based on or coupled with approximation schemes.

In the first attempt of this kind [243] (not meant to be applied to cosmology) Szekeres showed that linearized Einstein equations, i.e. when the metric is assumed to be a small correction superposed on the Minkowski background, are formally similar to the Maxwell equations. As a consequence of this, a macroscopic theory for gravitation could be formulated in analogy with Lorentz's theory in electrodynamics. In particular, estimates of the terms Π in equations (1.11) have been given

$$\Pi_{ab} = {Q'}_{ab}{}^{cd}{}_{,cd}, \tag{2.1}$$

with Π_{ab} determined as the double divergence of an effective quadrupole gravitational polarization tensor Q'_{abcd} with suitable symmetries $Q'_{abcd} = Q'_{[ab]cd} = Q'_{ab[cd]} = Q'_{cdab}$

(Q_{abcd} incorporates also any dipole polarization effects that may occur). Such models might cover the larger-scale transitions.

This approach does not invoke any specific kind of averaging. It was applied to the propagation of gravitational waves through a medium whose molecules were supposed to be harmonic oscillators. The result obtained was that gravitational waves slow down in such kind of a medium.

Another approach is the one due to Sibgatullin [231]. He does not give a definition of averaging either. The Einstein equations are averaged after the metric was decomposed into a "background" and a "fluctuation". No criterion to separate the metric into the background and fluctuation was provided. With the assumption that the characteristic scale of correlations between matter and geometry is small, with respect to the scale of variation of the smoothed-out geometry, the result was calculated approximately that fluctuations in matter do not influence the equations of zero-th and first order in the small parameter.

Independent considerations on the subject was presented by Bialko [28]. In this approach the metric was developed into a FLRW background and a high frequency perturbation. The Einstein equations were averaged over spatial volumes, under the assumption that the characteristic wavelength of the perturbation is small as compared to the curvature radius. Further, the equations governing the evolution of the averaged perturbation were found to differ from those for linearized perturbations by a logarithmically varying factor.

Another approach is due to Noonan [199]. He showed within the weak field slow motion approximation that when the Einstein equations are averaged by volume, the energy momentum tensor splits into three parts. The interpretation of the first part is kinematical due to averaged microscopic motions. The second contribution is mechanical due to averaged microscopic stresses, and the third one - gravitational, due to averaged small scale variations in the gravitational field. In [200] the author showed in addition, that the time-space components of the above macroscopic energy momentum tensor can be interpreted as the flux of gravitational energy of the microscopic field.

In the recent approach developed by Futamase [104, 105], the averaging is performed in the perturbation framework (see section 2.1.2 for details). The components of tensors are averaged over the spatial volume. In [107] the author considered inhomogeneous space-times with preferred slicings. Assuming that in the limit of zero perturbation, the preferred slicings go over into the homogeneous spaces of the FLRW models, the effect on the Friedmann equation of averaging by 3−dimensional volumes within the preferred slices was calculated. In [30] in the same approximation, the backreaction of inhomogeneities on the evolution was calculated, the result being that inhomogeneities slow down the expansion as compared to the standard Friedmann equation. Therefore the age of the Universe calculated from the Hubble law should be underestimated.

Another option mentioned in [106] could be statistical kind of averaging. Suppose that we have a statistical ensemble containing all possible density and velocity distribution of fluid elements (representing galaxies) with some constraints which characterize the universe we wish to approximate. Choosing a particular ensemble in which the density and velocity distribution satisfy the condition $<\delta\mu>=<v^i>=0$, and the averaging of any quantity with spatial derivative vanishes, then the averaging procedure obtained could be also appropriate to treat the situation where there are singularities, unlike within the spatial averaging concept [106].

2.1.1 Perturbation approaches

There is the "usual" perturbation approach, by which we mean here that one first introduces a fixed, i.e., unaffected by the perturbations background space-time, e.g. the FLRW metric, and assumes that the perturbation variables in the given background are small. With this assumption, one can expand these variables to higher orders, keeping only the zero and first order terms. This allows to tackle only weakly non-linear situations. The important fact is ignored in this approach, namely, that the material distribution itself determines the geometry, and in the presence of inhomogeneities one cannot really specify the background metric independently from the inhomogeneities - the backreaction problem.

The study of perturbations of the Einstein equations in the cosmological context started with the pioneering work of Lifshitz [169]. Of particular interest are the scalar perturbations since they are directly related to density fluctuations in the early universe and are thus relevant to the structure formation. Lifshitz's theory is however not easy to interpret due to its gauge ambiguity. This ambiguity is eliminated in the theory of gauge invariant perturbations due to Bardeen [18]. Both approaches are in reality valid only in the linear regime.

In principle, there exists a general method of determining the equations for any order of perturbations, but in practice the generalization of these schemes to the non-linear situations is not straightforward.

Recently, a new gauge invariant version of perturbation theory has been given [86, 88] (henceforth we call it EBH scheme) and we review it below (see also [44]). The standard $\delta\mu/\mu$ approach compares two evolutions (the actual one and a fictitious comparison one) along a world line. The covariant and gauge-invariant EBH scheme compares evolutions along neighbouring worldlines in the actual universe reflecting thus the spatial density variation in the fluid.

The advantages of this scheme are, firstly, that it does not necessarily assume the background geometry *a priori* as exact equations are given governing the evolution of density inhomogeneities in arbitrary space-time, without any reference to a background space-time. Secondly, it deals with exact quantities, like e.g. the comoving fractional gradient of the energy density orthogonal to the fluid flow (spatial projection of the energy density gradient). These are both directly observable and gauge invariant in the case of linear perturbations about FLRW universe. It has been shown that the linearized EBH equations are equivalent to the Bardeen gauge invariant equations [138]. The EBH equations are not however restricted only to the linear case[1].

We consider the exact covariant fluid equations for a general fluid flow in a curved space-time [82]. The 4-velocity vector tangent to the flow lines (the world-lines of fundamental observers in the universe which are at rest with respect to our volume element of fluid) is $u^a = dx^a/d\tau$, $(u^a u_a = -1)$, where τ is the proper time along the fluid flow lines.

[1]An extension of EBH scheme, combined with the spatial averaging, to tackle non-linear case can be also developed (see [108] for a sketch of the scheme).

The projection tensor into the tangent 3-spaces orthogonal to u^a (into the local rest frame of a comoving observer) is

$$h_{ab} \equiv g_{ab} + u_a u_b, \tag{2.2}$$

and $h_b{}^a h_c{}^b = h_c{}^a, h_a{}^b u_b = 0$.

The time derivative of any tensor $T^{ab}{}_{cd}$ along the fluid flow lines is $\dot{T}^{ab}{}_{cd} \equiv T^{ab}{}_{cd;e} u^e$, the covariant derivative along u^a (the rate of change of $T^{ab}{}_{cd}$ as measured by a fundamental observer).

The first covariant derivative of the 4-velocity vector is

$$u_{a;b} \;=\; \omega_{ab} + \sigma_{ab} + \frac{1}{3}\Theta h_{ab} - \dot{u}_a u_b, \tag{2.3}$$

where $\Theta \equiv u^a{}_{;a}$ is the expansion, $\omega_{ab} = \omega_{[ab]}$ is the vorticity tensor ($\omega_{ab}u^b = 0$), and $\sigma_{ab} = \sigma_{(ab)}$ is the shear tensor ($\sigma_{ab}u^b = 0,\ \sigma^a{}_a = 0$). A representative length scale S along the flow lines is defined by

$$\frac{\dot{S}}{S} = \frac{1}{3}\Theta. \tag{2.4}$$

The vorticity and shear magnitudes are defined by $\omega^2 = \frac{1}{2}\omega_{ab}\omega^{ab}$, $\sigma^2 = \frac{1}{2}\sigma_{ab}\sigma^{ab}$.

As we restrict our attention to the case of a perfect fluid, the matter stress tensor takes the form

$$T_{ab} = \mu u_a u_b + p h_{ab}, \tag{2.5}$$

where μ is the energy density, $\mu = T_{ab}u^a u^b$, and the pressure $p = \frac{1}{3}h^{ab}T_{ab}$ (in the local rest frame of a comoving observer). In general, μ and p will be related through an equation of state.

In a FLRW universe model the shear, vorticity, acceleration, and Weyl tensor vanish, and the energy density μ, the pressure p and expansion Θ are functions of the cosmic time t only. Three simple gauge invariant quantities give us the information we need to discuss the time evolution of density fluctuations.

The first is the *comoving fractional density gradient*

$$\mathcal{D}_a \equiv S h_a{}^b \frac{\mu_{,b}}{\mu}, \tag{2.6}$$

which is gauge-invariant and dimensionless, and represents the spatial density variation over a fixed comoving scale [86]. Note that S, and so $\mathcal{D}_a$, is defined only up to a constant along each world-line by equation (2.6); this allows it to represent the density variation between any neighbouring world-lines. The time variation of this quantity precisely reflects the relative growth of density in neighbouring fluid comoving volumes.

The second is the *pressure gradient*

$$\mathcal{Y}_a \equiv h_a{}^b p_{,b} \,. \tag{2.7}$$

The third is the *comoving expansion gradient*

$$\mathcal{Z}_a \equiv S h_a{}^b \Theta_{,b} \,. \tag{2.8}$$

We can determine exact propagation equations along the fluid flow lines for these quantities, and then linearize these to the almost-FLRW case. The linearized equations are those given in [130] (see equations (13) to (19) there) plus the linearized propagation equations for the gauge-invariant spatial gradients defined above (see [86, 88] and [44]).

The basic equations are: the energy and momentum conservation equations (the time and space components of the 4-dimensional equation $T^{ab}{}_{;b} = 0$)

$$\dot{\mu} + (\mu + p)\Theta = 0, \tag{2.9}$$

$$(\mu + p)\dot{u}_a + \mathcal{Y}_a = 0; \tag{2.10}$$

the Raychaudhuri equation

$$\dot{\Theta} + \frac{1}{3}\Theta^2 + 2(\sigma^2 - \omega^2) + \frac{1}{2}\kappa(\mu + 3p) - \dot{u}^a{}_{;a} = 0, \tag{2.11}$$

where $\dot{u}^a{}_{;a}$ is the acceleration divergence; and the propagation equations for the gauge-invariant variables $\mathcal{D}_a$, $\mathcal{Z}_a$:

$$h_c{}^a(\mathcal{D}_a)^{\cdot} = -\mathcal{D}_a(\omega^a{}_c + \sigma^a{}_c) + \frac{p}{\mu}\Theta\mathcal{D}_c - (1 + \frac{p}{\mu})\mathcal{Z}_c, \tag{2.12}$$

$$\begin{aligned} h_c{}^a(\mathcal{Z}_a)^{\cdot} \;&=\; -\Theta\mathcal{Z}_c + h_c{}^a(\frac{1}{3}\Theta\mathcal{Z}_a - \frac{1}{2}\mu\kappa\mathcal{D}_a - 2S(\sigma^2)_{,a} + 2S(\omega^2)_{,a} + \\ &\quad S\dot{u}^b{}_{;ba}) - \mathcal{Z}_b(\sigma^b{}_c + \omega^b{}_c) + \dot{u}_c S\mathcal{R}, \end{aligned} \tag{2.13}$$

where $\mathcal{R} \equiv \frac{1}{3}\Theta^2 - 2\sigma^2 + 2\omega^2 + \kappa\mu + \dot{u}^a{}_{;a}$.

In the above, $\dot{u}^b{}_{;ba}$ stands for the gradient of the acceleration divergence, and $\kappa = 8\pi G$. Once the equation of state of the fluid is known the evolution of $\mathcal{Y}_a$ will follow from that for $\mathcal{D}_a$.

For completeness, we give also the propagation equation for the acceleration $a_c \equiv \dot{u}_c$:

$$h_a{}^c(a_c)^{\cdot} = a_a\Theta(\frac{dp}{d\mu} - \frac{1}{3}) + h_a{}^b\left(\frac{dp}{d\mu}\Theta\right)_{,b} - a_c(\omega^c{}_a + \sigma^c{}_a) \tag{2.14}$$

($\frac{dp}{d\mu}$ is taken along the fluid flow lines).

Further equations can be used as the basis of various systematic approximation schemes. The major point to notice is that in using equations (2.12) and (2.13) in an approximation scheme, to determine propagation of density inhomogeneities to the nth. order, we only need to solve the other equations of the model to the $(n-1)$th. order [108]. This gives the behaviour of the coefficients in these equations (and the Christoffel terms implied in the covariant derivatives on the left) to that order; then they directly determine the behaviour of inhomogeneities at the nth. order.

As equations (2.12) and (2.13) are gauge-invariant as well as covariant, we can use any coordinates and any convenient choice of background FLRW model in their further investigation. However equations (2.9) and (2.11) are not; we can e.g. deal with them by using an averaging procedure to determine a background model (see footnote[1] in this chapter).

2.1.2 Approaches of Futamase and Kasai

Futamase's approach

In a series of papers [104, 105, 106] Futamase developed an approximation scheme for describing an inhomogeneous universe, valid in non-linear case, basically with an arbitrary density contrast. This is a perturbative approach and the averaging introduced gives a clean separation between the global and local quantities.

The aim is to construct the approximate, reliable metric representing the real, clumpy universe in General Relativity. Obviously, the averaged, smooth metric coincides nowhere with the real inhomogeneous metric, but we know that the FLRW description is valid only in some averaged sense (if so). It seems then natural, to suppose that the space-time is close to a FLRW space-time, i.e., the inhomogeneous space-time is in a sense a small deviation away from the averaged smooth space-time which is not *a priori* given.

The crucial observation is the fact that the size of the metric perturbation and that of the density contrast are independent of each other in the exact theory, as well as in post-Newtonian approximations. For example, in the Solar System the metric coefficients in nearly orthonormal coordinates deviate, from their special relativistic values, by no more than $\sim 2GM_{\odot}/c^2R_{\odot} \sim 10^{-6}$, whereas the density contrast between the interior of the Sun, planets and interplanetary space is $> 10^{20}$.

The *ansatz* for the metric is taken as:

$$g_{\mu\nu} = a^2(\eta)(\tilde{g}^{(b)}_{\mu\nu} + h_{\mu\nu}), \tag{2.15}$$

where $h_{\mu\nu}$ are generated by local matter inhomogeneities and the gravitational waves (we neglect the latter), assumed to be small; this does not imply the smallness of density contrast. The scale factor $a(\eta)$ describes, as usual, the global FLRW expansion (averaged). In other words, $g_{\mu\nu}$ is the standard FLRW metric when $h_{\mu\nu} = 0$, with $\tilde{g}^{(b)}_{\mu\nu} = -d\eta^2 + d\Omega_3^2(k)$, where $d\Omega_3^2(k)$ is the standard metric on S^3 if $k = 1$ and on $\mathbb{R}^3$ if $k = 0, -1$.

The *ansatz* for the metric is such that the deviations from the FLRW models are small, but this does not imply that the zero-th order space-time is the FLRW one. It depends on the approximation chosen, e.g., within the linear approximation the zero-th order space-time is indeed taken to be the FLRW space-time. The approximation chosen depends on the kind of physical situation that one deals with, for the case at hand, what we have in mind is the matter clumps of various scales interacting gravitationally with each other and the density contrast between them and the mean density is $\gg 1$.

Above all, we have to restrict our space-times to those in which there is a well-defined meaning of the spatial average. The spatial averaging is therefore defined in a family of geometrically preferred slices, i.e., such that the metric deviation away from the FLRW metric is small everywhere on them.

The scheme is worked out in harmonic gauge:

$$\overline{h}^{\mu\nu}_{|\nu} = 0, \tag{2.16}$$

where $|$ stands for covariant derivation with respect to $\tilde{g}^{(b)}$, and $\overline{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\tilde{g}^{(b)}_{\mu\nu}h$, with $h = \tilde{g}^{(b)\mu\nu}h_{\mu\nu}$, is the so-called trace reversed perturbation.

The equations are derived under the assumption that h is small and that the scale on which h varies is small compared to that of a and $\tilde{g}^{(b)}$. The above *ansatz* is used to calculate the Einstein equations as follows (a prime stands for the derivative with respect to the conformal time η)

$$\begin{aligned}(\tfrac{a'}{a})^2(4\tilde{g}^{(b)\mu\eta}\tilde{g}^{(b)\nu\eta} - \tilde{g}^{(b)\eta\eta}\tilde{g}^{(b)\mu\eta}) - 2\tfrac{a''}{a}(\tilde{g}^{(b)\mu\eta}\tilde{g}^{(b)\nu\eta} - \tilde{g}^{(b)\eta\eta}\tilde{g}^{(b)\mu\nu}) + A^{\mu\nu} + \\ \tfrac{a'}{a}(2\overline{h}^{\eta(\mu|\nu)} - \overline{h}^{\mu\nu|\eta} - \tilde{g}^{(b)\eta(\mu}\overline{h}^{|\nu)} + \tfrac{1}{2}\tilde{g}^{(b)\mu\nu}\overline{h}^{|\eta}) - \tfrac{1}{2}\overline{h}^{\mu\nu|\rho}{}_{|\rho} = 8\pi G\tau^{\mu\nu},\end{aligned} \tag{2.17}$$

where $A^{\mu\nu}$ is the background spatial curvature term given by $A^{\eta\eta} = -3k\tilde{g}^{(b)\eta\eta} = 3k$, $A^{ij} = -k\tilde{g}^{(b)ij}$, and $A^{\eta i} = 0$. $\tau^{\mu\nu} = a^4T^{\mu\nu} + t^{\mu\nu}$ may be regarded as the effective (total) stress-energy tensor, i.e., material stress-energy tensor $T^{\mu\nu}$ plus gravitational stress-energy pseudotensor $t^{\mu\nu}$ consisting of terms quadratic in $\overline{h}$.

Effectively, this means that the Einstein equations are expanded in terms of two small parameters, ϵ and κ, whose meaning is the following:

- ϵ is the size (amplitude) of the metric perturbation h and it is assumed that $h,\ h_{\mu\nu} \simeq \mathcal{O}(\epsilon^2)$, $h_{\mu\nu,\rho} \simeq \mathcal{O}(\epsilon^2/l)$; ϵ defined this way is also an amplitude of the peculiar gravitational potential Φ, the gravitational potential generated by the inhomogeneous distribution of matter, where $\Phi \simeq \mathcal{O}(\epsilon^2)$, as well.

- $\kappa = \frac{l}{L}$ is the ratio between the scale of the variation of h and that of a and $\tilde{g}^{(b)}$, where for L we can take the present horizon size $\sim 10^4\,Mpc$, and for l the typical scale of inhomogeneities; note that $\kappa \in [0,1]$, with small κ indicating condensed density contrast, large κ - diffused density contrast.

The relative size of κ and ϵ depends on the physical system considered. It is straightforward to see that the density contrast is of the order of ϵ^2/κ^2, and consequently the linear stage is characterised by $\kappa \gg \epsilon$, and in the non-linear stage we have $\epsilon \gg \kappa$. For galaxies we typically have $\epsilon > \kappa$, and the ratio ϵ/κ increases when we consider smaller regions. Basically, if the typical size of inhomogeneities is larger than a galactic scale, the approximation is valid in the parameter range $\epsilon^2 \ll \kappa$.[2]

In deriving (2.17), terms like $\frac{a''}{a}\bar{h} \simeq \mathcal{O}(\epsilon^2/L^2)$, $\frac{a'}{a}\bar{h}_{|\rho}\bar{h} \simeq \mathcal{O}(\epsilon^4/lL)$, $\bar{h}_{|\rho}\bar{h}_{|\sigma}\bar{h} \simeq \mathcal{O}(\epsilon^6/l^2)$ and of higher orders, were neglected, whereby it was taken that $\frac{a'}{a} \simeq \mathcal{O}(1/L)$ and $\frac{a''}{a} \simeq \mathcal{O}(1/L^2)$. Physically, it means that the effect of the self-gravity of clumps on their dynamics is more important than the expansion of the universe. The neglected terms are negligible as far as $\epsilon, \kappa \ll 1$ and $\kappa \gg \epsilon^2$.

A perfect fluid is taken for the material source which is characterized by the density field ρ, its peculiar velocity $\vec{v}$, and peculiar gravitational potential Φ;

$$T^{\mu\nu} = [\rho + p(\rho)]u^\mu u^\nu + p(\rho)g^{\mu\nu}. \tag{2.18}$$

One works with conformally rescaled variables: $\tilde{u}^\mu = au^\mu$, $\tilde{g}^{\mu\nu} = a^2 g^{\mu\nu}$. Then $\tau^{\mu\nu} = a^2\tilde{T}^{\mu\nu} + t^{\mu\nu}$, where $\tilde{T}^{\mu\nu} = [\rho + p]\tilde{u}^\mu\tilde{u}^\nu + p\tilde{g}^{\mu\nu} = a^2 T^{\mu\nu}$.

On very large scales the universe is assumed to be homogeneous in space. In the next step the spatial averaging is applied to the truncated Einstein field equations (2.17), assuming spatial periodicity of the material initial data[3] (as well as that of the free data for the gravitational field) and no coherent motion over the volume to be averaged[4].

The spatial average over a volume V is defined as:

$$< Q >= \frac{1}{V}\int_V Q\sqrt{\tilde{g}^{(b)}}d^3x, \tag{2.19}$$

where $\tilde{g}^{(b)}$ is the determinant of the spatial part of the background metric $\tilde{g}^{(b)}_{\mu\nu}$ and $\sqrt{\tilde{g}^{(b)}}d^3x$ is the invariant volume element in the background space, and for the density we have $< \rho >= \rho_b$ (background density). The only property used in a calculation is $< Q_{|i} >= 0$, which is implied by the spatial periodicity. Also, $< \tau^{\eta i} >= 0$, which just means no

[2]In [106] an approximate metric is constructed in the situation with strong gravity and/or smaller regions of inhomogeneity, where $\epsilon^2 \gg \kappa$, but we will not discuss this case here.

[3]Space-time averaging is another possibility; in [139] temporal periodicity for the inhomogeneities had to be assumed in order to get a separation between global and local evolution.

[4]This is almost always safe, i.e. with large enough averaging volumes and randomly distributed perturbations.

coherent motion over the averaging volume. The spatial average of (2.17), under the above requirements, implies $< \overline{h}^{\eta i} >= 0$, and of (2.16) $< \overline{h}_{\eta\eta} >= const.$ which under a redefinition of the time variable and scale factor can be put to zero. Also, $< \overline{h}^{k}{}_{k} >$ can be put to zero, since it expresses an additional isotropic expansion which can be absorbed into the scale factor upon its redefinition.

The averaged sources are used to calculate the global expansion and the following averaged Einstein equations are obtained from (2.17) to the first non-trivial order, (by non-trivial order we mean the order at which the first backreaction effect due to inhomogeneities on the expansion of the universe appears),

$$\begin{aligned} (\tfrac{a'}{a})^2 &= \tfrac{8\pi G}{3} < \tau^{\eta\eta} > -k \\ \tfrac{a''}{a} &= \tfrac{4\pi G}{3} < \tau^{\eta\eta} - \tau^{k}{}_{k} > -k \\ \tfrac{1}{a^2}(a^2 < \overline{h}^{ij} >_{|\eta})_{|\eta} &= 16\pi G < \hat{\tau}^{ij} >, \end{aligned} \tag{2.20}$$

where $\hat{\tau}^{ij} = \tau^{ij} - \frac{1}{3}\tilde{g}^{(b)ij}\tau^{k}{}_{k}$ is the trace free part of τ^{ij}.

The averaged line element,

$$< ds^2 >= a^2(\eta)(-d\eta^2 + (\delta_{ij}+ < \overline{h}_{ij} >)dx^i dx^j), \tag{2.21}$$

tells us that the averaged space-time expands anisotropically, except when $< \overline{h}_{ij} >$ vanishes identically, since $< \overline{h}_{ij} >$ expresses the deviation from the isotropic expansion due to the inhomogeneities $< \hat{\tau}^{ij} >$. Note that the first two equations of (2.20) are the Friedmann equations with source terms replaced by the effective stress-energy tensor, thus the effect of the local inhomogeneities on the global expansion can be partly expressed by the effective density and pressure $\rho_{eff} = a^2 < \tau^{\eta\eta} >$, $p_{eff} = \frac{1}{3}a^2 < \tau^{k}{}_{k} >$.

One can integrate the last equation of the system (2.20) to obtain the expression for $< \overline{h}^{ij} > (\eta)$, and see that a sufficient condition for global isotropic expansion is given by $< \overline{h}^{ij} >_{,\eta} (\eta_o) \equiv 0$ and $< \hat{\tau}^{ij} >\equiv 0$. The equations determining the evolution of the local inhomogeneities are derived by substituting the above equations into (2.17), additionally we have also the equations of motion (derived from the conservation of the stress-energy tensor).

Now, we can employ a particular approximation scheme. The evolution of the density perturbations in our picture is sufficiently well described by means of a post-Newtonian approximation. The post-Newtonian approximation is characterized by small parameter ϵ, of the order of a typical peculiar velocity divided by the speed of light. It is introduced by a coordinate transformation $\eta_N = \epsilon\eta$, η_N is the Newtonian time which means physically that a typical time scale gets longer as ϵ^{-1} as the velocity goes to zero. This parameter is identified with the already introduced ϵ. The orders for material variables are assumed to be $\rho_N = \epsilon^{(-2)}\rho$, $v^i_N = \epsilon^{(-1)}v^i$, $p_N = \epsilon^{(-4)}p$. The other small parameter, κ is associated with global cosmic expansion (when $\kappa \to 0$, the expansion of the universe slows down). ϵ and κ are our order parameters, in the sense, that they parameterize a sequence of space-times and we study the Newtonian limit on that sequence.

The evolution equations for the local inhomogeneities are then solved perturbatively up to the first non-trivial order.

An interesting outcome of the application of this approximation scheme is that the backreaction leads to an underestimation of the age of the universe, as inferred from a measurement of today's Hubble constant [30] (also see [29]). For a simple model within the framework of pancake theory for structure formation on a flat expanding background, it is shown in [30] that the age problem (severe in view of the recent determinations of globular cluster ages) may be solved by taking into account the backreaction of inhomogeneities in an averaged sense. This scheme can also be used for the correct interpretation of observations of gravitational lenses.

Kasai's approach

Kasai's scheme [149] to construct inhomogeneous relativistic universes which are homogeneous and isotropic on average, goes further than Futamase's approach. It is not assumed here that the deviations from a FLRW model are small to acquire FLRW-like behaviour on average.

Here also, spatial averaging is introduced, but the description is based on the deformation tensor and can give yet another possibility to obtain solutions to describe more realistic situations, and on the other hand, to formulate a relativistic version of the Zel'dovich approximation used to handle the evolution of the large scale structure in Newtonian cosmology.

In [149] the inhomogeneous irrotational dust universe models are constructed in the framework of General Relativity, with the property of being homogeneous an isotropic on average. The averaging is introduced for matter only, on the hypersurfaces Σ_t orthogonal to dust motion, and the mean ("background") density for the inhomogeneous universe model is written as

$$\rho_b =< \rho >\equiv \lim_{V\to\Sigma_t} \frac{1}{\int_V [\det(g_{ij})]^{1/2} d^3x} \int_V \rho[\det(g_{ij})]^{1/2} d^3x, \tag{2.22}$$

for $V \subset \Sigma_t$ (assuming that this limit exists).

The "scale factor" (averaged) is then defined by

$$\dot{\rho}_b + 3(\frac{\dot{a}}{a})\rho_b = 0. \tag{2.23}$$

The peculiar deformation tensor is introduced as

$$V^i_j \equiv u^i_j - \frac{\dot{a}}{a}\delta^i_j. \tag{2.24}$$

This quantity describes the deviation from a uniform Hubble expansion. As usual, u^μ is a 4-velocity and comoving coordinates are used. The deformation tensor $u^i_j \equiv u^i_{;j}$, describing the change of the relative position X^i between the world lines of neighbouring "particles" (galaxies), $\dot{X}^i = u^i_j X^j$ is at the same time, the extrinsic curvature of the $t = const.$ hypersurfaces Σ_t.

Another quantity introduced is a density contrast[5]

$$\Delta \equiv \frac{\rho - \rho_b}{\rho}. \tag{2.25}$$

Given this, the Einstein equations are the following:

$$\dot{\Delta} + (1-\Delta)V^i_i = 0 \tag{2.26}$$

$$(\frac{\dot{a}}{a})^2 = \frac{8\pi G}{3}\frac{\rho_b}{1-\Delta} - \frac{1}{6}{}^{(3)}R - \frac{2}{3}\frac{\dot{a}}{a}V^i_i - \frac{1}{6}[(V^i_i)^2 - V^i_j V^j_i] \tag{2.27}$$

[5]Note that it differs form the conventionally adopted $\delta \equiv \frac{\rho-\rho_b}{\rho_b}$.

$$\ddot{\Delta} + 2(\frac{\dot{a}}{a})\dot{\Delta} - 4\pi G\rho_b\Delta = -(1-\Delta)[(V_i^i)^2 - V_j^iV_i^j], \tag{2.28}$$

where $\dot{} \equiv \frac{\partial}{\partial t}$, ${}^{(3)}R = {}^{(3)}R_i^i$ and $i,j = 1,2,3$.

FLRW model (with ρ_b and a) is the background model. Note that (2.27) reduces to the Friedmann equation, $(\frac{\dot{a}}{a})^2 = \frac{8\pi G}{3}\rho_b - \frac{k}{a^2}$, when there are no inhomogeneities if and only if the condition

$$\frac{8\pi G}{3}\rho_b\frac{\Delta}{1-\Delta} - \frac{1}{6}{}^{(3)}R + \frac{k}{a^2} - \frac{2}{3}\frac{\dot{a}}{a}V_i^i - \frac{1}{6}[(V_i^i)^2 - V_j^iV_i^j] = 0 \tag{2.29}$$

holds.

On the other hand, when the left hand side of (2.28) is zero, one gets the evolution equation for $\delta \equiv \frac{\rho-\rho_b}{\rho_b}$ in the linear perturbation theory.

The nice property of this approach is the fact that equations similar to (2.26) and (2.28) appear in Newtonian cosmology in the context of extending the Zel'dovich-type approximations, with V_j^i spatial gradient of peculiar velocity $v^i_{,j}$. In particular, when $(V_i^i)^2 - V_j^iV_i^j = 2(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1) = 0$, where λ_i are eigenvalues of V_j^i, Δ obeys the same equation as δ and the solutions can be extrapolated from the results of the linear perturbation theory. Thus the present construction can represent "relativistic pancake solutions" analogous to those in Newtonian cosmology.

2.2 Exact non-covariant averaging

Despite the fact that quite often the averaging procedure is applied on top of a perturbative scheme, we will review in this chapter approaches in which the averaging (by volume in all cases) is defined (non-covariantly) without any recourse to approximations. There is no covariant definition of spatial averaging available in a general space-time.

The earliest attempt is due to Shirokov and Fisher [229], where they proposed to define the components of the averaged (macroscopic) metric as the volume averages of the corresponding components of the small scale (microscopic) metric. This procedure is explicitly non-covariant (volume integrals of tensor components do not constitute a tensor!) and has no geometrical interpretation. It is therefore not clear what is actually represented by a volume average of a metric component. Nevertheless, this definition was applied to metrics which are small perturbations of the FLRW models and from them the Einstein tensor was calculated, as well as the averages of all its components. They were then equated to the averages of the appropriate components of the energy momentum tensor. Terms non-linear in the small quantities were neglected. The result they obtained was a generalization of the FLRW solutions, with a repulsive term preventing the singularity for all three curvatures. At a maximal crunch, each particle fills the interior of a sphere of a radius equal to that of the particle's own gravitational radius.

Further in [230] (which is really only a short conference report on work in progress), the author considered the following deformation of the FLRW metrics:

$$ds^2 = dt^2 - (1 + \frac{1}{4}kr^2)^{-2}G^2(y)f(t,x,y,z)(dx^2 + dy^2 + dz^2), \tag{2.30}$$

where $G(y)$ and $f(t,x,y,z)$ are unknown functions. The consequences of averaging of the Einstein tensor (calculated from an "averaged" metric), with the assumption that it obeys Cosmological Principle, are calculated exactly without approximations. The term due to averaging was interpreted as a negative contribution to pressure capable of preventing the Big Bang.

In a series of papers, Saar [217] considered the influence of averaged rapid fluctuations on a slowly changing background metric. The metric (without considering any explicit form of it) was split into the background and a fluctuation times a small parameter. Einstein equations were then developed to the second order in the small parameter and averaged by volume. As a result, the contribution of averaged fluctuations can be interpreted as a negative pressure in the background and cosmological expansion can proceed slower, with more time available for structure formation.

The result of Nelson [197], on the other hand, seems to contradict the others. The metric was split into smooth background and small perturbations describing lumps, and then averaged by integrating metric components over volume. The background was obtained to be approximately equal to the average of the whole metric, with the average obeying *"the usual set of cosmological equations"*. The corrections to the average obey equations *"equivalent to instability equations. The large scale development of such a Universe is* *therefore shown to be almost independent of the formation of condensations provided the average of the energy-stress tensor is unaffected by the condensations".* The relevant assumptions were that the averaging volume contains many condensations, the number of condensations is big, they are evenly distributed in space and their radii large with respect to the Schwarzschild radius.

Interesting approach was put forward in [181]. A series of following papers by Marochnik and co-workers, initiated in fact properly a direction of research taking into account the small scale inhomogeneities.

The metric, the density and the velocity field were taken to be sums of an average (background) and a correction (called turbulence) whose average is zero. The following equalities were assumed to hold

$$\lambda_T^2 < L^2 \ll 1/\bar{R}, \tag{2.31}$$

where λ_T is the characteristic scale of the turbulence, L the averaging scale, and $\bar{R}$ the background curvature. The effect of averaging in terms of volume integrals on the Einstein equations was examined. They claimed that corrections to energy and pressure (due to averaging out small scale inhomogeneities) do not need to be positive and may be interpreted formally as antigravitation. This can further lead, for example, to a situation where perturbations of a FLRW background can yield a non-expanding, static universe; or the corrections due to averaging might prevent the Big Bang singularity.

Further on, in [182] the results of averaging on linearized perturbations of the FLRW models were calculated and it turns out that:

1. [6] For $p = \rho/3$ (valid in the FLRW background)
   - if the potential turbulence exceeds the vortical one [170], then the perturbations decelerate the expansion and the initial singularity can be avoided. (The authors notice however that before the minimal radius is approached the linear approximation is invalidated.)
   - if the vortical turbulence dominates the expansion is accelerated (the singularity remains).

[6] It was noted in [182] that this result is in contradiction to [28], where the main reason for deviation from the Friedmann equation was attributed to the gravitational waves.

2. For other equations of state, the turbulence could modify the evolution of the universe when $p \geq \rho/3$. In this regime ($p > \rho/3$) however, the linear approximation does not hold any more.

In [183] the influence of long-wave turbulence on the background expansion was examined, taking into account only those modes of turbulence that remain finite when close to singularity. It turned out that they influence the expansion in the same way as in the perturbative solution of Lifshitz and Khalatnikov [171].

In [184] the equations of [182] were written up to the second approximation in the turbulent perturbation and solved in [185], where also the effects of the perturbations on the background were studied. No pronounced qualitative effects were found though.

In [186] the changes in the most important cosmological parameters due to averaged out small scale inhomogeneities were calculated and the author found that this affects the transition moment between the hadron and lepton eras (by a factor of up to 1.4); the temperature in the transition moment (by a factor of 0.88); and the helium abundance[7].

One of the most recent papers considering the effect of averaging over spatial volumes on the Einstein equations is the one due to Zotov and Stoeger [266]. They simply compare an exact FLRW model with the one where galaxies, represented by a Schwarzschild metric, are superposed on the FLRW background with a constant number density. They find that upon averaging the metric components, the background FLRW model with the scale factor $R(t)$ changes to another FLRW model with the scale factor $S(t)$, and

$$S^2(t) = R^2(t)(1-K), \tag{2.32}$$

where $K = NV_1/V_2$, N is the number density of galaxies, V_1 the averaging volume, V_2 the volume per one galaxy in the space. The effect of matter, as calculated through the average values of metric components, is therefore to squeeze the space volume. If calculated through substituting the new average density in the Einstein equations the effect is the same, but now the dependence of $R(t)$ on the density parameter σ_o is given in the form of parametric equations

$$R_h(t) = H_o\sigma_o(1-2\sigma_o)^{-3/2}(\cosh 2\psi - 1),$$

[7] It may remain unchanged if the energy density of short wave fluctuations is smaller than $1.5\bar{\rho}$ (where $\bar{\rho}$ is the large scale average energy density).

$$ct = H_o\sigma_o(1-2\sigma_o)^{-3/2}(\sinh 2\psi - 2\psi).$$

The authors do not calculate nor discuss in their paper the terms due to averaging in the Einstein equations.

2.3 Exact covariant averaging

Isaacson was the first to consider the problem of covariant averaging involved in "coarse-grain" viewpoint. In [139] he considered the vacuum Einstein equations in the short-wave approximation, assuming that the metric of a gravitational wave space-time can be split into a low frequency background and high frequency wave, namely, $g_{\mu\nu} = g^{(B)}_{\mu\nu} + h_{\mu\nu}$. High frequency means small wavelength with respect to the curvature radius of the background, the smallness of the waves' amplitude was assumed. The Einstein equations were then linearized with respect to the high frequency correction and the waves were shown to obey a covariant generalization of the equation of massless spin 2 fields in flat background. They were shown to travel on null geodesics of the background, their amplitude, frequency and polarization modified by the background curvature.

Further in [140] the same considerations were carried out to a higher order of approximation. The author showed that the corrections of the first non-linear order to the vacuum Einstein equations provide a term ("stress-energy" tensor) that can be interpreted as the effective energy of the gravitational waves. It was then used to define the total energy and momentum carried off to infinity by the waves. The stress-energy tensor of gravitational waves was shown to be well defined only in a smeared out sense (cf. also [192]).

The energy momentum tensor of the gravitational waves arises once the metric is averaged. The average is defined by parallel-transporting the tensors from the point x' to the representative point x, along the geodesic between x' and x, and then integrating the resulting object with respect to x' with a weighting function $f(x,x')$ (this function was not defined explicitly):

$$< T_{\mu\nu}(x) >= \int_{all\ space} G^{\alpha'}_{\mu}(x,x') G^{\beta'}_{\nu}(x,x') T_{\alpha'\beta'}(x') f(x,x') d^4x', \tag{2.33}$$

where $G^{\alpha'}_{\beta}$ are the propagators of parallel displacement. From the weighting function it was demanded that:

1. $f(x,x') \to 0$, when the distance $d(x,x')$ obeys $\lambda \ll d \ll L$, where λ is the wavelength of (high frequency) waves and L is the wavelength of the low frequency background.

2. $\int_{all\ space} f(x,x')d^4x' = 1$.

This definition was applied within a perturbative scheme but it is perfectly covariant.

The evaluation of the effective stress-energy tensor for the gravitational waves requires averaging of various quantities over several wavelengths. From (2.33) one can derive Isaacson's averaging rules (he names the averaging scheme used as "Brill-Hartle averaging" after [41]) which are the following:

- covariant derivatives commute, e.g. $< h\, h_{\mu\nu|\alpha\beta} >=< h\, h_{\mu\nu|\beta\alpha} >$, where $|$ stands for the covariant derivative with respect to $g^{(B)}_{\mu\nu}$;
- gradients average out to zero, e.g. $< (h_{|\alpha} h_{\mu\nu})_{|\beta} >= 0$;
- one can integrate by parts flipping derivatives from one h to the other, $< h\, h_{\mu\nu|\alpha\beta} >=< -h_{|\beta} h_{\mu\nu|\alpha} >$.

In harmonic gauge, $\bar{h}_\mu{}^\alpha{}_{|\alpha} = \bar{h} = 0$, the stress-energy for the gravitational waves was shown to have the following form

$$T^{GW}_{\mu\nu} = \frac{1}{32\pi} < \bar{h}_{\alpha\beta|\mu} \bar{h}^{\alpha\beta}{}_{|\nu} >, \tag{2.34}$$

where $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} h g^{(B)}_{\mu\nu}$.

Further improvement was due to Matzner. In [188] he does not need the assumption that the wave is of high frequency in order to identify the background. Moreover, he gives another definition of averaging the geometry, namely, a metric is Lie-dragged along a specific vector field, defined by

$$\xi^{(\alpha;\beta)}{}_{;\beta} + \lambda \xi^\alpha = 0 \tag{2.35}$$

to a chosen point, averaged there over all points and then Lie-dragged back. Equation (2.35) is in fact a generalization of the Killing equation. The above definition was applied to the $t = const.$ sections of the Taub-NUT space (space of finite volume). It turned out, that the averaged metric has the FLRW algebraic form, however it is not in any simple way related to the original metric (e.g. the volume of the averaged space is different from the initial volume).

In [189] a measure of symmetry in a riemannian manifold with a positive-definite metric was put forward. Namely, it was defined as the minimum value of the functional:

$$\lambda[\xi] = (\int_{all\ space} \xi^{(\alpha;\beta)}\xi_{(\alpha;\beta)}dv)/(\xi^{\mu}\xi_{\mu}dv), \tag{2.36}$$

where ξ^{α} is a vector field, $dv = \sqrt{g}d^{n}x$. For compact space, or when the integral over the boundary at infinity is zero, the minima of $\lambda[\xi]$ obey (2.35); on a positive-definite manifold $\lambda = 0$ in (2.35) implies Killing equation. The above definition was applied to some space-times and the parameters λ, defined by (2.35), could be interpreted as averaged energy-density and averaged stresses of gravitational waves.

In [211] Rosen pointed out that for a stochastic stress-energy tensor associated with cosmic turbulence the Einstein equations imply fluctuations in the space-time metric tensor (of a purely classical-statistical character). He then showed that averaging the metric produces corrections to the energy momentum tensor and calculated them explicitly, assuming that: (1) the perturbed metric is conformally equivalent to the background metric, and (2) the averaged energy-momentum tensor has the algebraic form of a perfect fluid. The conclusion reached was that fluctuations in the metric always accelerate the expansion in a FLRW background (i.e. increase $\ddot{R}/R$). However, no definition of averaging was proposed and the author called his averages *"with respect to the statistical ensemble"*; the whole treatment was rather axiomatic. This paper nevertheless appears to be one of the first tackling the problem from a geometric viewpoint.

Further, there is the approach of Carfora and Marzuoli [51] (which will be reviewed in section 2.4.1). We will only mention here that they were the first to confirm exactly and in a covariant manner the result predicted by Shirokov and Fisher [229].

In [152] Khiet, influenced by Shirokov and Fisher paper, studied the results of averaging microscopic gravitational equations and without any definition of an averaging procedure, simply guessed the correction term $C_{\alpha\beta}(g)$ in the macroscopic equations $G_{\alpha\beta} + C_{\alpha\beta} =$

$\kappa T_{\alpha\beta}$ on the basis of covariance requirements. He obtained new, non-linear in curvature, field equations and studied their exact solutions for a general FLRW metric. These were however field equations of a new metric gravitation theory, instead of averaged Einstein equations.

There is yet another approach due to Arifov and Shayn [7] which is similar in spirit to Khiet's approach. In this approach it is in fact enough to introduce averages formally.

The average curvature $\bar{r}_{\alpha\beta\gamma\delta}$ and metric tensor $\bar{g}_{\alpha\beta}$ are no longer the curvature and metric tensors of the same riemannian manifold. It was proved that when dim $n \leq 4$, there exists the riemannian manifold with the metric tensor $G_{\alpha\beta}$ of the same signature as $g_{\alpha\beta}$ and curvature tensor $R_{\alpha\beta\gamma\delta} \equiv \bar{r}_{\alpha\beta\gamma\delta}$, which is not riemannian with $\bar{g}_{\alpha\beta}$. Effectively, this is a bi-metric gravitation theory. The macroscopic Einstein equations were obtained upon averaging out of the left hand side of Einstein's equations and extracting the Einstein tensor (made up of the averaged curvature tensor), namely,

$$R^{\alpha}_{\beta} - \frac{1}{2}\delta^{\alpha}_{\beta}R = -T^{\alpha}_{\beta} + D^{\alpha}_{\beta}, \tag{2.37}$$

where D^{α}_{β} is defined as

$$D^{\alpha}_{\beta} \equiv D'^{\alpha}_{\beta} + \delta^{\alpha}_{\beta}(H^{\mu\nu}R_{\mu\nu} + \frac{1}{2}H^{\rho\sigma}H^{\mu\nu}R_{\rho\mu\nu\sigma}) - H^{\mu\nu}R^{\alpha}_{\mu\nu\beta} - H^{\alpha\rho}R_{\rho\beta} - H^{\alpha\rho}H^{\mu\nu}R_{\rho\mu\nu\beta}, \tag{2.38}$$

with further definitions following simply

$$D'^{\alpha}_{\beta} \equiv \frac{1}{2}\delta^{\alpha}_{\beta}(\bar{r} - \bar{g}^{\rho\sigma}\bar{g}^{\mu\nu}\bar{r}_{\rho\mu\nu\sigma}) - (\bar{r}^{\alpha}_{\beta} - \bar{g}^{\alpha\rho}\bar{g}^{\mu\nu}\bar{r}_{\rho\mu\nu\beta}) \tag{2.39}$$

for the correlation tensor, and $H^{\alpha\beta} \equiv \bar{g}^{\alpha\beta} - G^{\alpha\beta}$, $\bar{g}^{\alpha\beta}\bar{g}_{\gamma\beta} = \delta^{\alpha}_{\gamma}$, $G^{\alpha\beta}G_{\gamma\beta} = \delta^{\alpha}_{\gamma}$.

It can be argued whether any real averaging has been achieved this way, in any way no geometrical considerations were offered. Consequently, the correlation tensor could only be modeled phenomenologically. For example, in the case of gravitational field of the spherically symmetric source $D^{\alpha}_{\beta} = \sigma C^{\alpha\mu\nu\sigma}C_{\beta\mu\nu\sigma}$, where $C_{\alpha\mu\nu\sigma}$ is the conformal curvature tensor and $\sigma = const.$ The solution found in this case was argued to yield a macroscopic metric that has neither an event horizon nor a singularity for certain values of the parameter σ (contrary to the usual Schwarzschild solution).

The problem of averaging is very much related to the issue of approximating a fine-scale cosmological model with a large-scale model. In [234] Spero and Baierlein proposed an independent approach, namely, to define an approximate symmetry of an inhomogeneous model by "best-fitting" to it a Bianchi-type model. This best-fit was defined in terms of a minimum of a functional with respect to variations of a triad of orthonormal vectors in the given space-time and variations of the structure constants, to be found. As shown, the resulting Bianchi type is not always unique unless it is one of the generic types: IX, $VIII$, VII_h or VI_h. The classification depends on the slicing and is not necessarily preserved in time. Moreover, the approximating type of space-time is not guaranteed to obey Einstein equations.

The second paper of the same authors [235] is basically an application of the earlier ideas to two slicings of the Gowdy solutions [117], providing approximants of Bianchi type I and VI_o, and to Kantowski-Sachs metric with the approximant Bianchi type I.

Another interesting paper is that by Stoeger, Ellis and Hellaby [238], where they proposed a criterion of continuous homogeneity of the universe, namely, if the mean mass density[8] in a sphere of volume V_L, centered at the point $\vec{r}$, is given by

$$\bar{\rho}_L(\vec{r}) = \frac{1}{V_L}\int_{V_L} \rho(\vec{r'})\, d^3\vec{r'}, \tag{2.40}$$

where V_L is assumed to be small enough so that the curvature inside V_L does not need to be taken into account. Then we say that the density distribution is spatially homogeneous on average at the level ϵ, on scales larger than L_c, if and only if, $\exists\ \epsilon \ll 1$ and L_c such that:

$$|\bar{\rho}_{L_1}(\vec{r_1}) - \bar{\rho}_{L_2}(\vec{r_2})| < \epsilon\bar{\rho}_{L_1}(\vec{r_1}) \tag{2.41}$$

for all $\vec{r_1}$, $\vec{r_2}$ and all L_1, $L_2 \geq L_c$. In principle, this criterion is falsifiable by observations. As an example, they considered galaxies randomly distributed in space according to Poisson distribution, and showed that without further assumptions such a distribution is not in agreement with (2.41) and so cannot be described by a FLRW model in this sense. In the observed Universe one probably has $L_c \approx 200\,Mpc$ and $\epsilon \leq 0.01$.

2.4 Smoothing of cosmological spacetimes

The task, we are interested in, can be rephrased as setting up a program for approximating the evolution of cosmological space-time solutions of Einstein's equations via the

[8] This criterion applies to any scalar.

development of a procedure for "smoothing" sets of initial data for such space-times.

Looked at this way, smoothing is equivalent to a physical approximation scheme for particular space-times. The idea is the following. Given an initial data set: the spatial metric g, the extrinsic curvature K, and matter fields ψ, one would like to build a new smooth (i.e. spatially homogeneous) initial data set $(\overline{g}, \overline{K}, \overline{\psi})$, so that the new initial data is more easily evolved than the old one, and at the same time the evolution of new initial data models certain aspects of the evolution of the original initial data.

On the other hand, one can think of smoothing as a mathematical method for making general statements about a collection of space-times. The smoothing procedure could then be used as a map from general space-times to the spatially homogeneous ones, in order to study the space of space-times, in particular, the collection of space-times whose large scale dynamics are closely represented by the dynamics of spatially homogeneous space-times [141].

The flows of the metric are an important part of the smoothing we have in mind, namely, the Ricci-Hamilton flow (facts concerning it are given in Appendix B) and the Renormalization Group flow (see section 3.6).

General mathematical preliminaries can be found in Appendix A.

2.4.1 Smoothing-out spatially closed cosmologies

In [51] (see also [53]) a specific smoothing-out procedure was put forward, deforming a family of locally inhomogeneous and anisotropic spatially closed space-times into closed FLRW universes. These space-times are associated with gravitational configurations which can be considered near to the standard ones generating closed FLRW cosmological models. This class is large, it contains solutions to the Einstein field equations that are not just perturbations of closed FLRW space-times.

The smoothing-out procedure is employed in the full theory, and a precise content to the averaging hypothesis, by providing explicitly the correction terms to the physical sources induced upon smoothing-out the space-time geometry, can thus be given.

The idea is the following. We pick up an appropriate initial data set, on a closed spacelike hypersurface, which upon the Cauchy evolution is going to be the space-time to be averaged out. Such data set is then smoothly deformed into a FLRW initial data set, by the action of parabolic-type operators. This deformation is constructed in such a way as to make the deformed data satisfy the four constraints associated with Einstein's equations. It follows then that the flow of deformed initial data generates a one parameter family of solutions to the field equations, which interpolates between the original space-time and a closed FLRW space-time, considered to be the smoothed-out counterpart of the given universe model.

To make the above precise, let $({}^{(4)}V \overset{\phi}{\simeq} \mathcal{M} \times I,\, {}^{(4)}g)$ be the space-time manifold, the Cauchy evolution of a regular initial data set $(\mathcal{M}, g, K)$, where ϕ is a diffeomorphism mapping ${}^{(4)}V$ to $\mathcal{M} \times I$ ($I \subset \mathbb{R}$), with $\mathcal{M}$ being the (closed) 3-manifold carrier of the initial data (i.e., a space-like 3-hypersurface in the space-time manifold) and $g, K \in S^2\mathcal{M}$, representing in the final space-time the induced riemannian 3-metric on $\mathcal{M}$, and the second fundamental form of the embedding $\mathcal{M} \to ({}^{(4)}V,\, {}^{(4)}g)$, respectively. We assume that $\mathcal{M}$ is topologically a 3-sphere S^3 and that the class of initial data supported by $\mathcal{M}$ is such that $Ric(g)$ is a positive definite bilinear form for them ($Ric(g)$ is the Ricci tensor associated with g).

Due to the results of R. Hamilton, (see Appendix B), $({}^{(4)}V,\, {}^{(4)}g)$ resulting from the time evolution of data from the above class, can be taken as modeling a locally anisotropic and inhomogeneous universe not too far from a closed FLRW space-time. A smoothing-out mapping associates with the given initial data set a one parameter family $(\mathcal{M}, g(\beta), K(\beta))$ with $0 \leq \beta < \infty$, $g(0) = g, K(0) = K$, approximating, closer and closer, the standard initial data set for a closed FLRW model and reaching it uniformly as $\beta \to \infty$ (β is the parameter labelling the family (flow) of 3−metrics).

According to Hamilton's theorem, we can deform the metric g into the constant-curvature metric $\overline{g}$ on S^3 by the flow of metrics $g(\beta)$, $0 \leq \beta < \infty$, solution to the non-linear, weakly parabolic, initial value problem

$$\frac{\partial}{\partial \beta} g_{ab}(\beta) = \frac{2}{3} < R(\beta) >_\beta g_{ab}(\beta) - 2R_{ab}(\beta), \tag{2.42}$$

with $g_{ab}(0) = g_{ab}$, $(a, b = 1, 2, 3)$, where $< R(\beta) >_\beta$ is the average scalar curvature over $(\mathcal{M}, g(\beta))$, and $R_{ab}(\beta)$, $R(\beta)$ are the components of the Ricci tensor and the scalar curvature associated with $g(\beta)$, respectively. We recall here that the family (2.42) has the following properties (for more details see Appendix B):

1. the volume of $\mathcal{M}(\beta)$ is independent of β;

2. any symmetries of $g_{ab}(\beta_o)$ are inherited by all $g_{ab}(\beta)$ with $\beta \geq \beta_o$, and

3. the limiting smoothed metric $\bar{g}_{ab} = \lim_{\beta\to\infty} g_{ab}(\beta)$ has positive constant curvature.

Now, g_{ab} is the inhomogeneous metric to be smoothed-out. Equation (2.42) (with the initial condition) defines a smooth family of deformations of the initial manifold, deforming it into a 3−space of constant curvature and of the same volume as the initial manifold.

In order to smooth-out the whole data set, we need to average out the second fundamental form K, as well. Obviously, we have for the smoothed metric $\overline{g}_{ab} = \lim_{\beta\to\infty} g_{ab}(\beta)$ (presuming that the flow converges), where $g_{ab}(\beta)$ satisfies the Ricci flow equation (2.42). Given (g, K), let us then define a nearby flow $\tilde{g}_{ab}(\beta;\epsilon)$, with initial condition

$$\tilde{g}_{ab}(\beta=0;\epsilon) \equiv g_{ab}(\beta=0) + \epsilon K_{ab}(\beta=0). \tag{2.43}$$

These flows evolve with β yielding as "connecting vector" the bilinear form

$$K_{ab}(\beta) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}(\tilde{g}_{ab}(\beta;\epsilon) - g_{ab}(\beta)), \tag{2.44}$$

so we can define

$$\overline{K}_{ab} = \lim_{\beta\to\infty} K_{ab}(\beta). \tag{2.45}$$

The β evolution of $K(\beta)$ is found by linearizing (2.42). Formally we define it as

$$\frac{\partial}{\partial\beta} K_{ab}(\beta) = \frac{2}{3} g_{ab}(\beta)(\frac{1}{2} < R(\beta) K^c_c(\beta) >_\beta - \frac{1}{2} < R(\beta) >_\beta < K^c_c(\beta) >_\beta -$$
$$< R_{ab}(\beta) K^{ab}(\beta) >_\beta) + \frac{2}{3} < R(\beta) >_\beta K_{ab}(\beta) - \Delta_\beta K_{ab}(\beta) - L_Y g_{ab}(\beta), \tag{2.46}$$

with $K_{ab}(0) = K_{ab}$, and where Δ_β denotes the DeRham-Lichnerowicz Laplacian associated with $g(\beta)$, $\Delta_\beta K_{ab}(\beta) = -\nabla^c\nabla_c K_{ab} + R_{ac}K^c_b + R_{bc}K^c_a - R^d_{abc}K^c_d$, and L_Y is the Lie derivative along the vector field Y.

It can be shown that the flow $K(\beta)$, solution of (2.46), is such that $\frac{\partial}{\partial\beta}(< K^a_a(\beta) >_\beta) = 0$, i.e., the average over $(\mathcal{M}, g(\beta))$ of the trace of K is constant during the deformation.

Also, as (2.46) is the formal linearization of (2.42), we have $\lim_{\beta\to\infty} K_{ab}(\beta) = \frac{1}{3} < K^a_a >_o \overline{g}_{ab}$; $< \cdots >_o$ stands for the space average of the original physical quantity. Thus, the flow $K(\beta)$ deforms the given K by eliminating its shear: $K_{ab} - \frac{1}{3}K^c_c g_{ab}$, and replacing the original (position-dependent) rate of volume expansion K^a_a with its average value.

For each β the space average of matter density $< \rho(\beta) >_\beta$ can be defined analogously as the average of scalar curvature, namely,

$$< \rho(\beta) >_\beta = (V(\beta))^{-1} \int_{\mathcal{M}(\beta)} \rho(\beta) dV(\beta) \tag{2.47}$$

so then $\bar{\rho} = \lim_{\beta\to\infty} < \rho(\beta) >_\beta$ is the matter density in the limiting FLRW model.

The smoothing flow of regular initial data sets has to be such that for each value of β the four constraints of the Einstein equations:

$$R(\beta) - K_{ab}(\beta)K^{ab}(\beta) + (K^a_a(\beta))^2 = 2\rho(\beta), \tag{2.48}$$

$$\nabla_a K^{ab}(\beta) - \nabla^b K^a_a(\beta) = J^b(\beta), \tag{2.49}$$

have to be satisfied; where $\rho(\beta), J(\beta)$ are the mass and momentum density, respectively (of the external sources as described by a system of observers instantaneously at rest on $\mathcal{M}$), referred to the β dependent measure associated with $g(\beta)$.

For $\beta = 0$, (2.48) and (2.49) hold true, since $\rho(\beta = 0) = \rho$ and $J(\beta = 0) = J$ are the physical densities of sources of a given gravitational configuration $(\mathcal{M}, g, K)$. The averaging flows $\rho(\beta)$ and $J(\beta)$ cannot be defined independently once g and K are given and deformed, according to (2.42) and (2.46), in order for the constraints (2.48) and (2.49) to remain valid. In other words, to properly average the sources one has to take into account the backreaction of the geometry, determined by the constraints.

The constraints are thus interpreted as actually defining $\rho(\beta)$ and $J(\beta)$. Indeed, then $\overline{\rho} \equiv \lim_{\beta\to\infty} < \rho(\beta) >_\beta$ and $\lim_{\beta\to\infty} < J(\beta) >_\beta \equiv \overline{J} = 0$ (from (2.48) and (2.49) and the properties of the Ricci-Hamilton flow). One can show explicitly that

$$\overline{\rho} = [< \rho >_o + \frac{1}{2} < (K_{ab} - \frac{1}{3}K^c_c g_{ab})(K^{ab} - \frac{1}{3}K^c_c g^{ab}) >_o + \frac{1}{2}\overline{R}(\eta + \sigma^2)]/(1 + \sigma^2), \quad (2.50)$$

where $\sigma \equiv (< (K^a_a)^2 >_o - < (K^a_a) >^2_o)/ < K^a_a >_o$, $\eta \equiv (\overline{R}- < R >_o)/\overline{R}$, i.e., σ is the standard deviation describing the fluctuations of the original (position dependent) value of K^a_a with respect to its average (conserved) value $< K^a_a >_o$; η, $0 \leq \eta < 1$, denotes the relative function of the physical scalar curvature with respect to the averaged one $\overline{R} = \lim_{\beta\to\infty} < R(\beta) >_\beta$. $(K_{ab} - \frac{1}{3}K^c{}_c g_{ab}$ is the traceless part of K_{ab} on the initial manifold $\mathcal{M}(0)$).

Now, we can build an effective stress tensor modeling the dynamical effects of deviations from a spatially homogeneous geometry, (also those which are too big to be handled by perturbation techniques), that have been smoothed-out.

If $(\mathcal{M}, g_t)$ defines a normal geodesic slicing of $({}^{(4)}V, {}^{(4)}g)$ (for sufficiently small t) the stress tensor enters only into the evolution part of Einstein's equations,

$$\frac{\partial}{\partial t}K_{ab} = R_{ab} + K^c_c K_{ab} - 2K_{ac}K^c_b - (T_{ab} - \frac{1}{2}T^c_c g_{ab}) - \frac{1}{2}\rho g_{ab}. \quad (2.51)$$

The smoothing flow $T(\beta)$ of the spatial stress tensor is defined by requiring that for each t for which the evolution of the data $(g(\beta), K(\beta))$ is defined, the flows $(g_t(\beta), K_t(\beta))$ (resulting from the evolution equations) are Ricci-Hamilton flows, with initial conditions $g_t(0) = g_t$, $K_t(0) = K_t$, respectively.

The physical meaning of the presented results enables us to state precisely what is meant by the requirement that the original physical model universe and its FLRW smoothed-out ideal should behave as close as possible under their own gravitation. Namely, for $\beta \to \infty$, the volume $V(S^3, \overline{g}_t) = V(\mathcal{M}, g_t)$; this shows how the dynamics of the closed FLRW model is related to the dynamics of the original space-time. The fact that $V(\mathcal{M}, g_t) = V(\mathcal{M}, g_t(\beta))$ implies that $\frac{\partial}{\partial t}(< K^a_a >_o) = \frac{\partial}{\partial t}(< K^a_a(\beta) >_\beta)$ along the flow $(g_t(\beta), K_t(\beta))$.

The smoothed-out pressure $\overline{p} \equiv \lim_{\beta\to\infty} < p(\beta) >_\beta$ (in the final FLRW model on the surface of homogeneity $t = 0$, $\overline{p}_t$ (as well as $\overline{\rho}_t$ if the equation of state is known) can be determined by the evolution equation) is shown to be (taking into account (2.51))

$$\begin{aligned}\overline{p} &= \frac{1}{3} < T^a_a >_o + \frac{2}{3} < (K_{ab} - \frac{1}{3}K^c_c g_{ab})(K^{ab} - \frac{1}{3}K^c_c g^{ab}) >_o - \frac{4}{9}\sigma^2 < K^a_a >^2_o \\ &\quad + \frac{1}{3}(< \rho >_o - \overline{\rho}),\end{aligned} \tag{2.52}$$

with $\overline{\rho}$ given by (2.50). We see from (2.50) and (2.52) that $\overline{\rho} > 0$ and $|\overline{p}| \leq \overline{\rho}$, i.e. the dominant energy condition is satisfied.

In case when $\sigma^2 \approx 0$ (homogeneous expansion) and $\eta \ll 1$ (fluctuations of the physical curvature with respect to FLRW background curvature small, on average) the closed FLRW universe is the proper model, only if we add to the physical sources, $< \rho >_o$ and $< T^a_a >_o$, the term: $< (K_{ab} - \frac{1}{3}K^c_c g_{ab})(K^{ab} - \frac{1}{3}K^c_c g^{ab}) >_o$, taking this way into account the contribution of cosmological gravitational radiation. This term can influence the dynamics of the universe, and there is no evidence by now that the relative magnitude of this term with respect to $< \rho >_o$ is $\ll 1$.

Let us note that the smoothed stress tensor is defined by requiring that the smoothing commutes with the Einstein evolution, i.e., at each stage of this smoothing we, in fact, appeal to the standard form of Einstein's equations, though the effect of smoothing on those equations was being investigated. Therefore in this scheme one cannot say anything about what the effect of smoothing is on the form of the equations.

Let us also stress that this smoothing program has in general a few unresolved issues, like e.g. dependence on the spacelike slice chosen, making an identification between hypersurfaces of the original space-time and those of the smoothed space-time[9]. Secondly, the issue is how to define the smoothed stress tensor T_{ab} if one uses a more complicated form for $T_{\mu\nu}$ than a perfect fluid.

One possible and interesting application of this averaging procedure could be to examine the conjecture that all cosmological models with S^3 spatial topology have a time of maximum expansion.

[9]This issue appears in a different light in view of the approach in section 3.6.

Hemmerich [134] criticized this approach by raising, first of all, the slice dependence problem, namely, that the space-time metric in the limit $\beta \to \infty$ is non-unique as it depends not only on $g_{ab}(\infty)$, but also on the foliation, i.e. the lapse and shift functions in the ADM formalism. In particular, the limiting 4−metric may not be FLRW one at all. It was shown in [134] that the lapse and shift can be limited, so that $g^{(4)}_{\mu\nu}(\infty)$ is indeed the FLRW geometry.

The second objection was that the equations of Carfora–Marzuoli approach, among them (2.50), may be obtained directly from ADM formalism without employing the Ricci-Hamilton flow (2.42). To answer this, we can point out that equation (2.42) provides not only a relation between the average values of scalars on the initial manifold $\mathcal{M}(0)$ and the smoothed-out one $\mathcal{M}(\infty)$, but also a mapping of points of $\mathcal{M}(0)$ into points of $\mathcal{M}(\infty)$ which is lacking in other approaches. Moreover, with (2.42) one can consider any intermediate scale smoothing, from $\beta = 0$ to $\beta = \beta_o < \infty$, in addition to global smoothing from $\beta = 0$ to ∞.

The only real limitation of this scheme seems to be the assumption that all $\mathcal{M}(\beta)$ are closed with their Ricci tensors positive definite.

2.5 Specific exact solutions

An alternative in a situation when there are spherically symmetric inhomogeneities in the lumpy universe model is to use some exact solutions, e.g. in the form of Swiss Cheese models. These are made by combining matched sections of the FLRW expanding universe with spherically symmetric (Schwarzschild) vacuum or Lemaître-Bondi-Tolman [167, 33, 248] dust solutions[10], to yield an exact inhomogeneous universe model representing growth of inhomogeneities with a spatially homogeneous and isotropic background.

This kind of treatment, in a certain sense, achieves the goals of averaging without averaging *per se*. It is in fact a conceptually different approach to the fitting problem.

The first model is due to Bonnor [34]. He examined the Einstein-Strauss vacuole as representing a bound cluster of galaxies embedded in a standard pressure free cosmological

[10]This solution, though first proposed by Lemaître, is often called the Tolman-Bondi or just Tolman model.

model, using the Darmois junction conditions. The key point of his analysis is that the average energy density of the whole cluster is not the same as the density of the smoothed-out model. Therefore the field equations, with the energy density being the average density over a cluster of galaxies, are not satisfied.

The case when the lumpy universe model is an exact Tolman spherically symmetric dust universe was considered by Hellaby [132]. The approach taken was that of volume matching. The idea is that the behaviour of a certain region in space can be found to yield a particular FLRW model. One can then compare the density and equation of state of the resulting FLRW model with the matter content of the original volume. The results do not need to correspond to the mean density and pressure of the region derived by a simple averaging procedure. And if they do not, then we cannot count on this averaging of data to provide the parameters of a suitable homogeneous equivalent of our universe. Hellaby showed that there is a family of parabolic Tolman models in which the volume-averaged behaviour of the energy density and pressure is identical to that of a $k = 0$ pressure free FLRW universe (however local matching over infinitesimal volumes does not satisfy this condition). But in hyperbolic and elliptic cases, the effective (macroscopic) equation of state is not the same as the average of the Tolman values, e.g. although the pressure is zero in the inhomogeneous Tolman models the effective pressure in the FLRW (averaged) model is non-zero.

Recently, Moffat and Tatarski [193] studied a local void embedded in the globally FLRW model and discussed observational properties of such a model. The inhomogeneity was described using the Lemaître-Bondi-Tolman solution (the same as in [132]) with the spherically symmetric matter distribution taken to be dust, based on the faint galaxies number counts. The model has a property of being very similar to the FLRW one at the beginning of the expansion but becomes observationally different at later times. The authors studied its effects on the cosmological time scale, the measurement of the Hubble constant and the redshift-luminosity distance relation, which were shown to be fully compatible with cosmological observations. However, if we happened to live in such a void and insisted on interpreting cosmological observations through the FLRW model, then e.g. the Hubble constant measurement could give results depending on the separation of the source and the observer, and quasars could be younger than we think and also less distant (consequently less energetic).

Worth mentioning here, for its own sake, is Lindquist and Wheeler approach [174] where the idea of Schwarzschild–cell method was elaborated. (The method is similar in spirit to that of Wigner and Seitz in the solid state problem.) A number of mass concentrations is considered, such that the zone of influence of each can be approximated by a sphere. Inside each cell the actual gravitational potential is replaced by the Schwarzschild expression. Its important feature is that its normal derivative at the boundary of each lattice cell is non-zero and, moreover, does not go to zero at finite distance. Due to this fact, the mass concentrations on either side of the cell boundary accelerate towards that boundary, at such a rate as to nullify the discontinuity in matching of the normal derivative of the gravitational potentials that would otherwise occur. This feature expresses the equation of motion of the mass at the center of a cell as a dynamic condition on the boundary of the cell. When applied to the problem of the expanding universe, this idea enables one to derive the whole of the dynamics of the expansion and subsequent contraction from the elementary static Schwarzschild solution.

2.6 The Green Function approach

This is a perturbative approach which does not rely on any averaging procedure.

With the aim of studying the effects of a given matter distribution on the metric, and hence on the radiation, Jacobs and colleagues put forward a new scheme of determining the realistic metric of our universe [144, 143].

The idea is of solving the field equations through the use of scalar harmonics as spatial basis functions, while avoiding the use of any averaging procedure for the metric perturbations. Small metric perturbations are assumed (again, this does not restrict the size of perturbations to the matter variables), and the global expansion rate is that of FLRW model. Also assumed is the matter distribution and its evolution (known from observations and/or theory), but the results do not assume a particular model for the formation of structure in the matter distribution, and are valid everywhere in our universe outside of strong field regions.

In the presence of inhomogeneities we have

$$ds^2 = a^2(\eta)[\gamma_{\mu\nu}(\vec{x}) + h_{\mu\nu}(\eta, \vec{x})]dx^\mu dx^\nu, \tag{2.53}$$

where $h_{\mu\nu}$ describes the metric perturbations. It is assumed that $h_{\mu\nu} \equiv \mathcal{O}(\epsilon^2) \ll \gamma_{\mu\nu} \equiv \mathcal{O}(1)$ (background terms are of order 1), then $\Delta_\delta h_{\mu\nu} \simeq \mathcal{O}(\epsilon^2/\kappa)$.[11] Also, $\epsilon^2 \ll 1$ and $\epsilon^2 \ll \kappa$. The latter means that the matter inhomogeneities move non-relativistically and the effective stress-energy of metric perturbations is small.

Since the background is homogeneous and isotropic one can perform a separation of space and time dependencies in the field equations, enabling perturbations to be written, not as functions of $h_{\mu\nu}$, but as harmonic decomposition. The spatial dependence of perturbations is then expanded as eigenfunctions (normal modes) of the covariant Laplacian ${}^{(3)}\nabla^2$ on the 3-dimensional static background γ_{ij}. The field equations are reduced this way to the equations for the time dependent amplitudes of the modes.

Only scalar harmonics (scalar modes Q) are considered, in terms of solutions of

$${}^{(3)}\nabla^2 Q(\vec{x}, \vec{q}) = -q^2 Q(\vec{x}, \vec{q}) \tag{2.54}$$

and obviously, $Q \simeq \mathcal{O}(1)$ and $\nabla_i Q = \mathcal{O}(q) \simeq \mathcal{O}(\kappa^{-1})$.

The metric perturbations (h_{oo}, h_{oi}, h_{ij}) are then expanded in terms of scalar harmonics Q [18]. The longitudinal gauge is assumed and the Einstein tensor is written, including terms linear in $h_{\mu\nu}$ and its derivatives. Non-linear terms of $\mathcal{O}(\epsilon^4)$, $\mathcal{O}(\epsilon^4/\kappa)$, $\mathcal{O}(\epsilon^4/\kappa^2)$, or smaller are neglected. With this, we retain non-linear interactions of energy density of perturbations and their backreaction on the FLRW component. The stress-energy tensor is constructed in the usual way, taking perfect fluid as the background model, and perturbations (scalar) to the energy density μ, pressure p, and velocity v_i. The components of $T_{\mu\nu}$ are then written to the first order in the velocity.

By exploiting the harmonic decomposition of the field equations one can solve them by taking their spatial projections against different scalar modes. The important result is

[11] The parameters ϵ and κ are the same as already described in Futamase's approach.

then

$$ds^2 = a^2[-(1+2\phi)d\eta^2 + (1-2\phi)\gamma_{ij}dx^i dx^j], \tag{2.55}$$

where $\phi(\eta,\vec{x}) = -\frac{1}{2}h_{oo} = -\int d\mu(\vec{q})Q(\vec{x},\vec{q})H(\eta,\vec{q}) + \mathcal{O}(\epsilon^4)$, and $H \simeq \mathcal{O}(\epsilon^2)$ is the amplitude ($d\mu$ is the measure associated with the eigenvalue spectrum). $\phi(x^\mu)$ is the effective quasi-Newtonian potential of the inhomogeneities, characterizing metric perturbations.

In terms of the matter variables the equation for ϕ can be obtained from the following equation for H:

$$3(\frac{a'}{a})H' + (q^2 + 8\pi a^2\mu - 6k)H = 4\pi a^2\mu\Delta, \tag{2.56}$$

where $\Delta(\eta,\vec{q})$ is a suitable density fluctuation variable describing perturbations to the energy density μ, and *a priori* $|\Delta| > 1$; as usual $' = \frac{d}{d\eta}$.

To this level of approximation the Friedmann equation holds and there are two additional equations relating the metric perturbations to the pressure and velocity perturbations.

The estimation of the orders of magnitude of matter variables perturbations allows us to conclude that in any allowed regime (linear $\epsilon/\kappa \ll 1$, non-linear $\epsilon/\kappa \gg 1$) the pressure and velocity perturbations are much weaker than the density fluctuations. In other words, the metric perturbation $H(\eta,\vec{q})$ are determined primarily by $\Delta(\eta,\vec{q})$, i.e., hydrodynamically the density fluctuations can be treated as the source.

To any order of magnitude arguments we have to consider effects on the scale factor a, since it makes an implicit contribution. On physical grounds:

$$a(\eta) = a_{FLRW}[1 + \mathcal{O}(<\epsilon^4/\kappa^2>)], \tag{2.57}$$

and clearly, $\mathcal{O}(<\epsilon^4/\kappa^2>) \ll \epsilon^4/\kappa^2 \ll 1$, so using the background scale factor does not alter the arguments about the matter variables perturbations.

Solving equation (2.56) for $H(\eta,\vec{q})$ we can obtain the pseudo-Newtonian potential:

$$\begin{aligned}\phi(\eta,\vec{q}) \;\; = \;\; & \int dV(\vec{y})G(\eta_o,\eta,\vec{x},\vec{y})\phi(\eta_o,\vec{y}) - \frac{4\pi}{3}\int_{\eta_o}^{\eta} du\frac{a^3\mu}{a'}\int dV(\vec{y}) \\ & G(u,\eta,\vec{x},\vec{y})\Delta(u,\vec{y}) + \mathcal{O}(\epsilon^4),\end{aligned} \tag{2.58}$$

where dV is a coordinate volume element, and $G(u, \eta, \vec{x}, \vec{y})$ is a Green function for metric perturbations due to the scalar density fluctuations in a FLRW background. Formula (2.58) offers a relativistically correct way of calculating the metric perturbations, taking into account the cosmological expansion, non-linear density evolution and also, e.g. deviations from the thin-lens approximation.

In case of flat spatial sections ($k = 0$) it can be proved that

$$G_{k=0}(u, \eta, \vec{x}, \vec{y}) = \frac{a(u)}{a(\eta)} \frac{1}{[4\pi C(u, \eta)]^{\frac{3}{2}}} \exp[-\frac{|\vec{y} - \vec{x}|^2}{4C(u, \eta)}]. \tag{2.59}$$

The latter gives an interesting analogue with the diffusion [143].

The Green function expression for the potential can be reduced to a Newtonian form $\phi_{Newt} \simeq -\int dV(\frac{3}{8\pi}(\frac{a'}{a})^2 \frac{\Delta}{|\vec{y}-\vec{x}|}) \simeq -\int a^3 dV \frac{\mu\Delta}{a|\vec{y}-\vec{x}|}$ under appropriate conditions. But more interesting are situations where the time evolution of density fluctuations makes a significant contribution to the metric, e.g. post-Newtonian ones.

3 The Renormalization Group approach in Gravitation

3.1 Introduction

Many phenomena in Nature are so complicated that they do not succumb to reductionist approaches. For systems with many (infinitely many) degrees of freedom new types of collective behaviour emerge and their large-scale behaviour cannot be predicted from their microscopic origin. The phenomena we have in mind are complex precisely because they contain events and information over a wide range of length- and time-scales [256].

Luckily, physics is accustomed to ignoring inconvenient details and making use of large simplifications to get to the heart of issues. For example, in the standard models of magnetism many details of multi-spin and long range interactions, magneto-elastic coupling are omitted, as a rule, and yet the theory captures much of the essence of magnetic phenomena. Usually, a successful theory enables us to isolate some limited range of length-scales, or select a not too big set of variables, to render the problem tractable and at the same time preserve its essence. In many circumstances, fortunately, it is not necessary to resolve the details associated with each scale since generally phenomena at each scale can be treated independently. For example, in hydrodynamics there is no need to specify the motion of each water molecule in order to describe waves as a disturbance of a continuous fluid. However, in the complex phenomena, where each length-scale's contribution is of equal importance, one would need to take into account the entire spectrum of length scales, dealing with fluctuations of practically any wavelengths and consequently many coupled degrees of freedom. Such problems are thus generally intractable. Examples of this class of phenomena are critical phenomena, turbulent flow, the internal structure of elementary particles and confinement in QCD, to name a few.

There exist however a general class of theories (and methods) known as the Renormalization Group (RG) approach, which enabled to make progress in understanding the dynamics of complex phenomena. The value and significance of RG ideas, that have pervaded much of today's Statistical Mechanics and Quantum Field Theory, should not be underestimated since not only is the RG approach a method (unlike, e.g. High-Temperature Series Expansion method to calculate the values of critical exponents), but also a theory with essential physical ideas behind which can explain phenomena like scaling and universality, observed in various facets in Nature. From an even more open point of view it is also the RG *philosophy*, interpreted broadly to include various kinds of "multi-length-scale" and "coarse-graining" arguments.

What is the RG then? Usually, when we speak of a group we are thinking of symmetry operations, i.e. transformations that leave the physics invariant. In particular, this means that the RG procedure (whatever it means at the moment) can be iterated. Actually, the RG should be properly called a semi-group, because the inverse of the transformation is not defined. The RG approach can be very loosely described as follows: our aim is to study some properties of a certain function H and we perform a change of variables transforming the initial problem into an identical one in terms of a new function, now H', such that $H' = \mathcal{R}H$. The transformation $\mathcal{R}$ has to be chosen in such a way that after a few iterations (or at least in the limit $n \to \infty$) $H^{(n)} = \mathcal{R}^{(n)}H$ becomes tractable by some other techniques. In the successive changes of coordinates there is some loss of information, due to the fact that the changes of coordinates will not be one to one or everywhere defined - as a rule, since otherwise the problem could not really become simpler in the new coordinates.

Generally, the subject of the RG is then the modification of the fundamental laws of physics with the change of the observational length scale, and one may probe the dependence of the effective (RG-improved) couplings on the characteristic length through the RG flow equation(s). The otherwise complicated flow pattern becomes particularly simple in the vicinity of the fixed points where the linearized RG flow and scaling holds.

3.2 A short history of RG

The RG equations were first introduced into particle physics. It was observed that the conformal invariance, a space-time symmetry of the classical electro-magnetic action[1], can be violated by quantum effects in the energy regime $E \gg 0.5 MeV$, where the electron mass can be ignored. In Gell-Mann and Low's treatment of the short distance behaviour of QED, the theory is not scale invariant when the electric charge was renormalized at very large distances. In the seminal papers of [241] and [112], it was shown that quantum effects induce a scale dependence in the electro-magnetic charge and that the derivative of the electro-magnetic coupling with respect to the scale is an analytic function of the coupling itself, the so-called β function.

Further, there are physical amplitudes in QFT, which depend on the couplings and also on points in space, labeled by the coordinates x_i (e.g. Cartesian in flat space). In [48] it was shown that the variation in the couplings under a change of scale (at which they are defined) can be always compensated for by a rescaling of the coordinates x_i, so that the vacuum amplitudes remain invariant. This results in the RG equations - an inhomogeneous partial differential equation for the amplitudes, further extended in [253, 135] to a homogeneous equation. In QFT the RG is an *exact continuous symmetry group* (of a solution formulated in its natural variables).

The issue of course is that QED or any renormalizable field theory is plagued by "infinities" which can be renormalized. This requires an introduction of "bare" couplings $g_{bare}(g, \epsilon)$, which are analytic functions of the renormalized couplings, g_i, and a regularization parameter or parameters, ϵ, e.g. for a cut-off, Λ, $\epsilon = \kappa/\Lambda$, where κ is a renormalization point and for dimensional regularization $\epsilon = 4 - d$, where d is the dimension of space or space-time. One can prove then that to any given order in perturbation theory it is possible to choose the couplings in a cut-off dependent way so as to make physics at momenta much smaller than the cut-off independent of it. As eventually the cut-off is sent to infinity, the above means physics at any finite momentum. The change in the cut-off accompanied by a suitable change in the couplings is an invariance of the theory.

These transformations form a group, with the law that changing the cut-off by a factor

[1] This symmetry is only angles preserving, not lengths.

s_1 and then by s_2, should be equivalent to a change by a factor $s_1 s_2$. Writing

$$s = e^{-t} \tag{3.1}$$

i.e., $\Lambda(t) = \Lambda_o e^{-t}$ (Λ_o is some fixed number), the above composition rule just means that when two transformations are implemented in a row, the parameters t add. The β-function is defined as

$$\beta(g) = \frac{dg}{dt}, \tag{3.2}$$

where g stands for the coupling(s).[2] In the case of Yang-Mills theory we have

$$\frac{dg}{dt} = cg^3 + \text{higher orders}, \quad c > 0. \tag{3.3}$$

Integrating this from $t = 0$ to $t = t$ (i.e. from Λ_o to $\Lambda_o e^{-t}$) we find

$$g^2(t) = \frac{g^2(0)}{1 - 2g^2(0)ct}. \tag{3.4}$$

What this means is that as we send the cut-off to infinity ($t \to -\infty$) we have to reduce the coupling to zero logarithmically $g^2(t) \simeq 1/|t|$.

In statistical physics the RG equation was raised to central importance through the work of Kadanoff [147] and Wilson [257] who transferred RG philosophy from the relativistic quantum field to the analysis of phase transition phenomena in spin lattice systems. There, one contemplates changing the cut-off (and the couplings) even in a problem where nature provides a natural cut-off, such as the inverse lattice spacing, $\Lambda \simeq 1/a$, and there are no UV infinities. In this view-point, the cut-off is not to be viewed as an artifact to be sent to infinity, but as the dividing line between the modes we are interested in and the ones we are not interested in. We may change the cut-off and the couplings without affecting the slow mode physics (for a clear exposition, see e.g. the introduction to [228]).

This rationale lies behind Kadanoff's *spin blocking* (or decimation), with simultaneous finite renormalization of a coupling constant. Blocking is performed by averaging over the constituents in such a way that the number of degrees of freedom diminishes, simplifying the system and preserving, at the same time, its long-range properties. This in fact refers

[2] There are different conventions - in condensed matter physics increasing t decreases Λ, opposite to the field theory convention.

to the creation of another theoretical model for the same physical phenomena thus RG is here an approximate, discrete *semi group.* To get, after averaging, a system formally similar to the initial one, one neglects some terms, in the new effective Hamiltonian, that are unimportant for infra-red properties. The new model at the end formally differs by some elementary scale and the coupling constant value. The blocking operation can be repeated.

The RG method turned out to be powerful in establishing connections between physics at different scale levels, understanding universality of various types of critical behaviour and finding a stable infra-red fixed point. In QFT this same approach was used to discover a stable ultra-violet fixed point by suppression of irrelevant low energy degrees of freedom. Further, the RG algorithms were assimilated in the field of polymers, transfer phenomena and percolation. The physics of renormalization transformation in turbulence is related to a change of ultra-violet cut-off in the wave number variable.

More recently the idea inspired by O'Connor and Stephens [201] that the RG equations, from a geometric point of view, should be viewed in terms of a coordinate transformation on the space of couplings, was further pursued by Dolan [71] who showed that the RG equations for vacuum amplitudes can be interpreted as a Lie derivative on the space of couplings.

3.3 Real-space RG and critical phenomena

Critical points (second-order phase transitions) occur in liquid-gas transitions, ferromagnetic transitions, binary alloys transitions, etc. There are close analogies relating all these critical points, which is one of the fascinations of the subject.

The critical temperature (Curie temperature) T_c of a ferromagnet marks the onset of spontaneous magnetization in the absence of an external field. Above the critical temperature T_c, it is zero (the ferromagnet is in a paramagnetic state), below (near) T_c, varies as $(T_c - T)^\beta$. The exponent β is an example of a critical exponent. Theory of critical phenomena should be able to predict the values of the exponents. Other critical exponents characterize other power laws near the critical point. The critical point of the ferromagnet is at zero applied magnetic field $\vec{H}$ and at $T = T_c$. The derivative of the magnetization $\vec{M}$ diverge at T_c. The phase transition has no associated latent heat and can be described as a critical phase transition.

The idea of possible application of RG method in the phase transition context in Statistical Mechanics is connected with the homogeneity of thermodynamic functions near the critical temperature, which was first described in [254], and with the behaviour of correlations functions in the critical region ($T = T_c$), studied by Kadanoff [147]. The important idea is that only long wave length fluctuations in the order parameter, are responsible for the critical singularities. Near to a critical point the correlation length diverges and because the degrees of freedom are correlated over macroscopic distances, the system looks very similar under a transformation that divides the lattice into smaller blocks and replaces the degrees of freedom inside each of them by a block-average. In fact, Kadanoff has brilliantly hypothesized that if the block lattice is considered, then at the critical point the block-lattice Hamiltonian may be reduced to the initial lattice Hamiltonian by scaling transformations. Intuitively, the effects of this kind of averaging are to cut down the correlation length by the scale factor b, where b is the linear dimension of the block and the interactions between the new variables will be different. If the blocks are small then we expect that the interactions between the block variables (spins for a ferromagnet) will be local if the interactions we started with were local. Notice, that the correlation length decreases (by a factor of b) since it is measured in lattice units, which change in our transformation from being the distance between the original lattice points to that between blocks now.

The RG theory of critical phenomena owes its great success to the remarkable works of Wilson. His formulation provided a systematic way of implementing the integration over a finite fraction of degrees of freedom in a system near its critical point, and quantifying the effect on the remaining variables, providing in this way all the mathematical infrastructure explaining scaling and universality. More precisely, the effect of the long wave length fluctuations can be calculated by using self-similarity properties of a critical system under scale transformations, where scaling amounts to the integration over the short wave length components and determining the effect on the long wave components. When this effect is vanishing the scale transformation has a fixed point, which in fact determines the critical singularities that are universal, i.e. insensitive to the details of the molecular interaction. In [259, 258] it was shown how to embody at the microscopic level, Kadanoff's hypothesis of universality into precise differential equations and also how to obtain the explicit block lattice Hamiltonians from initial Ginzburg-Landau Wilson Hamiltonian. He has also found the recursion relations and utilized the scaling transformation and the fixed point method to solve them.

In the real space approach the self-similarity is realized for a microscopic Hamiltonian, and scaling now means that the variables of the system are combined to new similar variables. For lattice systems (Ising model) this seems possible. Universality emerges through the properties of the (self similar or) fixed point Hamiltonian. It is attractive in the irrelevant (or non-singular) directions and a critical fixed point dictates the singular properties for a whole surface of critical Hamiltonians to be the same.

Generally, some problems appear with the above procedure in the applications, but first let us remark that according to what was said above, RG transformation is a mapping from Hamiltonians or actions defined in a certain phase space, to actions in the same space. More generally, RG map is defined as a rule (deterministic or stochastic) for generating a configuration ω' of "block spins", given a configuration ω of "original spins". In mathematical language this is given by a probability kernel $T(\omega \to \omega')$, with which one can define a probability distribution $\mu'(\omega')$ of block spins from any given probability distribution $\mu(\omega)$ of original spins [94]

$$\mu'(\omega') = (\mu T)(\omega') \equiv \sum_{\omega} \mu(\omega) T(\omega \to \omega'). \tag{3.5}$$

RG map is therefore defined as a map from measures to measures, in fact in applications we assume it to be defined from Hamiltonians to Hamiltonians[3].

If we represent the initial action as a point in a coupling constant space, it will flow under RG transformation to another point in the same space, therefore a possibility of a *fixed point* of the group action exists, that is to say, the action function which reproduces itself after RG transformation. Geometrically speaking the fixed point does not flow under

[3] We will not enter into a discussion of any problems which this assumption can engender; interested reader can consult [94].

RG transformation. The sequence of theories specified by the couplings generated under successive application of RG is called the RG flow. The critical behaviour of the model can be gotten from the RG flows [101].

Now, the problems are also, (1) to give the rules relating the new variables to the old ones, and (2) to work out the properties of the relation between the Hamiltonians for the new and old variables. Technically, one has problems with determination of the properties of renormalization transformation. It is hoped that approximations, which fail at the critical point for the partition function, may be carried out successfully for a renormalization transformation. This crucially depends on the choice of the renormalization transformation. In fact, many aspects of the renormalization procedure are obscure.

In order to highlight the concepts involved, it is useful to consider a simple example, namely the Ising model. The Hamiltonian can be generally written as

$$\mathcal{H}(s) = \sum_{\alpha} K_\alpha s_\alpha, \tag{3.6}$$

where s_α stands for a spin function of type α, e.g.

$$\begin{aligned} s_h &= \sum_i s_i, \\ s_\epsilon &= \sum_{(ij)} s_i\, s_j, \end{aligned}$$

for the order parameter and for the energy, respectively, where s_i is the spin of site i (and (ij) nearest neighbour (NN) sites). K_α are the coupling constants, namely, K_h is the magnetic field, K_ϵ the NN coupling for the standard Ising model.

The RG transformation is defined as

$$e^{\mathcal{H}'(s')} = \sum_{\{s\}} P(s', s) e^{\mathcal{H}(s')}, \tag{3.7}$$

where $P(s', s)$ is the weight function satisfying

$$P(s', s) \geq 0, \quad \sum_{\{s'\}} P(s', s) = 1. \tag{3.8}$$

Equation (3.7) is the defining relation for the new Hamiltonian and $\mathcal{H}'(s')$ can be written again in the form (3.6), defining new coupling constants K'_α, such that the relation

$$K'_\alpha = K'_\alpha(K_\beta) \tag{3.9}$$

is equivalent with the formal definition (3.7). Many transformations are possible, in fact (3.8) guarantees only that the K'_α are real and the partition functions of $\mathcal{H}$ and $\mathcal{H}'$ are the same (of the same functional form).

The derivative matrix

$$T_{\alpha\beta} = \frac{\partial K'_\alpha}{\partial K_\beta} \tag{3.10}$$

yields directly the critical exponents.

In the case of an exactly solvable $1d$ Ising model with NN interactions, the averaging process is easily defined to consist of exact integration over every other spin in the chain. This procedure preserves the partition function and the interaction between the thinned-out degrees of freedom remains NN, and moreover is an analytic function of the original NN interactions. The integration procedure can be iterated therefore, and new NN couplings K'_α obtained in terms of the original coupling K_α by the same function as $K' = \mathcal{R}(K)$.

In the case of $2d$ Ising model with NN interactions one can integrate over spins, say, on every even site but this reduction of degrees of freedom induces other types of interactions between the remaining spins, in addition to renormalizing the original NN coupling. The new couplings are analytical functions of the original NN coupling but now, the integration cannot be iterated to yield closed form solutions since at each step longer and longer range interactions get generated.

Now, in the Wilson's formulation of the RG one considers from the outset, a general class of theories defined by the infinite set $\{K_\alpha\}$ of couplings corresponding to interactions of all possible ranges and Hamiltonian's complexity. An appropriate scale transformation $\mathcal{R}_b$, which integrates over a fraction of the original degrees of freedom, maps the theory into another one with the renormalized couplings: $K'_\alpha = \mathcal{R}_b(K_\alpha)$. For most models one cannot find out exactly this relationship and must resort to approximate numerical methods. When performing the real-space RG procedure one stipulates that the effective Hamiltonian $\mathcal{H}^{(n+1)}$ must take the same functional form as $\mathcal{H}^{(n)}$, so that the model is exactly the same at every stage, except for a change in the parameters in the effective Hamiltonian (this condition is in almost all cases impossible to meet exactly).

A useful RG transformation has to have a number of properties; we will mention just one, the correlation length decreases under $\mathcal{R}_b : \xi \to \xi' = \xi/b$. It is preserved under RG transformation if the starting Hamiltonian H[4] has $\xi = 0$ or $\xi = \infty$. Theories with $\xi = 0$ are trivial, e.g. $T = 0$ or $T = \infty$ limit of most statistical mechanics models.

Critical Hamiltonian $\mathcal{H}_c$, the one with the infinite correlation length, when acted on with $\mathcal{R}_b$ produces another critical Hamiltonian, since obviously $\xi' = \xi/b = \infty$. The set of critical points define a hypersurface in the infinite dimensional space $\{K_\alpha\}$. The RG flow on this surface can: a) meander randomly, b) go to some limit cycle or strange attractor or, c) converge to a fixed point $\mathcal{H}^*$. At the fixed points the renormalized couplings are exactly equal to the original couplings, $K'_\alpha = K_\alpha$, and the theory reproduces itself al all length scales. In the case of critical points only the long distance behaviour is reproduced. In studying the critical phenomena we are interested in RG transformation which has a critical fixed point $\mathcal{H}^* = \mathcal{R}_b(\mathcal{H}^*)$ with $\xi = \infty$. Each such fixed point has a basin of attraction, i.e. the set of $\mathcal{H}_c$ that converge to it under $\mathcal{R}_b$, which defines the universality class since the long distance behaviour of all theories corresponding to these $\mathcal{H}_c$ is governed by the same fixed point. What this means is that the critical (static) exponents are rather insensitive to the details of the system and for a continuous phase transition they depend only on: the dimensionality of the system d, its symmetry, the dimensionality of the order parameters, whether the forces are of short or long range. The values of critical exponents can be obtained from the eigenvalue equation for the linearized RG transformation in the vicinity of the fixed point.

We would like now to give an example of the scaling hypothesis in the case of a ferromagnet, represented by Ising model. It says that close to the critical point the singular part of the Gibbs free energy[5] is a generalized homogeneous function of its variables

$$G(\lambda^{a_t}t, \lambda^{a_h}h) = \lambda G(t,h), \tag{3.11}$$

where a_t, a_h are two parameters (scaling powers), λ is arbitrary and $t = T - T_c$. The scaling laws (algebraic equalities between the critical exponents) can also be derived in RG approach. We do not present the details here.

[4]We refer to $\mathcal{H}$ as the effective Hamiltonian. For the starting system, the temperature dependence of $\mathcal{H}$ is through $\mathcal{H} = \beta H$, $\beta \sim 1/T$, where H is independent of temperature.

[5]The same property holds for all the other thermodynamic potentials.

For a magnet near its critical point, there are two quantities which measure the deviations from the critical point: a dimension-less magnetic field h and $t = T - T_c/T_c$, and near the critical point $h, t \ll 1$. The statement of scaling is that the powers of the temperature deviation provide a characteristic scale of all physical quantities, e.g. the magnetic field always appears in the theory in the combination h/t^{Δ} (where Δ is the critical index for the magnetic field). Correspondingly, the magnetization appears in the combination m/t^{β}. The content of the scaling hypothesis (which after all should not be called a hypothesis any more) is then that the magnetization appears in the scaling form

$$m(h,t) = t^{\beta} m^{\star}(\frac{h}{t^{\Delta}}). \tag{3.12}$$

If h is of the same order of magnitude as t^{Δ}, then $m^{\star}$ is of order one and m is of order t^{β}. Statements like this can be verified experimentally.

More literature on the subject for interested reader can be found in [17, 72, 31, 123], see also [228].

Before closing this section we offer some comments on the renormalization of fluids. The fluid is believed to be in the same universality class as the Ising model. Conceptually the problem is whether a renormalization on a microscopic level is possible and whether a lattice structure is necessary for a successful renormalization. In [136] a map was constructed by which the fluid properties are associated with a (general) Ising model. In fact, in some respect the fluid is an advantage, as for the fluid the elimination of an infinitesimal fraction of the particles is achieved most easily. On the other hand, the continuous potentials are much more difficult to handle than the discrete interaction constants for a lattice.

3.4 Critical phenomena in Cosmology

In this chapter we review the recently discovered critical behaviour in the gravitational collapse situations.

We suspect that these are not the only critical phenomena possible on GR grounds. We will deal with the other possible scenarios in section 3.6 after we develop the necessary tools to tackle the problem. All of them borrow from the previously discussed concepts of RG.

In our opinion, there are at least a few facts–hints, pointing to the broadly understood criticality in the universe. Firstly the distribution of cosmic structures in the Universe, galaxies and clusters of galaxies, exhibits certain scaling properties [36, 146], namely, the two-point correlation function of galaxies, clusters and quasars has a power law[6] behaviour $r^{-\gamma}$, with $\gamma \sim 1.8$, up to present day scales of about $300\,Mpc$ (but with different correlation lengths for different cosmic structures). *This fact seems to be a clear hallmark of the phase transition underlying the origin of structure formation.*

Scaling behaviour is in general common to systems that obey non-linear dynamical equations with possibly a stochastic driving term, or more generally, even to chaotic systems and phenomena (see [214] for the excellent review on chaos in the Einstein equations). In this respect it is interesting to notice that spirals are a dominant type of pattern in spatio-temporal chaos in non-equilibrium systems [115]. A general explanation of the phenomenon is lacking, though many examples[7] show that pattern formation in spatially extended non-linear systems near the critical point, specially when chaos is involved, yields spirals. Depending on the value of parameters, the spiral pattern dominates or not and the mechanism by which the pattern conversion occurs seems to be through interaction with defects in the patterns. Their rôle in non-equilibrium patterns is by no means clear; the defect propagation and interaction can be crucial to understanding non-equilibrium patterns and chaos. We have mentioned this fact due to the spiral galaxies which abound in the universe; in our opinion this issue deserves further investigation.

No doubt, the universe is an extremely large dynamical system and moreover it is quite complex, since it contains events and information over a wide range of length- and time-scales.

Recently, a new approach has emerged, namely, *self-organized criticality* (SOC), that

[6] Power laws are essential for scale invariant phenomena since re-scaling of the variables, e.g. $x \to ax$, does not change the shape of the distribution e.g. $N(ax) = a^{-b}x^{-b}$.

[7] E.g. spirals in thermal convection near the critical point of a fluid, Rayleigh-Bénard convection near the thermodynamic critical point. In this case whether the spiral pattern dominates or not depends on the Prandtl number (related to the viscosity).

might be a mechanism leading to complexity[8]. It is pertinent to systems with many interacting degrees of freedom which operate presumably far from equilibrium. For them small increments in energy input can trigger an arbitrarily large *avalanches* (activity) with power law spatial and temporal distribution functions limited only by the size of the system, whereby they self-organize themselves into a critical state[9] [16].

3.4.1 Critical phenomena in GR gravitational collapse

Recently, Choptuik [60] (see also [59]) and Abrahams & Evans [1] discovered numerically new solutions of Einstein equations at the threshold of black hole formation, exhibiting non-linear dynamical behaviour suggestive of critical phenomena. Before going to the detailed examples we will describe some of their characteristics in general terms.

Critical phenomena become evident as variations in the properties of space-times across a parameter space of space-times. Suppose that a single parameter p_k describes each space-time. There is then a correspondence between parameter spaces $\{\mathcal{G}_k\}$ and the space-times $\{S_k[p_k]\}$. For each $\mathcal{G}_k$ a critical value $p_k^\star$ of the parameter p_k, separates the parameter space into a half-space $\mathcal{G}_k^+$ describing space-times that contain a black hole and the other that do not. The parameters p_k are associated, as can be seen, with the variations in the strength of the gravitational self-interaction. The critical behaviour occurs in space-times that are just on the "edge" of forming a black hole, i.e. when $|p_k - p_k^\star|$ is small. In particular, in $\mathcal{G}_k^+$, the black hole mass was found to fulfill a power-law dependence $|p_k - p_k^\star|^\beta$ with a critical exponent β. For p_k close to $p_k^\star$ each space-time develops a strong field region $\mathcal{R}$, where the gravitational field (and any coupled field) develops an oscillatory character, revealing the existence of (discrete) scaling relations due to which the successive oscillations are echoes of each other on progressively smaller spatial and temporal scales. Also in the case of [60] the universality was demonstrated. This refers to the fact that in the case of scalar

[8] Similar ideas built on chaos, scaling, etc. find their way in other branches of science as well, e.g. economics, biology, environmental sciences (see e.g. [213]), though mostly they are only implemented in simple numerical toy-models of sandpile type. This situation is similar to the one that regards fractals – not much is known about the *physics* of fractal dynamics.

[9] As an example, we can give pulsar glitches caused by changes in the speed of neutron star rotation (presumably due to an abrupt change in internal structure). When the changes in frequencies are translated to changes in rotational energy and the cumulative distribution of that energy plotted, it turns out to be a power law.

field collapse the shape of the fields in the critical space-time $S_k[p_k^*]$ in the strong field region, and the value of critical exponent for black hole mass (and the scaling relation) do not depend on which parameter space, $\mathcal{G}_k$, is examined. To put it differently, the critical behaviour is generic and independent of the details of the initial data.

3.4.1.1 Scalar field collapse

For massless scalar field (both minimally coupled $\zeta = 0$ and non-minimally coupled) the critical behaviour was found by computer simulation of the collapse of spherically-symmetric wave-packets of scalar field.

The equation of motion is

$$\phi_{;\mu}{}^{;\mu} = \zeta R \phi\,, \tag{3.13}$$

and the line element

$$ds^2 = -\alpha^2(r,t)dt^2 + a^2(r,t)dr^2 + r^2 d\Omega^2, \tag{3.14}$$

where α is the lapse function and a the radial metric function[10]. Lapse is fixed by the polar time slicing condition and the shift vector $\beta^r = 0$, fixing the spatial coordinate trajectories. This gauge generalizes Schwarzschild coordinates for dynamical space-times as $a^2 = (1 - 2m(r,t)/r)^{-1}$.

Choptuik introduces auxiliary variables $\Phi = \phi'$, $\Pi = a\dot{\phi}/\alpha$ and solves the following equations ($\zeta = 0$):

$$\dot{\Phi} = (\tfrac{\alpha}{a}\Pi)'$$

$$\dot{\Pi} = \tfrac{1}{r^2}(r^2\tfrac{\alpha}{a}\Phi)'$$

$$\tfrac{\alpha'}{\alpha} - \tfrac{a'}{a} + \tfrac{1-a^2}{r} = 0 \tag{3.15}$$

$$\tfrac{a'}{a} + \tfrac{a^2-1}{2r} - 2\pi r(\Phi^2 + \Pi^2) = 0, \tag{3.16}$$

where a dot stands for $\partial/\partial t$ and a prime for $\partial/\partial r$. The radial coordinate r (covariantly defined) measures proper surface area; time coordinate t has no relevant geometrical interpretation, except as $r \to \infty$ where it measures proper time. In fact, though critical

[10] $G = c = 1$ units are used.

phenomena were discovered via computations in r, t, they are better described in r, T, where T is proper time of an observer fixed at $r = 0$, namely, $T \equiv \int_0^t \alpha(0, \tilde{t}) d\tilde{t}$.

The finite difference code is based upon an adaptive-mesh-refinement algorithm which is very well able to resolve very fine spatial and temporal features.

The scalar field ϕ has an initial profile

$$\phi(r, 0) = \phi_o r^3 e^{-[(r-r_o)/\Delta]^q} , \tag{3.17}$$

and a one-parameter space of solutions is generated from this Cauchy data with a condition on Π that initially the scalar radiation is purely ingoing. The initial data are specified if also the slicing condition (3.15) is solved for α, and Hamiltonian constraint (3.16) for a. Once r_o, Δ and q are considered fixed then ϕ_o serves as a single parameter characterizing the sequence of solutions, and $\mathcal{G}_k = \phi_o$, for a particular k. The parameter ϕ_o is related to the strength of the field's self-interaction. For small ϕ_o, the wave-packet implodes and disperses to infinity, the scalar and gravitational fields decouple (dynamics is described by flat space-time solution of the spherically-symmetric wave equation); for large ϕ_o its implosion leads to a black hole formation. Black hole is detected by monitoring $\frac{2m}{r}$. For black hole space-times' $\frac{2m}{r} \rightarrow 1$ (for some specific $r = R_{BH}$), and the mass of black hole $m_{BH} = 2R_{BH}$ can be calculated. A critical value of the parameter $\phi_o^\star$ separates supercritical ($\phi_o > \phi_o^\star$) from subcritical ($\phi_o < \phi_o^\star$) solutions.

In regions of parameter space close to $\phi_o^\star$, Choptuik found that critical features in solutions (close to the critical point) tend to depend linearly on $\ln|\phi_o - \phi_o^\star|$ and therefore exponentially on the initial conditions, and that structures with increasingly finer spatial and temporal scales develop as $\phi_o \rightarrow \phi_o^\star$. In terms of two new variables $X = \sqrt{2\pi} r \Phi / a$ and $Y = \sqrt{2\pi} r \Pi / a$ (they are invariant with respect to rescalings of the length and time coordinates ($r \rightarrow \kappa r$ and $t \rightarrow \kappa t$) and hence to rescaling of the mass of the space-time) one can better describe the echoing and scaling behaviour. This is so because critical dynamics is most naturally expressed in terms of the variables that are form-invariant with respect to these rescalings, which on their own, express the absence of any intrinsic mass/length scale in the model. Obviously, the equations to solve and any solutions thereof are invariant under these rescalings.

In solutions close to the critical ones in the strong field region, the scalar field oscillates and the number of oscillations is proportional to $|\ln|\phi_o - \phi_o^\star||$. The conjecture therefore is that every critical solution contains an infinite number of echoes. Let us introduce the logarithmic spatial and temporal coordinates ρ and τ defined as follows

$$\begin{aligned}\rho &= \ln r \\ \tau &= \ln(T^\star - T).\end{aligned} \tag{3.18}$$

(The constant $T^\star$ is the finite accumulation time of the echoes in the precisely critical solution and can be determined in solutions close to the critical ones by fitting).

Choptuik finds that approximate scaling relations hold

$$\begin{aligned}X(\rho - \Delta, \tau - \Delta) &= X(\rho, \tau) \\ Y(\rho - \Delta, \tau - \Delta) &= Y(\rho, \tau).\end{aligned} \tag{3.19}$$

Critical dynamics is unique (up to trivial rescalings $r \to kr$, $t \to kt$) and invariant under a *discrete* scaling symmetry. The scale periodicity is similarly conjectured for all other form-invariant quantities, including ϕ, $r^2\phi^{,\mu}\phi_\mu$, $\frac{dm}{dr}$, m/r. Physically it means that for a precisely critical configuration, an infinite series of "echoes" is generated from the recurrence of strong-field evolution on ever decreasing spatio-temporal scales. The scalar field oscillations appear as echoes of one another on scales finer by a factor of $e^{-\Delta} \simeq 1/30$. Critical dynamics "accumulates" at some critical central proper time $T^\star$. In other words, if the radial profiles of X and Y are observed at some time T_1, with a small interval $\delta T_1 = T^\star - T_1$, and again at a second time T_2, with even smaller interval $\delta T_2 = e^{-\Delta}\delta T_1$ before $T^\star$, then a new detail will have appeared in the later profiles on a finer scale, but upon rescaling radially by a factor of e^Δ the new profiles are in fact identical to the earlier ones. Note that the scaling relations observed is always approximate because even in the precisely critical solution the initial oscillations near the outer edge of the strong field region will contain information on the initial data. With each echo this information is washed out and scaling relation holds tighter. Besides, any near-critical solution produces only a finite number of echoes as $T^\star$ is approached before it "decides" whether to form a black hole or not. The value of Δ was found to be universal $\Delta \simeq 3.4$. The profiles of $X(\rho,\tau)$ and $Y(\rho,\tau)$ were also found to be universal, i.e. independent of the family of initial data.

For solutions in the half-spaces $\mathcal{G}_k^+$ (for which $\phi_o \to \phi_o^\star$ from above) a power law has been found for the black hole masses

$$m_{BH} \simeq C|\phi_o - \phi_o^\star|^\beta, \tag{3.20}$$

with $\beta \simeq 0.37$ (and C family dependent constant). This power-law behaviour was also found to be universal, with possibly weak dependence, if any, on the coupling constant ζ. The controversial conjecture one can make based on the above is that a black hole first appears along any sequence at $p = p^\star$ with infinitesimal mass, i.e. the black hole transition point is generically massless. Since the associated oscillations in the scalar field occur on increasingly finer scales, both the "kinetic energy" of ϕ, $\phi^{,\mu}\phi_\mu$, and the scalar curvature of the space-time, get driven to infinite values at the critical event $r = 0$, $T = T^\star$. The precisely critical space-time is necessarily singular and most likely naked.

3.4.1.2 Axisymmetric gravitational wave collapse

The second example of critical phenomena in GR was demonstrated in [1] in the axial collapse of gravitational wave-packets. This model is source free, $T^{\mu\nu} = 0$, and less symmetric (one Killing vector) from the Choptuik's example. A dynamical degree of freedom of the gravitational field is necessarily involved here.

Abrahams and Evans compute axisymmetric, asymptotically flat vacuum space-times, using a $3+1$ formalism. They adopt the maximal time-slicing condition ($K_i^i = 0$; K_j^i is the extrinsic curvature) and the quasi-isotropic spatial gauge, which fix the coordinates. The line element has the form

$$ds^2 = -\alpha^2 dt^2 + \phi^4[e^{2\eta/3}(dr + \beta^r dt)^2 + r^2 e^{2\eta/3}(d\theta + \beta^\theta dt)^2 + e^{-4\eta/3} r^2 \sin^2\theta d\varphi^2], \tag{3.21}$$

where α is the lapse function, β^r and β^θ are shift vector components, ϕ is the conformal factor, and η is the even-parity "dynamical" metric function. Numerical solutions are computed to the following set of equations

$$\partial_t\hat{\lambda} = \mathcal{D}_\beta[\hat{\lambda}] - \phi^6(D^r D_r\alpha + 2D^\varphi D_\varphi\alpha) + \alpha\phi^6(R_r^r + 2R_\varphi^\varphi) + \hat{K}_\theta^r/r[r\partial_r\beta^\theta - \partial_\theta(\beta^r/r)],$$

$$\partial_t \hat{K}^{\varphi}_{\varphi} = \mathcal{D}_\beta[\hat{K}^{\varphi}_{\varphi}] - \phi^6 D^{\varphi} D_{\varphi}\alpha + \alpha\phi^6 R^{\varphi}_{\varphi},$$
$$\partial_t(\hat{K}^r_\theta/r) = \mathcal{D}_\beta[\hat{K}^r_\theta/r - \phi^6 D^r D_\theta \alpha/r + \alpha\phi^6 R^r_\theta/r + (2\hat{\lambda} - 3\hat{K}^{\varphi}_{\varphi})[\partial_\theta(\beta^r/r) - \alpha K^r_\theta/r],$$
$$\partial_t \eta = \beta^r \partial_r \eta + \beta^\theta \partial_\theta \eta + \partial_\theta \beta^\theta - \beta^\theta \cot\theta + \alpha\lambda,$$
$$\Delta^{(3)}_f \psi = -\frac{1}{4}\psi(\Delta^{(2)}_f \eta + \frac{1}{2}\psi^{-8} e^{-2\eta}\hat{K}^i_j \hat{K}^j_i),$$
$$\Delta^{(3)}_f(\alpha\psi) = -\frac{1}{4}\alpha\psi(\Delta^{(2)}_f \eta - \frac{7}{2}A^2 K^i_j K^j_i),$$
$$r\partial_r(\beta^r/r) - \partial_\theta \beta^\theta = \alpha(2\lambda - 3K^{\varphi}_{\varphi}),$$
$$r\partial_r \beta^\theta + \partial_\theta(\beta^r/r) = 2\alpha K^r_\theta/r,$$

where $K^i_j K^j_i = 2\lambda^2 - 6\lambda K^{\varphi}_{\varphi} + 6(K^{\varphi}_{\varphi})^2 + 2(K^r_\theta/r)^2$, and the transport operator is defined by $\mathcal{D}_\beta[u] = \frac{1}{r^2}\partial_r[r^2\beta^r_u] + \frac{1}{\sin\theta}\partial_\theta[\sin\theta\beta^\theta u]$; $\lambda = K^r_r + 2K^{\varphi}_{\varphi}$, $\hat{K}^i_j = \phi^6 K^i_j$, D_k is the spatial covariant derivative, R^i_j spatial Ricci tensor, $A = \phi^2 e^{\eta/3}$, $B = \phi^2 e^{-2\eta/3}$, $\psi = B^{1/2}$ and $\Delta^{(3)}_f$ and $\Delta^{(2)}_f$ are the three- and two-dimensional flat space Laplacians, respectively.

To find Cauchy data for the gravitational field, η and K^r_θ (freely specifiable fields) are taken in the form of a linear ingoing gravitational wave-packet with quadrupolar ($l = 2$) angular dependence. The general linear $l = 2$ solution is described by a quadrupole moment $I(v)$ of arbitrary profile in advanced time v (or retarded u). The linear solution involves $I(v)$, its first two derivatives, $I^{(1)}(v) \equiv dI/dv$ and $I^{(2)}(v)$, and its integrals, $I^{(-1)}(v) \equiv \int^v dv' I(v')$ and $I^{(-2)}(v)$.

Appropriate expressions for η and K^r_θ have been found

$$\eta = (\frac{I^{(2)}}{r} - 2\frac{I^{(1)}}{r^2})\sin^2\theta, \tag{3.22}$$
$$\frac{K^r_\theta}{r} = (\frac{I^{(2)}}{r^2} - 3\frac{I^{(1)}}{r^3} + 6\frac{I}{r^4} - 6\frac{I^{(-1)}}{r^5})\sin 2\theta. \tag{3.23}$$

These Cauchy data will be at least slightly non-linear (depending on the initial amplitude and radius), since a wave-packet of finite amplitude confined within a finite radius will generate a finite mass. In order to find proper data, the exact Hamiltonian and momentum constraints are solved for ϕ, λ and K^{φ}_{φ}, subject to the choice above of η and K^r_θ. Nearly linear Cauchy data are still found to generate ingoing solutions. The form of $I(v)$ is chosen, such that a wave-packet has polynomial radial dependence of the form $I^{(-2)}(v) = a\kappa_p L^5[1 - (v/L)^2]^6$, for $|v| = |r - r_o| < L$ at $t = 0$. Here, κ_p is a constant, a is an amplitude parameter, L a width parameter and r_o a centering parameter. Each of these might serve as useful parameters of spaces $\mathcal{G}_k$. Initially, L and r_o have been fixed, while a was chosen to parameterize the Cauchy data, i.e. the solutions. In the limit $a \to 0$ the mass of the wave-packet is $M_p^{linear} = a^2 L/(2\pi)$, therefore a strength parameter in the form $\theta(a) = 2\pi M_p/L \simeq a^2$ is a good choice. A wave-packet with $\theta \ll 1$ weakly self-interacts, escaping to infinity virtually unaffected, but for $\theta \geq 1$ a black hole forms with $m_{BH} \to M_p$ as $\theta \to \infty$. The critical value along the sequence was found to be $\theta^\star \simeq 0.80$ ($a^\star \simeq 0.93$).

Here again, supercritical collapse of gravitational wave-packets generates black hole masses that are found to satisfy a power law

$$m_{BH} \simeq C(a - a^\star)^\beta. \tag{3.24}$$

The critical exponent value is also $\beta \simeq 0.37$, presently indistinguishable from that seen in scalar field collapse. Similarly, a scaling relation holds on the gravitational field in the strong field region $\mathcal{R}$. The gravitational field oscillates on progressively finer spatial and temporal scales, as is evident from figure (3.1) [96], which shows radial profiles of η (along the equatorial plane $\theta = \pi/2$). As can be seen, η exhibits an echoing in $\rho = \ln r$ of the form

$$\eta(\rho - \Delta, t_n) \simeq \eta(\rho, t_{n+1}). \tag{3.25}$$

The times t_n are found here, by using the central value of the lapse function $\alpha(t, r = 0)$ as a diagnostic to determine the completion of successive oscillations. The value of the scaling constant Δ was found $\Delta \simeq 0.6$, so a radial scale ratio $e^\Delta \simeq 1.8$ differs in this case from the corresponding value $e^\Delta \simeq 30\,(\Delta \simeq 3.4)$ in scalar field collapse. This result, having been obtained in numerical simulations with several different resolutions, appears to be robust.

3.4.1.3 Other examples

Recently Evans and Coleman [97] reported new research work into critical phenomena in the gravitational collapse of a relativistic perfect fluid ($p = \rho/3$). Their model employs a radiation fluid ($\gamma_{adiabatic} = \frac{4}{3}$) in the spherical symmetry. In this case some analytic progress was possible, starting from an *ansatz* of self-similarity (i.e. scale-invariance rather than scale-periodicity). In numerical calculations the power law dependence of black hole

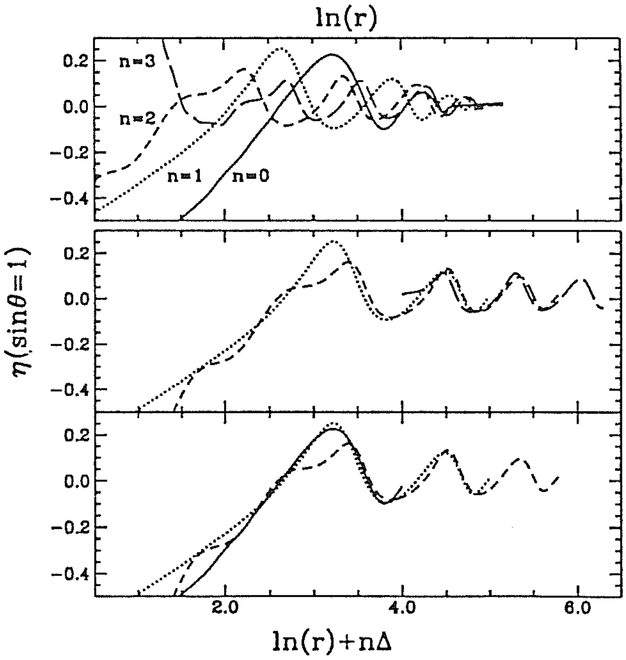

Figure 3.1: Scaling property of a near critical solution of axisymmetric gravitational wave collapse. Radial profiles of the metric function η (along $\theta = \pi/2$) plotted at four times corresponding to alternate maxima of the central value of the lapse function, α_c. The upper panel depicts all four profiles (labelled sequentially $n = 0 \div 3$) plotted versus $\rho = \ln r$. The two lower panels illustrate scaling by overlapping profiles that are shifted by $\rho \longrightarrow \rho' = \rho + n\Delta$ with $\Delta \simeq 0.6$. Profiles $n = 0, 1, 2$ are plotted in the bottom panel and $n = 1, 2, 3$ in the middle panel.

mass on $|p - p^{\star}|$ was obtained again with $\beta \approx 0.36$, as well as the evidence for a unique, self-similar critical solution in near-critical computations.

Also, Choptuik has found the same behaviour for massive scalar field in the spherical geometry [59].

Garfinkle [110] independently verified Choptuik's results for the case of a spherically symmetric scalar field. In the numerical calculations the parameter p has to be tuned to $p^{\star}$ to a great accuracy and the features must be resolved on extremely small scales. One might thus worry whether the results are not numerical artifacts. Garfinkle's algorithm uses the null initial value formulation of the problem (rather than the space-like one as Choptuik used) without adaptive mesh refinement. He finds results in agreement with those of Choptuik. In particular, the critical solution was verified to have the property of periodic self similarity: the scalar field evolves, after a certain amount of time, to a copy of its profile with the scale of space shrunk. A critical exponent $\beta \approx 0.38$ was found.

Strominger and Thorlacius [240] have reported the discovery of universality and mass-scaling in the context of the $2\,dim$ semiclassical RST [216] model, which employs null matter as a matter source. This model is used as a simplified model of quantum black hole evaporation and is exactly soluble. It was analytically demonstrated that the system exhibits universal power law mass-scaling at the critical point with $\beta = \frac{1}{2}$. Near critical scaling solution interpreted to describe the formation and evaporation of an arbitrarily small black hole was also found. In this model there is no analogue of the self-similar oscillations.

3.4.2 Discussion and conclusions

The present state of knowledge of critical phenomena in gravitational collapse is rather rudimentary. However they do indicate a connection between GR and scaling at critical point in statistical systems. One could ask how close this association really is to standard critical phenomena and how seriously one should take it. One would like to have analytic models and more results of numerical simulations of other kinds of space-times, with different symmetries and sources, performed in particular with the adaptive-mesh-refinement scheme for two and three dimensional cases.

Some analytic considerations on the subject were attempted, though so far they did not use any of the insights the above relation might provide. Let us be fair – the problem is extremely difficult. In [40] a spherically symmetric Einstein equations coupled to a massless (minimally coupled) scalar field were solved, assuming self-similar collapse (homothetic), and a self-similar solution with a critical parameter and critical regime was obtained. An exact one-parameter family of solutions was shown to exhibit a type of critical behaviour as discussed by Choptuik, though it is not clear that this is the universal strong field solution. For super-critical evolutions a quantity related to the mass of black hole exhibits a power law dependence on a parameter α, $m_{BH} \simeq |\alpha - \alpha_{crit}|^{1/2}$. The solution supports the conjecture that black hole formation initially occurs at infinitesimal mass.

The same model is considered in [202], where also Choptuik's scaling relations were derived. Critical behaviour was obtained which exhibits mass evolution $M_h \simeq (\frac{p-1}{8})^{1/2} v$ on the apparent horizon (M_h is the gravitational mass on apparent horizon, v the advanced time and p is the critical parameter and critical regime occurs for $p \to 1$). This equation exhibits mass evolution in some dynamical stage of gravitational collapse, instead of the asymptotic stage (at the future null infinity) $v \to \infty$, as evaluated by Choptuik.

If $X \equiv v/r$ the logarithmic evolution

$$\frac{1}{(p-1)X^2} = \log t - \log r$$

in a neighbourhood of $r = 0$ was derived. So for the self-similar solution the scaling relation $\phi(\log r - \Delta, \log t - \Delta) = \phi(\log r, \log t)$ clearly holds with a continuous parameter Δ. Choptuik however observed Δ discrete, namely, $n\Delta^*$, $\Delta^* \approx 3.4$, $n = 1, 2, 3, \ldots$. But this is not crucial in the region where $\log r, \log t \gg \Delta^*$. The above in fact holds in the strong field region when $p \to 1$ as then the apparent horizon is arbitrarily close to $r = 0$, and the spacing of $n\Delta^*$ is unimportant to recover approximately self-similarity outside the apparent horizon near $r = 0$. Generally, it is not clear how a continuous self-similarity could be at the center of previously obtained results. Both works used the exact self-similar solution.

The other approach is that of [153]. The analysis there shows that the one-parameter class of solutions decomposes into the subcritical and supercritical regimes. A proposal is made to develop a perturbation theory with a dimensionless expansion parameter $1/d$

to tackle the problem in $4d$ gravity. Critical exponent is found to be 0.5 at the limit of $d = \infty$.

Results of somewhat different flavour were provided in [157] where a simple model exhibiting critical behaviour in black hole formation (though not space-time self-similarity) was studied, namely, that of a thin shell coupled with outgoing null flux. A critical exponent was found analytically to depend rather strongly on the constants specifying the system.

As we saw the value of the relevant critical exponent was not obtained as 0.37. The difference (from 0.5), though small, is of crucial importance since this number, β, should reflect a *deep property of the gravitational field equations.* Suffice it to notice, that the experimental values for the critical exponent β (its explicit definition depends of course on the system considered) range from $0.305 \div 0.37$ for a variety of systems, like binary fluid, He I - He II transition, $\beta-$brass, a magnet (Fe) Ni and fluid (Xe), that are believed to belong to the same universality class [31]. It might be that the correct values of critical exponents can only be obtained when RG theory is made use of. This situation parallels the study of phase transitions in Condensed Matter before RG theory was available, when the mean field theory calculations were invariably giving critical exponents in form of rational numbers, in disagreement with experiment.

The results discussed in the previous sections can be described from a point of view which treats Einstein's equations as generating a RG flow on the space of initial conditions. Interestingly, it seems that the black hole mass m_{BH} plays a rôle of the order parameter for these critical phenomena[11], like spontaneous magnetization for ferromagnets (below the

[11] Note that there also seem to exist phase transitions (of second order) when we cool down black hole with respect to the corresponding Schwarzschild temperature $T_S = (8\pi m)^{-1}$, by increasing its charge Q and angular momentum J at fixed total mass. The heat capacity c_{JQ} passes then from negative (for a Schwarzschild black hole) to positive values (for Kerr-Newman) through an infinite discontinuity [176]. In [164] a possible second order phase transition of Reissner-Nordstorm black hole was studied, namely, at the thermally stable phase at which the heat capacity at constant charge is positive, the mean square fluctuations of mass and entropy of massive RN b. h. are divergent as certain critical temperature is approached. In the thermally unstable phase at which the heat capacity at constant charge is negative, the fluctuation probability diverges exponentially as the crit. temp. is approached. It turns out that the fluctuation-dissipation properties of the b.h. near the crit. temp. bear some resemblance to other thermodynamical systems. Could it be that the existence of a thermally unstable phases suggest a new kind of crit. phenomenon peculiar to gravitation?

Curie temperature) or $|\rho - \rho_c|$ for liquid-gas transition in the co-existence region. Indeed, black holes appear only for solutions with $p > p^\star$ and a black hole of infinitesimal mass is conjectured to exist at $p = p^\star$. This analogy is supported by the fact that in numerical simulations the solutions approach universal scaling solutions at criticality.

Choptuik demonstrated that details inherent in the original data are "washed out" in the strong field region $\mathcal{R}$ in near critical evolution. With each echo as $r \to 0$ and $T \to T^\star$ information may be steadily lost and its rate per echo may depend on the value of Δ. Likewise, an analogue of the correlation length ξ in statistical mechanics systems seems to be now the ratio of the radii of the outer edge of the scaling region, r_{max}, and the inner edge, r_n, of the innermost echo i.e., $\xi \sim r_{max}/r_n \sim e^{n\Delta}$. As $p \to p^\star$ an ever larger region (in terms of the scale r_n) becomes "correlated" with self similar echoes and $\xi \to \infty$, as it should for "critical" systems.

One can attempt to classify the self-similar (scaling) solutions to Einstein's equations by viewing the evolution equations as generating a RG flow on the space of initial conditions. We follow here a general discussion in [6].

If we define the RG transformation with parameter Λ as (in the scalar field case)

$$\phi(r,t) \to \phi_\Lambda(r,t) \equiv \Lambda^{\beta_\phi}\phi(\frac{r}{\Lambda^{\alpha_r}}, \frac{t^\star - t}{\Lambda^{\alpha_t}}), \tag{3.26}$$

the self-similar solutions of the equations would then correspond to a fixed point of this transformation. E.g. a continuously self-similar solution $\phi^\star$ satisfies $\phi^\star(r,t) = \phi^\star_\Lambda(r,t)$ for all Λ. When the solution is discretely self-similar it would correspond to a *limit cycle* of RG flow $\phi^\star(r,t) = \phi^\star_{\Lambda_0^n}(r,t)$ for all integers n. Only the perfect fluid collapse gives a simple fixed point; the others have limit cycles.

To obtain the critical exponents one could approach using a linear perturbation theory around the critical (scaling) solution. So far however it has not been understood however. Moreover, the critical point is in the strongly coupled regime. This poses additional problems. One should thus look for a simpler system in the same universality class.

Interesting connections seem to exist between the universality in gravitational collapse and no-hair behaviour [205].

3.5 RG in cosmology – motivation

The usual RG transformations invoking averaging over square blocks are designed mostly having ferromagnetic systems in mind. However, there are many problems suitable for RG methods but it seems they have not yet been expressed in such a way that they can be solved. They are amongst the hardest problems known in physics where their difficulty can be traced to a multiplicity of scales [259, 258]. The important issue in the proper application of RG theory to a particular problem at hand is the choice of variables of the model and the RG map. For each new physical situation one has to "custom-make" the RG map, as Michael Fisher puts it clearly [101]:

> For any given Hamiltonian or class of Hamiltonians there is not just one renormalization group - "the renormalization group" as some people say - but rather there are many that might be introduced, and one must question, for example, whether the process is best carried out in real space or momentum space and so on. A "good" renormalization group must be "apt" or appropriate for the problem at hand, and it must, in particular, "focus" properly on the critical phenomena of interest.

Some form of the Renormalization Group is active on any system where there are fluctuations present. This is so, because one can integrate the fluctuations out of the physical quantities of interest, e.g. the partition function, and depending on the "scale" up to which one is integrating, the same quantities that emerge are different. The functional relation between them provide recursion relation between the physical parameters, the coupling constants, which characterize the physics at each scale, and this is precisely the RG. The nature of the fluctuations does not need to be quantum. For example, they could be thermal fluctuations as in Statistical Mechanics, or they could be due to collisions as in many body perturbed motion in GR, such as the ones due to "frictional" processes in the Universe, be them ones in the epochs of large entropy production [21], or those originating in dynamical friction, or purely chaotic processes due to the many body nature of the gravitational system.

If the universe can be considered as a complex general relativistic many body system (GRMBS), the question then arises whether one could apply the methods, much as in statistical mechanics to understand some of its features and undertake the study of its collective behaviour, local morphological problems, etc., by taking advantage of their suitability for the study of this class of problems. A recognition of powerful methods, deriving from Renormalization Group, that can be used to study complex systems without loosing their physical picture encourages us to undertake this task. A common feature of many-body systems is that (under certain conditions) they may exhibit condensation-like phenomena (e.g. formation of Cooper pairs in superconductors). One can thus anticipate there would be a possibility of forming extended (elementary excitations) as well as localized states (solitary waves).

In the case of quantum fluctuations we have the Gell-Mann-Low version of the RG, which for discrete iterations (or blockings) is related to the Wilsonian RG. In field theory the use of the Gell-Mann-Low version of the RG is based upon perturbation theory, whilst the Wilson one has a direct geometrical interpretation (as already stressed before) and is in principle non-perturbative. For thermal fluctuations, we have RG which seems similar to the encarnation appearing in the Ising model (Heisenberg model, more probably). In the case of many body perturbed motions we have to develop ideas (see next sections).

3.6 RG approach to averaging in cosmology

Contents of this section is based on [57].

Some notions and facts relevant to the contents of this section are given in Appendix C (in particular, a definition of the Gromov distance can be found there), which the reader is encouraged to consult. General mathematical preliminaries can be found in Appendix A.

3.6.1 Coarse-graining in Cosmology

A possible solution to the averaging problem would be to explicitly construct a procedure for carrying out the smoothing process in the full theory. Almost all existing attempts were concerned with the linearized theory, with a possible exception of [51] (see also [141]).

In [51] a covariant smoothing-out procedure was put forward for the space-times associated with gravitational configurations which may be considered near to the standard ones, generating closed FLRW universes (see section 2.4.1). The procedure makes use of Hamilton's theorem about smooth deformations of three-metrics and is adapted for smoothing-out an initial data set for cosmological solutions to the Einstein equations. While interesting in its own right, this approach seemed rather *ad hoc* and not yet capable of resolving the issues of actual limits of validity of the FLRW models in cosmology.

Now, our hope is that there is a smart and simpler way to the heart of the problem, borrowing from the known theories and methods of statistical mechanics, based on the real-space Renormalization Group (RG) approach to study critical phenomena in lattice models [147, 257, 259]. The averaging problem in cosmology can be looked at and studied as belonging to precisely the kind of many length scales problems. These are difficult problems where the reductionist approach fails and where the effective degrees of freedom of a physical system are scale dependent. The difficulty of this kind of issues lies, as already mentioned, in a multiplicity of scales and moreover there can presumably be a gross mismatch between the largest and smallest scales in the problem.

Often a major step consists in finding a way of looking at things. Therefore we stress that the problem we face with the averaged description in cosmology is effectively a question of how a system behaves under changes of "scales". As such it is most naturally addressed using RG approach, understood here as a general strategy to handle problems of multiple length scales[12] which allows us to extract the long distance behaviour of the system by making the scale successively coarser. In cosmology, we have the curvature inhomogeneities and to consistently tackle this problem, we will have to consider a procedure operating on the metric, not only on or apart from the matter present.

To provide even more support in favour of the above idea notice that one can also be guided by scaling ideas, of which the power law behaviour of the two-point correlation function for galaxies, clusters and quasars, is a fair example. Scaling, on its own right, is deeply understood within the underlying mathematical scaffold which is RG. These are

[12] Although the renormalization procedure might seem purely formal there are important physical ideas behind it, namely, that of *scaling* and *universality*.

hints therefore that one can regard the universe as a gravitational dynamical system not far from criticality (understood intuitively by analogy with e.g. a ferromagnet). Later one can also try to qualify the precise nature of the critical behaviour within the phase transition context; for a simple example see section 3.6.3.3.

The real space renormalization techniques are mostly applicable to discretized models, based on a lattice. Therefore we now turn to describing a suitably discretized manifold model we are going to work with.

3.6.1.1 Discretized manifold model

The approach taken is that of a $(3+1)$ formulation of General Relativity (GR) [9]. Let us suppose we have a differentiable, compact three-manifold (without a boundary) $\mathcal{M}$. Generally, in this case we will always assume that these manifolds possess certain natural constraints on their diameter and a suitably defined notion of curvature. The point of this requirement is that the manifolds, or more precisely the riemannian structures, can then be classified according to how they can be covered by small metric balls (to be defined later). Moreover the space of riemannian structures has some remarkable compactness properties. This is a classical result obtained by M. Gromov [119] (see also [109]). On a set of riemannian structures it is possible to introduce a distance function, the Gromov distance, which roughly speaking enables one to say something about how close particular manifolds are to each other. For the riemannian manifolds which can be considered close to each other (in the sense of Gromov distance) it is possible to cover them with the balls arranged in similar packing configurations [52].

In order to define such coverings [122], let us parameterize the geodesics by arc-length, and for any point $p \in \mathcal{M}$ let $d_{\mathcal{M}}(x,p)$ denotes the distance function of the generic point x from the chosen one p. Then for any given $\epsilon > 0$ it is always possible to find an ordered set of points $\{p_1, \ldots, p_N\}$ in $\mathcal{M}$, so that [122]

i) the open metric balls (the geodesic balls) $B_{\mathcal{M}}(p_i, \epsilon) = \{x \in \mathcal{M} | \; d_{\mathcal{M}}(x, p_i) < \epsilon\}$, $i = 1, \ldots, N$, cover $\mathcal{M}$; in other words the collection $\{p_1, \ldots, p_N\}$ is an *ϵ-net* in $\mathcal{M}$.

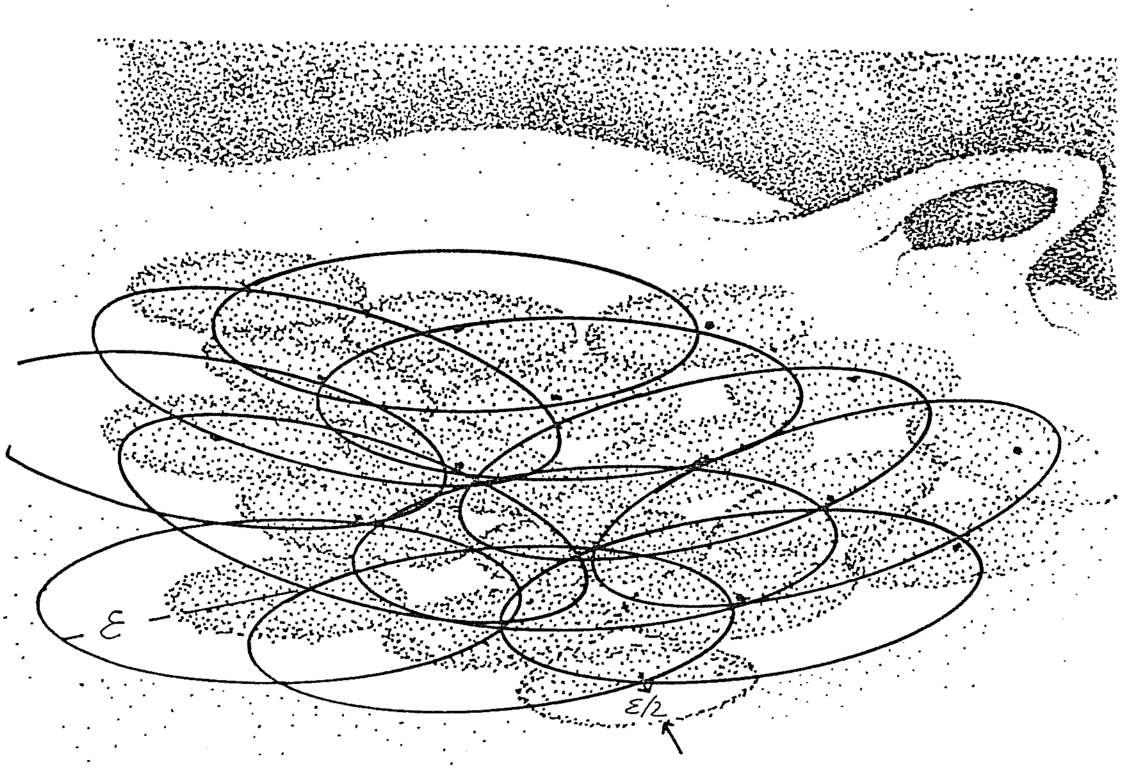

Figure 3.2: A portion of a minimal geodesic balls covering. Dotted disks are the $\frac{\epsilon}{2}$-balls which pack a given region. The larger, undotted disks, represent the ϵ-balls which provide the covering.

ii) the open balls $B_{\mathcal{M}}(p_i, \epsilon/2)$, $i = 1, \ldots, N$ are disjoint, i.e., $\{p_1, \ldots, p_N\}$ is a *minimal* ϵ-net in $\mathcal{M}$.

It is fair to say that as a consequence of the compactness properties of the set of riemannian structures that we consider, for each "length scale ϵ" there exists a finite number of "model" geometries which describe with an $\epsilon-$approximation any given riemannian geometry. Namely given a ball of a certain radius $> \epsilon$ in any riemannian manifold (with suitable restrictions on their volume, diameter and sectional curvature, as we have said earlier) there exists a ball metrically similar (up to an ϵ scale) in one of the model geometries which does not retain the details of the original manifold on scales smaller than ϵ. Roughly speaking, ϵ is a measure of the typical curvature inhomogeneity with respect to the model background. Let us stress that this is a highly non-trivial result, in the sense that the metrical properties of the manifolds from an infinite dimensional set are, up to an ϵ scale, described by the metrical properties of just a finite number of model riemannian manifolds.

The $\epsilon-$nets underlying the balls-coverings precisely provide the discretized manifold model. This coarse-graining of a manifold according to Gromov is the most natural coarse-graining one can think of, pertinent for manifolds with a lower bound to the sectional curvature. This assumption does not limit the generality of our analysis which is basically motivated by a concrete physical problem, whose nature allows us to deal from the beginning with manifolds that are already in a certain sense quasi-homogeneous (Cf. comments on the solvability of the Ricci-Hamilton flow, later on).

In what follows, when speaking of balls we will always mean geodesic balls here.

3.6.1.2 An empirical averaging procedure

We assume that we have chosen a particular space-like hypersurface Σ of the $4-$dimensional manifold, on which the average of a scalar function $f : \Sigma \rightarrow \mathbb{F}$ is given as

$$< f >_{\Sigma(g)} = \frac{\int_\Sigma f d\mu_g}{vol(\Sigma, g)}, \tag{3.27}$$

where $vol(\Sigma, g) = \int_\Sigma d\mu_g$ and μ_g is the riemannian measure associated with the three-metric g of Σ. Since at this stage we simply wish to put forward a few elementary geometrical considerations, we do not specify yet the choice of the hypersurface Σ and we do not attribute any particular physical meaning to the function f.

If the geometry of Σ is not known on a large scale, we cannot take (3.27) as an operational way of defining the average of f. From a more pragmatic point of view, supposing that we can only experience geometry in sufficiently small neighborhoods of a finite set of instantaneous observers, it makes much more sense to replace (3.27) with a suitable average based on the geometrical information available on the length scale of such observers.

For simplicity, given a finite set of instantaneous observers, located at the points $x_1, \ldots, x_N \in \Sigma$, we may assume that these susceptible to observation regions are suitably small geodesic balls of radius ϵ, scattered over the hypersurface Σ so as to cover it. In other words, we assume that $\{x_1, \ldots, x_N\}$ is a minimal $\epsilon-$net in Σ. Further, we denote by

U_ϵ the corresponding set of geodesic balls $\{B_\Sigma(x_i, \epsilon)\}, i = 1, \ldots, N$. Then we can bound (3.27) as

$$< f >_{\epsilon/2} \frac{\sum_h vol(B_h, \epsilon/2)}{vol\Sigma(g)} \leq < f >_{\Sigma(g)} \leq < f >_\epsilon \frac{\sum_h vol(B_h, \epsilon)}{vol\Sigma(g)}, \tag{3.28}$$

where

$$< f >_\epsilon \equiv \frac{\sum_i \int_{(B_i, \epsilon)} f d\mu_g}{\sum_i vol(B_i, \epsilon)}, \tag{3.29}$$

and

$$< f >_{\epsilon/2} \equiv \frac{\sum_i \int_{(B_i, \epsilon/2)} f d\mu_g}{\sum_i vol(B_i, \epsilon/2)}, \tag{3.30}$$

and where we explicitly indicated the dependence of the averages on a particular "covering". This suggests the consideration of $< f >_\epsilon$ as a suitable scale dependent approximation to $< f >$.

There are certain problems lurking that we have to clear up. Obviously, there are "unwanted" details affecting the average function over the discretized manifold, as given by its partition with a collection of geodesic balls, the immediate one being the underlying discretization. The important question to ask is what happens to the average when we change the length scale. Depending on whether we are actually increasing or decreasing it, respectively less or more details of the underlying geometry, will be felt by the average values. The natural philosophy is that over scales big enough no details should be discerned since the homogeneity and isotropy prevails. This is the reason why on constant curvature spaces averaging is well defined since there one can move the balls freely and deform them, but by so doing no new geometric details that measure the inhomogeneities will be felt in the averaged values of quantities we are interested in.

A natural question to ask now is then how the geometry, specifically curvature inhomogeneities, should depend on scale so that the average over the balls is scale independent, or equivalently, how do we have to deform the geometry in order to achieve the scaling limit when size of the balls matters no more?

First, we give some details on how we calculate the averages according to (3.29). In order to do this, we employ a preferred system of coordinates on $\{B_i\}$ given by the local diffeomorphism

$$\exp_x : T_x\Sigma \to \Sigma \tag{3.31}$$

i.e., we make use of the exponential mapping

$$\varphi_i \equiv \exp|_{\exp^{-1} B_i \equiv D_i} : D_i \to B_i\,, \tag{3.32}$$

where $D_i = D(x_i, \epsilon)$ is the ball in $T_{x_i}\Sigma$.[13]

On D_i we use polar coordinates and pull-back the riemannian measure accordingly, namely,

$$\varphi_i{}^*(\mu_g) = \theta(t, x_i) dt \otimes dx_i, \tag{3.33}$$

where dx_i denotes the canonical measure (euclidean volume form) on the unit sphere $D(x_i, 1) = S^2_1 \subset T_{x_i}\Sigma$ and where dt is the Lebesgue measure on $\mathbb{R}$ $(t \geq 0)$.

For t small enough one can prove Puiseux' formula

$$\theta(t, x) = t^{n-1}(1 - \frac{1}{3} r(x) t^2 + \mathcal{O}(t^2)), \tag{3.34}$$

where $n = dim\,\Sigma$ and $r(x)$ is the Ricci curvature $Ric(g)$ (at the point x).

Using this result we have

$$\int_{B(x_i,t)} f d\mu_g = \int_{S^2_1} f\theta(t, x) dx_i dt. \tag{3.35}$$

Let us consider the asymptotic expansion with respect to t [166]

$$\int_{B(x_i,t)} f d\mu_g = \omega_n t^n [f(x_i) + \frac{t^2}{2(n+2)}(\Delta f(x_i) - \frac{R(x_i)}{3} f(x_i)) + \mathcal{O}(t^2)], \tag{3.36}$$

where ω_n is the volume of the unit ball of $\mathbb{R}^n$, R is the scalar curvature at the center of the ball and Δ the Laplacian operator relative to the manifold.

[13]The transition from (Σ, g), at a given moment of time t_o, to a tangent space parallels the prescription in riemannian geometry for measuring [219]. Let U be a given neighborhood of an instantaneous observer in Σ and suppose it is so small that there is a neighborhood A of $0 \in T_x\Sigma$ such that $\exp_x : A(\subset T_x\Sigma) \to U(\subset \Sigma)$ is a diffeomorphism. One can then replace the considerations in (U, g) by those in A (with the riemannian measure pulled back) via $\exp_x^{-1}$. Namely, we can say that an instantaneous observer in (Σ, g) observes the universe with the help of the exponential mapping, which just means projecting structures from an open neighborhood $U \subset \Sigma$ of x by $\exp_x^{-1}$ and treating them as structures on $T_x\Sigma$.

Substituting $f = 1$ in the above formula we get the asymptotic expansion of the volume of a geodesic ball:

$$vol(B(x_i, t)) = \omega_n t^n (1 - \frac{R(x_i)}{6(n+2)} t^2 + \mathcal{O}(t^2)). \tag{3.37}$$

These standard formulae are what we need in order to calculate how the average value behaves when we change the scale.

Since we are interested in discussing how $< f >_\epsilon$ behaves upon changing the radius of the balls $\{B(x_i, \epsilon)\}$, let us consider the average $< f >_{\epsilon_o+\eta}$, with η a positive number with $\eta/\epsilon_o \ll 1$. According to the formulae recalled above, we can write

$$< f >_{\epsilon_o+\eta} = \frac{\sum_i [f_i + (\frac{\Delta f_i - R_i f_i/3}{2(n+2)})(\epsilon_o + \eta)^2]}{\sum_i [1 - \frac{R_i}{6(n+2)}(\epsilon_o + \eta)^2]} + \mathcal{O}((\epsilon_o + \eta)^4), \tag{3.38}$$

where we have introduced somewhat simplified but otherwise obvious notation. Upon expanding this expression in η, we get to leading order ($\mathcal{O}(\epsilon_o^4)$ in ϵ_o, and $\mathcal{O}([\frac{\eta}{\epsilon_o}]^2)$ in $\frac{\eta}{\epsilon_o}$),

$$< f >_{\epsilon_o+\eta} \simeq < f >_{\epsilon_o} + \frac{1}{n+2} < \Delta f >_{\epsilon_o} {\epsilon_o}^2 \frac{\eta}{\epsilon_o} + \frac{1}{3(n+2)} [< R >_{\epsilon_o} < f >_{\epsilon_o} - < Rf >_{\epsilon_o}] {\epsilon_o}^2 \frac{\eta}{\epsilon_o}, \tag{3.39}$$

where $< f >_{\epsilon_o}$ is the average of the function f over the set of N instantaneous observers U_{ϵ_o}, (with similar expressions for $< R >_{\epsilon_o}$, $< Rf >_{\epsilon_o}$ and $< \Delta f >_{\epsilon_o}$). Thus, under a change of the cutoff we can write

$$\epsilon_o \frac{d}{d\eta} < f >_{\epsilon_o+\eta}|_{\eta/\epsilon_o=0} = \frac{1}{n+2} < \Delta f >_{\epsilon_o} + \frac{1}{3(n+2)} [< R >_{\epsilon_o} < f >_{\epsilon_o} - < Rf >_{\epsilon_o}] \tag{3.40}$$

to leading order. In the next sections we will discuss the consequences of these formulae and the connection of our averaging procedure with the Ricci-Hamilton flow.

3.6.2 The Renormalization Group view

3.6.2.1 Block variables and recursion relations

The real space RG technique is based on the recursive introduction of *block variables.* A method of "blocking" is trivial to introduce on regular lattices. In particular, in the

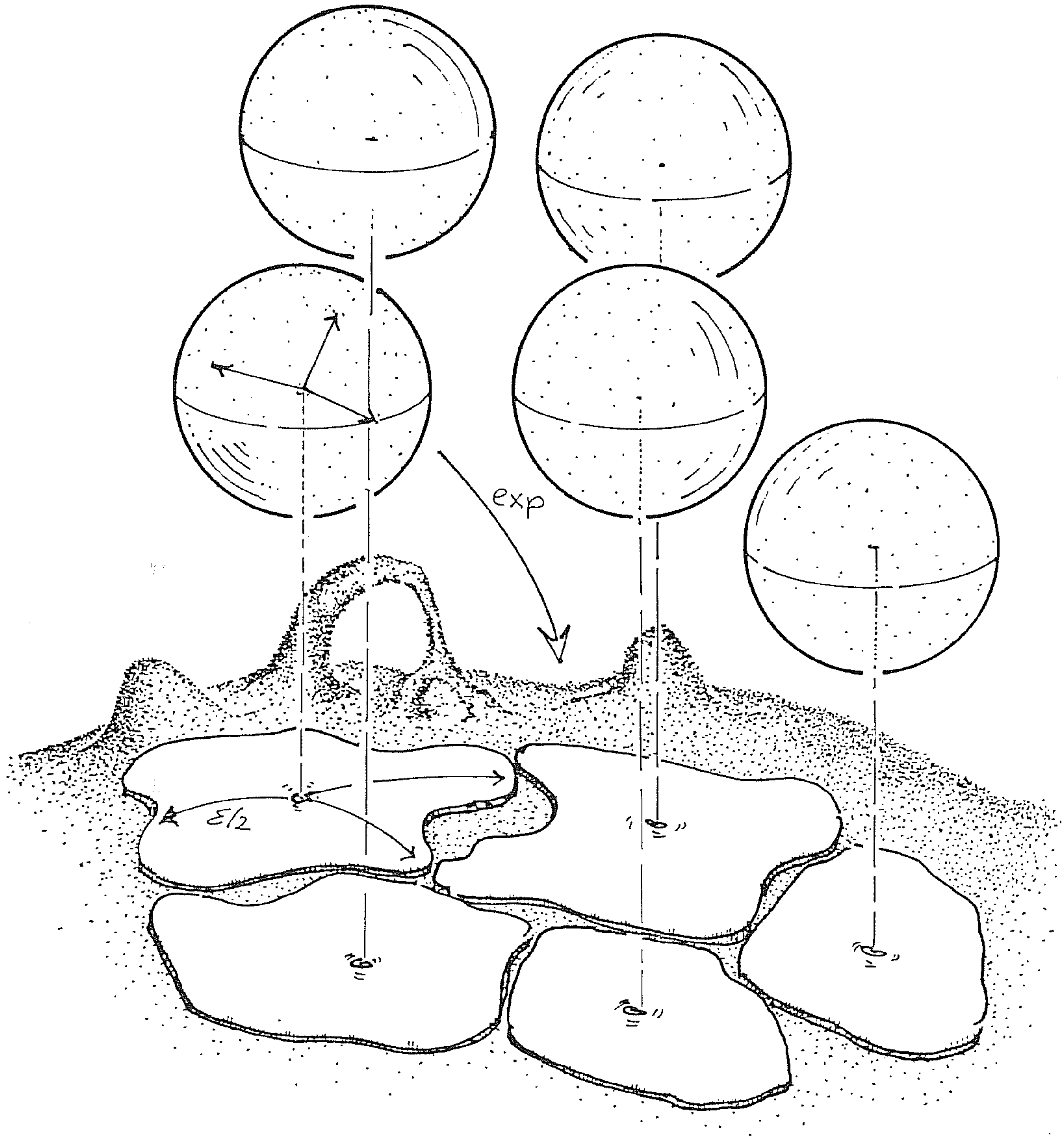

Figure 3.3: We can approximate the average over Σ by averaging over euclidean balls whose Lebesgue measure has been locally weighted through Puiseux' formula. In this drawing, the $\frac{\epsilon}{2}$-geodesic balls are represented by curved disks on Σ, while the corresponding euclidean balls are correctly depicted as 3-dimensional.

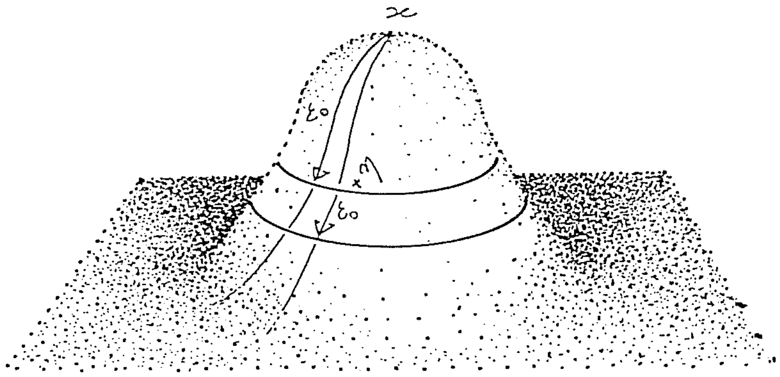

Figure 3.4: The local average of a function f over a geodesic ball $B(x, \epsilon_o)$ feels the underlying curvature of the manifold through Puiseux' formula. In particular, in passing from $B(x, \epsilon_o)$ to the larger ball $B(x, \epsilon_o + \eta)$ we get correction terms which depend on the fluctuations in curvature.

case of the Ising model (2−dim) it consists of a subdivision of the spin system into cells which supposedly interact in a similar way as the original spins. This can be done by introducing the block-spin variables via the majority rule or decimation procedure [31]. The behaviour of blocked lattice on large scales is equivalent to the behaviour of the original lattice corresponding to a different temperature. The slopes of the parameters' surface close to the *critical fixed point* determine the macroscopic characteristics of the model (see e.g. [31]).

In the context of the problem we are considering, the application of RG method forces us to invent some analogue of Kadanoff's blocking (block-spin transformation) applied to the geometry itself. This appears to be a difficult problem since in a general case when the geometry is curved the lattice itself takes on a dynamical rôle. Moreover, since the manifolds Σ we are dealing with are compact, (closed, without boundary), we must implement a Renormalization Group strategy in a finite geometry, and thus the relevant phenomena are here related to *finite size scaling.* Roughly speaking, the size of our manifold is characterized by a length scale, say L, (actually the volume, the diameter and possibly bounds on curvatures), which is large in terms of the microscopic scale (the radius of the typical geodesic ball coverings we shall use in the blocking procedure). A continuous theory, describing the (universal) properties of the field f on Σ, arises when the correlation length associated with the distribution of f is large (of the order of L). This being the underlying rationale, let us proceed and be guided by the formula (3.40).

Let us consider our system as nearly infinite, i.e. the manifold Σ divided by the collection of the geodesic balls of radius $\epsilon_m \equiv (m+1)\epsilon_o$ for $m = 0, 1, 2, \ldots$, and where ϵ_o, the chosen cutoff, is much smaller than the typical length scale associated with Σ, (this length scale can be identified with the injectivity radius of the manifold). Each ball will be labelled by k in the sequel and the original geodesic packing-covering is for $m = 0$. We can introduce this way a convenient notation for the integral of the generic function f over the ball $B(x_k, \epsilon_m)$ as

$$\psi_m(k; f) \equiv \int_{B(x_k, \epsilon_m)} f d\mu_g, \tag{3.41}$$

which can be seen as the block variable since it allows us to eliminate from the distribution of the field f all fluctuations on scales smaller than the cutoff distance ϵ_m. We wish to emphasize that if the geometry of the ball $B(x_k, \epsilon_m)$ is not flat, then the definition of $\psi_m(k; f)$ can be interpreted as that of a *weighted* sum over a flat ball, namely,

$$\psi_m(k; f) \equiv \int_{\exp^{-1} B(\epsilon_m)} f\theta(t, x) dt \otimes dx\,, \tag{3.42}$$

where the weight $\theta(t, x)$ is provided by Puiseux' formula (3.34).

If we consider the covering of $B(x_k, \epsilon_{m+1})$ induced by the geodesic balls $\{B(x_j, \epsilon_m)\}$, i.e., the collection of $N_k(m, m+1)$ open sets $\{B(x_k, \epsilon_{m+1})\} \cap \{B(x_j, \epsilon_m)\}$, then the above definition of block variables can be written recursively in terms of the values the function f takes correspondingly to balls of larger and larger radii. To this end, let us consider a partition of unity $\{\xi_h\}_{h=1,\ldots,N_k(0,m+1)}$, subordinated to the covering of the generic enlarged ball $B(x_k, \epsilon_{m+1})$, induced by the geodesic balls $\{B(x_h, \epsilon_o)\}$. Namely, a set of smooth functions such that: $0 \leq \xi_h \leq 1$, for each h; the support of each ξ_h is contained in the corresponding $B(x_h, \epsilon_o)$; and $\sum_h \xi_h(p) = 1$, for all $p \in B(x_k, \epsilon_{m+1})$.

Under such assumptions, the block variables $\psi_m(k; f)$ can be written recursively as

$$
\begin{aligned}
\psi_o(k;f) &= \int_{B(x_k,\epsilon_o)} \xi_k f d\mu_g, \\
\psi_{m+1}(h;f) &= \sum_{k}^{N_h(m,m+1)} \psi_m(k;f).
\end{aligned} \tag{3.43}
$$

Indeed, we have

$$
\begin{aligned}
\psi_1(h;f) &= \int_{B(y_h,\epsilon_1)} f d\mu_g = \sum_k^{N_h(0,1)} \int_{B(x_k,\epsilon_o)} \xi_k f d\mu_g \\
&= \sum_k^{N_h(0,1)} \psi_o(k;f),
\end{aligned} \tag{3.44}
$$

and

$$
\begin{aligned}
\psi_2(j;f) &= \int_{B(z_j,\epsilon_2)} f d\mu_g = \sum_h^{N_j(1,2)} \int_{B(y_h,\epsilon_1)} \xi_h f d\mu_g \\
&= \sum_h^{N_j(1,2)} \sum_k^{N_h(0,1)} \int_{B(x_k,\epsilon_o)} \xi_k \xi_h f d\mu_g \\
&= \sum_h^{N_h(1,2)} \psi_1(h;f),
\end{aligned} \tag{3.45}
$$

where we have exploited the fact that the functions $\{\xi_k \xi_i\}$ have $\xi_k \xi_i(p) = 0$, except for a finite number of indices (k,i), and $\sum_k \sum_i \xi_k \xi_i(p) = 1$, for all $p \in \Sigma$. The above expressions readily generalize for every m.

More explicitly, we can write these block variables as

$$
\psi_{m+1}(k;f) = \int_{B(x_k,\epsilon_{m+1})} f d\mu_g =
$$
$$
\omega_n \epsilon_m^n \sum_h^{N_k(m,m+1)} \left[f(h)\xi_h + \left(\frac{\Delta(f\xi_h)(h) - R(h)f(h)\xi_h/3}{2(n+2)} \right) \epsilon_m^2 + \mathcal{O}(\epsilon_m^4) \right]. \tag{3.46}
$$

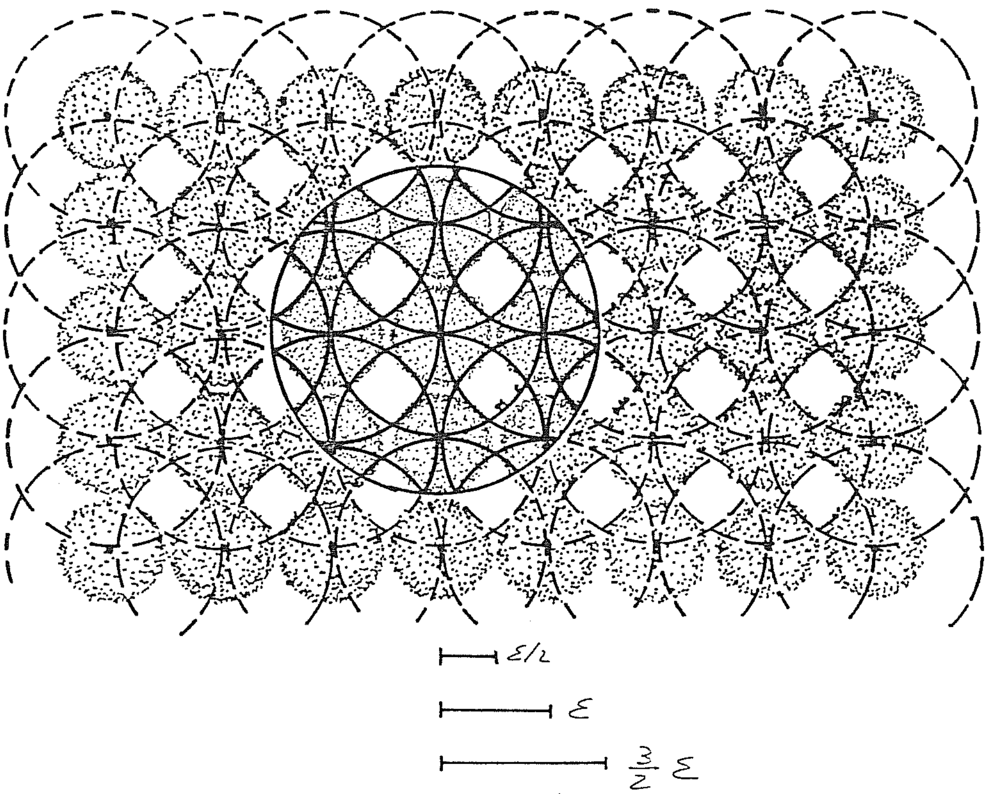

Figure 3.5: The intricate but symmetrical intersection pattern of an array of $\frac{\epsilon}{2}$-balls in the $2D$-plane. Here, the $\frac{\epsilon}{2}$-balls are dotted while the overlapping ϵ-balls, providing the covering, are the dashed circles. The solid circles and the arcs of circles describe the intersection pattern of a $\frac{3}{2}\epsilon$-ball. On a curved manifold of bounded geometry, the pattern is more complicated and not symmetrical at all. Nonetheless, one may obtain a recursive definition of blocking by exploiting a partition of unity argument, explained in the text.

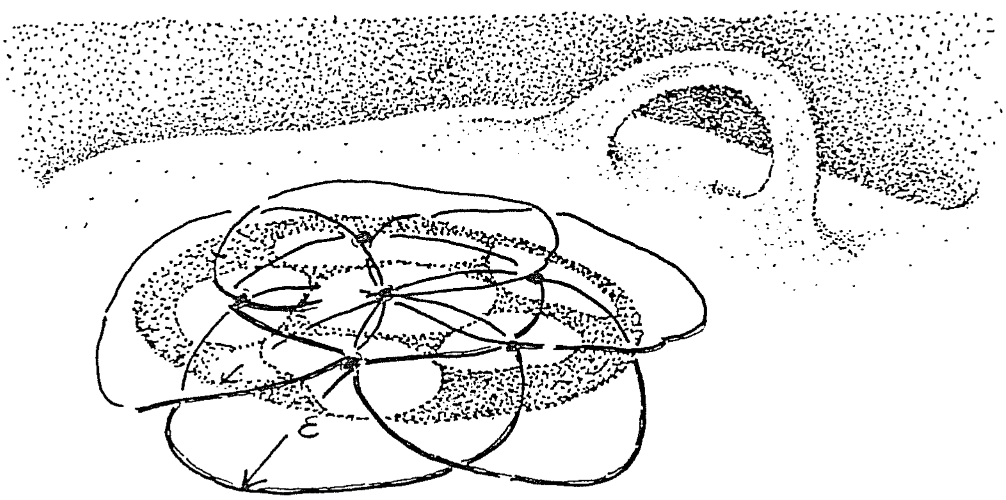

Figure 3.6: The intersection pattern of a 2ϵ-ball $B(x_j, 2\epsilon)$ with five ϵ-balls $B(x_h, \epsilon)$. The 2ϵ-ball is represented by a dotted disk, while the ϵ-balls are represented by solid circles. The white disks represent the packing of $B(x_j, 2\epsilon)$ with $\frac{\epsilon}{2}$-balls. The standard representation of the integral Bover $B(x_j, 2\epsilon)$, through the partition of unity subordinated to the $B(x_h, \epsilon)$-covering, yields the recursive relation $\psi_{m+1}(j) = \sum_h \psi_m(h)$.

Notice that in terms of the block variables $\psi_m(k; f)$ we can rewrite the empirical averaging (3.29) as

$$< f >_{\epsilon_m} (\psi_m) = \frac{\sum_k^{N(\epsilon_m)} \psi_m(k; f)}{\sum_k^{N(\epsilon_m)} vol[B(x_k, \epsilon_m)]}, \tag{3.47}$$

where $N(\epsilon_m)$ denotes the number of distinct $B(x_k, \epsilon_m)$ balls providing a minimal ϵ_m-covering of the manifold Σ. Thus, when m is sufficiently large, the variation in $< f >_{\epsilon_m}$ (ψ_m) under a block transformation $\psi_m(k; f) \to \psi_{m+1}(h; f)$ is given (to leading order) by

$$\begin{aligned} &< f >_{\epsilon_{m+1}} (\psi_{m+1}) - < f >_{\epsilon_m} (\psi_m) \simeq \\ &\frac{1}{n+2} < \Delta(f\xi_h) >_{\epsilon_m} {\epsilon_m}^2 \frac{1}{m+1} + \frac{1}{3(n+2)} [< R >_{\epsilon_m} < f\xi_h >_{\epsilon_m} - < Rf\xi_h >_{\epsilon_m}] {\epsilon_m}^2 \frac{1}{m+1}. \end{aligned} \tag{3.48}$$

The above choice of block variables brings out the coupling between averaging a scalar field over a manifold and the presence of fluctuations in the curvature of the underlying geometry.

In order to be more precise, let us assume that the variables $f(k)$ are randomly distributed according to some probability law $P(\{f(k)\})$ (later on we shall come back to this point with a definite prescription). Upon blocking the system and thus renormalizing the variables $f(k)$ by increasing the scale size, the probability distribution $P(\{f(k)\})$ induces a corresponding probability distribution on the variables ψ_m, *viz.*, $P(\{\psi_m(k; f)\})$.

From equation (18) it is clear that if the geometrical properties of any two balls, $B(x_i, \epsilon_{\bar{m}})$ and $B(x_j, \epsilon_{\bar{m}})$ (with $B(x_i, \epsilon_{\bar{m}}) \cap B(x_j, \epsilon_{\bar{m}}) = \emptyset$), are not correlated then the corresponding block variables, $\psi_{\bar{m}}(i; f)$ and $\psi_{\bar{m}}(j; f)$, are uncorrelated. Such $L \equiv \epsilon_{\bar{m}}$ characterizes the correlation (or persistence) length of the manifold (Σ, g). It is a measure of the typical linear dimension of the largest ball exhibiting a correlated spatial structures. This correlation length can be seen in close analogy with the usual correlation length (usually denoted by ξ) in condensed matter systems. It depends there upon the coupling constants in particular upon temperature, and diverges to infinity at the phase transition point.

Since g plays here the rôle of a running coupling, or if you prefer, of "temperature", the existence of a finite correlation length corresponds to a rather "irregular", crumpled geometry (as seen on scales of the order of L), or, equivalently, a *high temperature phase* of our system.

According to the central limit theorem it follows, for m large enough, ($\epsilon_m \gg L$), that the block variables $\psi_m(k; f)$, being the sum of uncorrelated random variables, are normally distributed, (let us say around zero, for simplicity), with a variance

$$E_P(\psi_m^2(i; f)) = N^{(\epsilon_o, m)}\bar{\chi}, \tag{3.49}$$

where $E_P(\ldots)$ denotes the expectation according to the probability law $P(\{\psi_m\})$, $\bar{\chi}$ is related to the variance of the variables $f(k)$, $R(k)$, and $N^{(\epsilon_o, m)}$ denotes the number of ϵ_o-balls in the m-ball $B(x_i, \epsilon_m)$.

Thus, irrespective of the details of the local distribution of the random variables $f(k)$ and $R(k)$, we can write for the distribution of $\{\psi_m(k; f)\}_1^N$ over (Σ, g)

$$dP(\{\psi_m\}) = \prod_k \left[d\psi_m(k;f) \left(2\pi N^{(\epsilon_o,m)}\bar{\chi}\right)^{-1/2} \exp\left[-\psi_m^2(k;f)/2N^{(\epsilon_o,m)}\bar{\chi}\right]\right]. \qquad (3.50)$$

This shows that by rescaling the block variables $\psi_m(k;f)$ according to

$$\phi_m(k;f) \equiv \left[N^{(\epsilon_o,m)}\right]^{-1/2} \psi_m(k;f), \qquad (3.51)$$

we get new block variables with a finite variance as $m \to \infty$, and for *random* metrics (Σ, g) we can write

$$dP(\{\phi_m\}) = \prod_k \left[d\phi_m(k;f)(2\pi\bar{\chi})^{-1/2} \exp\left[-\phi_m^2(k;f)/2\bar{\chi}\right]\right]. \qquad (3.52)$$

The above remarks, paradigmatic of the real space Renormalization Group philosophy, show that the definition of a sensible blocking procedure, in our geometrical setting, consists of a transformation increasing the scale size, realized by passing from the variables $f(k)$ to the variables $\psi_m(k;f)$, (namely by taking the average over all values of f in a larger and larger ball) followed by a rescaling, obtained by dividing $\psi_m(k;f)$ by a suitable power of the number, $N^{(\epsilon_o,m)}$, of elementary ϵ_o-balls contained in the ϵ_m-ball considered, (for random geometries this power is $1/2$).

Following standard usage, and in order to arrive at an interesting geometrical notion of blocking, we assume that for a generic metric g, this rescaling follows by dividing $\psi_m(k;f)$ by $\left[N^{(\epsilon_o,m)}\right]^{\omega_m}$, where ω_m will in general depend on m. Thus, the rescaled blocked variables of relevance are

$$\phi_m(k;f) = \left[N^{(\epsilon_o,m)}\right]^{-\omega_m} \psi_m(k;f). \qquad (3.53)$$

The value of ω_m will be fixed by the requirement that, as $m \to \infty$, and for some (critical) metric g_{crit}, (in general for an open set of such metrics), such normalized large scale block variables have a limiting probability distribution with a finite variance. Namely,

$$\begin{aligned} \lim_{m\to\infty} P_m(\{\phi_m(k;f)\}) &= P_\infty(\{\phi_\infty(f)\}), \\ E_{P_\infty}(\phi_\infty^2(f)) = 1. & \end{aligned} \qquad (3.54)$$

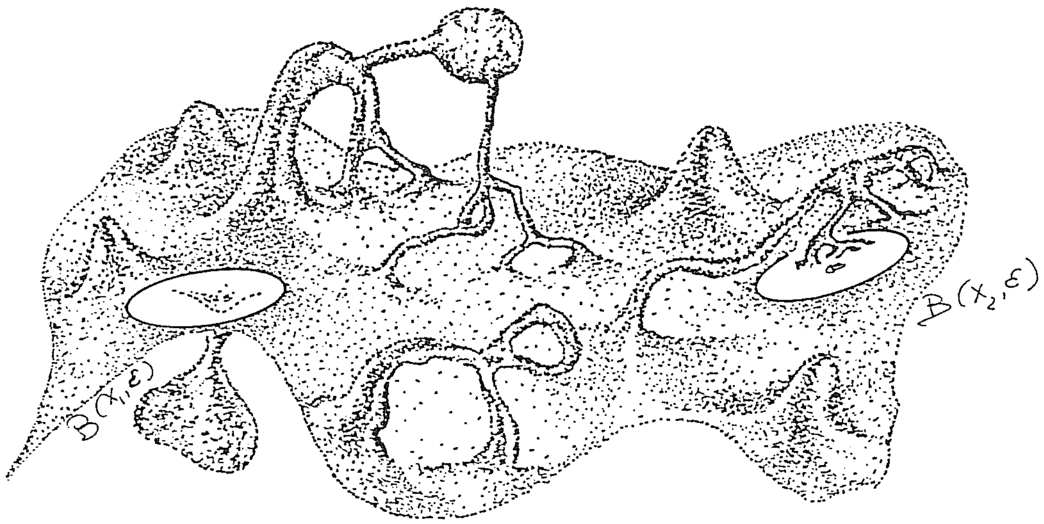

Figure 3.7: Even if the function f is constant over Σ, its average values over the geodesic balls, $B(x_1, \epsilon)$ and $B(x_2, \epsilon)$, are generally uncorrelated owing to the random fluctuations in the geometry of the balls.

Notice that if (Σ, g) is a *nice* manifold, e.g. a constant curvature simply connected three-manifold, we are obviously expecting that the corresponding ω_m is

$$\omega_m(\Sigma, g) = 1. \tag{3.55}$$

In general, we can assume that there is a set of critical metrics and a corresponding ω, such that the above requirements (3.54) are satisfied. Such critical metrics are not necessarily constant curvature metrics and the corresponding ω is not necessarily 1, (one may conjecture that $1/2 \leq \omega \leq 1$, as happens in the Renormalization Group analysis of many magnetic systems). At this stage, this is only a tentative assumption in order to arrive at an interesting concept of geometric Renormalization Group in our setting. Later on, we shall see that such assumptions are justified by exhibiting examples of such non-trivial metrics.

3.6.3 Averaging matter and geometry

Until now, our discussion has addressed mainly geometrical issues and the function f entering the cut-off dependent averaging $< f >_\epsilon$ was not specified. Now we wish to apply the results of the previous paragraph to the averaging of the matter sources, namely, the matter density ρ, the spatial stress tensor s_{ab}, and the momentum density J_a, entering in the phenomenological description of the matter energy-momentum tensor with respect to the instantaneous observers comoving with Σ, *viz.*

$$T_{ab} = \rho u_a u_b + J_a u_b + J_b u_a + s_{ab}, \tag{3.56}$$

where, $\mathbf{u}$ is a unit, future directed normal to the slice Σ, $(u^a u_a = -1)$.

In fact, since in the present epoch the universe is mainly matter dominated, i.e. the pressure can be safely neglected, it would in principle be justified to start our analysis with

$$T_{ab} = \rho u_a u_b + J_a u_b + J_b u_a, \tag{3.57}$$

for the stress-energy tensor. Even the terms involving the momentum density can be eliminated if we pick up a sensible slicing Σ_t, (the comoving frame), provided that the cosmological matter fluid is an irrotational fluid in equilibrium. In any case, wishing to maintain the following discussion to a sufficient degree of generality, without particular restrictive assumptions on the matter sources, we assume in our analysis that the matter energy-momentum tensor has the perfect fluid form

$$T_{ab} = \rho u_a u_b + J_a u_b + J_b u_a + g_{ab} p, \tag{3.58}$$

with a pressure p which is *a priori* not-vanishing, and where g_{ab} is the three-metric of Σ, namely, $g_{ab} = g^{(4)}_{ab} + n_a n_b$, $g^{(4)}_{ab}$ denoting the space-time metric[14].

As argued in the previous paragraphs, smoothing-out the matter sources as described by a set of instantaneous observers (represented by the three-dimensional hypersurface Σ), means eliminating from the distribution of such sources on Σ all fluctuations on scales smaller than the cutoff distance ϵ, leaving an effective probability distribution of fluctuations for the remaining degrees of freedom. The underlying philosophy being that this

[14]The fact that we are adopting a stress tensor with a non-vanishing energy flux can be traced back to the observation that even in an exact $k = 0$ or $k = +1$ FLRW universe there can be a non-zero particle flux tilted with respect to the frame of the fundamental observers, even though T_{ab} has the perfect fluid form. This is possible if the macroscopic quantities (e.g. a particle flux) are derived from kinetic theory with anisotropic distribution function $f(x^i, p^a)$ [90].

effective distribution has the same properties as the original one at distances much larger than ϵ (i.e. for fluctuations with wavelengths much larger than ϵ).

In order to implement this idea along the geometrical lines discussed in section 3.1, we first need to better specify what we mean by the assignment of a collection of instantaneous observers (endowed with clocks) on Σ. Since we are adopting a Hamiltonian point of view, such observers are specified by the assignment, on the (abstract) three-manifold Σ, of the lapse function α and the shift vector field α^i. The former provides the local rate of the coordinate clocks of such observers, while the latter is the three-velocity vector of the observers with respect to the set of instantaneous observers at rest on Σ.

The macroscopic variables of interest characterizing the matter sources are in the present framework, the matter density ρ, and the momentum density $\mathbf{J}$. Actually, it would make more sense to consider the cosmological fluid phenomenologically, described by a pressure p, a baryon number density n_b, energy density ρ, specific entropy s, and temperature T; such variables being related by

$$dp = n_b dh - n_b T ds \tag{3.59}$$

and

$$h = \frac{p+\rho}{n_b}, \tag{3.60}$$

where h is the specific enthalpy. However, for simplicity, we shall in the sequel consider a barotropic fluid.

Thus, the field f characterizing the matter sources, as described by the instantaneous observers on Σ, is given by

$$f \equiv \alpha\rho + \alpha^i J_i. \tag{3.61}$$

Notice that $2\alpha p(\rho)$ is the hamiltonian density of the fluid in the Hamiltonian formulation of Taub's variational principle for relativistic perfect fluids [194]; also we assume that the *dominant energy condition* holds.

The averaging of the matter sources along the lines described in the previous section would then require considering a finite set of instantaneous observers, $\{x_1, \ldots, x_N\}$, located on Σ and setting a standard for the cutoff distance ϵ_o over which the (experimental) distribution of matter sources, (i.e. the probability that the matter variables $\rho(i)$ and $\mathbf{J}(i)$ conform to a given distribution $\rho(i)d\rho$, $\mathbf{J}(i)d\mathbf{J}$), is determined. Then we proceed with the blocking prescription for eliminating unwanted degrees of freedom and consider the behaviour as the averaging regions become larger and larger.

We may decide to treat the riemannian geometry of Σ as uncoupled with the matter sources if we are simply interested in smoothing-out the sources, or if we wish to consider fluctuations of the matter as essentially uncoupled to the fluctuations in geometry. Namely, the curvature fluctuations appearing in the definitions of the block variables, $\psi_m(k; f)$ and $\phi_m(k; f)$, can be thought of, under appropriate circumstances, as independent random variables with a given distribution.

In general, however, the sources are coupled to the gravitational field, and we ought to treat the full dynamical system, the cosmological fluid plus the geometry of Σ, in the procedure of blocking. Moreover, we should bear in mind that without taking explicitly into account the backreaction of the geometry one cannot really provide a reasonable averaging procedure for the sources.

In order to do so, we consider for a given cutoff distance ϵ_o, the variables $\rho(k)$ and $\mathbf{J}(k)$ associated with a minimal geodesic ball covering $\{B(x_k, \epsilon_o)\}_{k=1}^{N}$. The change of the cutoff is naturally realized by considering balls $\{B(x_l, \epsilon_m = (m+1)\epsilon_o)\}$ with $m = 0, 1, \ldots$. The block variables, $\psi_o(k; \rho, \mathbf{J})$ and $\psi_m(k; \rho, \mathbf{J})$, are defined according to (3.43) with f given by (3.61). This transformation can be seen as thinning-out the degrees of freedom which is at the heart of any coarse-graining.

The original cutoff ϵ_o is chosen to set the scale over which General Relativity is experimentally verified and it can be taken as, say, the scale of planetary systems. We can thus safely assume that the Einstein field equations hold on that scale. It is however impossible to provide a mathematical model of the distribution of matter in the universe going down to such fine scales; besides this task would be impractical. What one does instead is to use continuous functions, assuming that they represent appropriately "volume averages". The results of such an averaging in an inhomogeneous medium obviously depend on the scale. The point is that if the Einstein equations hold on the scale where they have been verified (here taken to be that of the planetary scale), then they do not seem to hold *a priori* on larger, cosmological, scales that require averaging. To see this, notice that the Einstein tensor $\tilde{G}_{\mu\nu}$, calculated from an "averaged" metric $\bar{g}_{\mu\nu}$, cannot be equal to the Einstein tensor $\bar{G}_{\mu\nu}$ which was first calculated from the fine-scale metric $g_{\mu\nu}$ and then averaged. This is so due to the non-commutativity of "averaging the metric" with calculating the Einstein tensor being strongly non-linear in the metric components.

Below we shall assume a Hamiltonian point of view. Then the probability $P(\{\rho(k), \mathbf{J}(k)\})$ that the matter variables, according to the records of the instantaneous observers $\{B(x_k, \epsilon); \alpha(x_k), \alpha^i(x$ in Σ, take on some particular set of values, $\{\bar{\rho}(k), \bar{\mathbf{J}}(k)\}$, is given by the equation

$$P(\{\psi_o(k; \rho, \mathbf{J})\}_{1,\dots,N}) = \frac{1}{Z} \exp[-H(\{\rho(k), \mathbf{J}(k)\})], \tag{3.62}$$

where Z is a normalization factor, and $H(\{\rho(k), J(k)\})$ is the Hamiltonian associated with the matter variables ρ and $\mathbf{J}$. Namely,

$$P(\{\psi_o(k; \rho, \mathbf{J})\}_{1,\dots,N}) = \frac{\exp\left[- < \alpha\rho + \alpha^h J_h >_{\epsilon_o}\right]}{\sum_{\rho,J} \exp\left[- < \alpha\rho + \alpha^h J_h >_{\epsilon_o}\right]}, \tag{3.63}$$

where $\sum_{\rho,J} \dots$ denotes summation over all possible configurations of the matter variables, ρ and $\mathbf{J}$, that can be experienced by the instantaneous observers $\{B(x_k, \epsilon); \alpha(x_k), \alpha^i(x_k)\}$ in Σ.

The validity of General Relativity at this scale implies that this Hamiltonian is a part of $\mathcal{H}_{ADM}(\{\rho(k), \mathbf{J}(k)\})$, the Arnowitt-Deser-Misner Hamiltonian, associated with the data of the three-geometry g_{ab} of Σ, its conjugate momentum π_{ab} and of the matter variables ρ and $\mathbf{J}$, evaluated in correspondence with the $\{B(x_k, \epsilon)\}$-approximation associated with the net of points $\{x_k\}$. Namely,

$$\mathcal{H}_{ADM}(\{\rho(k), \mathbf{J}(k)\})|_\epsilon = < \alpha\mathcal{H}(g, \pi, G, \rho) + \alpha^h \mathcal{H}_h(g, \pi, G, \mathbf{J}) >_\epsilon \sum_i vol(B(x_i, \epsilon)), \tag{3.64}$$

where

$$\mathcal{H}(g,\pi,G,\rho) = (det(g))^{-1/2}[\pi^{ab}\pi_{ab} - \frac{1}{2}(\pi^a_a)^2] - (det(g))^{1/2}R(g) + 8\pi(det(g))^{1/2}G\rho,\tag{3.65}$$

and

$$\mathcal{H}_a(g,\pi,G,\mathbf{J}) = -2\pi^b_{a;b} + 16\pi(det(g))^{1/2}GJ_a,\tag{3.66}$$

and where the momentum conjugate to the three-metric g_{ab} is given in terms of the second fundamental form K_{ab} of the embedding of Σ in the resulting space-time, as

$$\pi^{ab} = (det(g))^{1/2}(K^{ab} - K^c_c g^{ab}).\tag{3.67}$$

We wish to recall that the lapse and the shift appear in the Hamiltonian as arbitrary Lagrange multipliers (their evolution is not specified by the equations of motion), and as such they enforce the constraints

$$\mathcal{H}(g,\pi,G,\rho) = (det(g))^{-1/2}[\pi^{ab}\pi_{ab} - \frac{1}{2}(\pi^a_a)^2] - (det(g))^{1/2}R(g) + 8\pi(det(g))^{1/2}G\rho = 0\tag{3.68}$$

and

$$\mathcal{H}_a(g,\pi,G,\mathbf{J}) = -2\pi^b_{a;b} + 16\pi(det(g))^{1/2}GJ_a = 0.\tag{3.69}$$

As is well known, these constraints are related to the invariance of the theory under the (four-dimensional) diffeomorphism group of the space-time resulting from the evolution of the initial data, satisfying them. The momentum constraint, $\mathcal{H}_a(g,\pi,G,\mathbf{J}) = 0$, generates the (spatial) diffeomorphisms into Σ, while the Hamiltonian constraint, $\mathcal{H}(g,\pi,G,\rho) = 0$, generates the deformation of the manifold Σ in the resulting space-time, i.e. the dynamics. (Notice that such deformations can be interpreted as four-dimensional diffeomorphisms only after that space-time has been actually constructed).

The Hamiltonian $\mathcal{H}_{ADM}(\{\rho(k),\mathbf{J}(k)\})$, apart from the lapse and the shift, depends on the three-metric g_{ab} on Σ and its conjugate momentum π_{ab}, the gravitational coupling G, and on the configuration which the matter variables, $\rho(k)$ and $\mathbf{J}(k)$, take on the set of instantaneous observers, $\{B(x_k,\epsilon);\alpha(x_k),\alpha^i(x_k)\}$, chosen to describe the distribution of matter at the given length scale ϵ.

The basic question to understand is how the block transformations, $\{\phi_m(k;\rho,\mathbf{J})\}\rightarrow\{\phi_{m+1}(k;\rho,\mathbf{J})\}$, followed by rescaling, affect the Hamiltonian associated with the matter variables ρ and $\mathbf{J}$, and then discuss how this in turn affects the full Hamiltonian $\mathcal{H}_{ADM}$.

Starting from the probability distribution $P(\{\psi_o(k;\rho,\mathbf{J})\}_{1,\ldots,N})$ we can inductively define the probability distribution $P^{(m+1)}(\{\psi_{m+1}(k;\rho,\mathbf{J})\})$ of the block variables $\{\psi_{m+1}(k;\rho,\mathbf{J})\}$ (and of the corresponding rescaled variables $\{\phi_{m+1}(k;\rho,\mathbf{J})\}$). Since the block variables $\{\psi_{m+1}(k;\rho,\mathbf{J})\}$ are recursively obtained from the knowledge of the block variables at the m^{th} stage, $\{\psi_m(k;\rho,\mathbf{J})\}$, the distribution $P^{(m+1)}(\{\psi_{m+1}(k;\rho,\mathbf{J})\})$ only depends on the knowledge of $P^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})$. We can formally write

$$P^{(m+1)}(\{\psi_{m+1}(k;\rho,\mathbf{J})\})=\sum_{\{\psi_m(k;\rho,\mathbf{J})\}}P^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\}), \tag{3.70}$$

where the sum is over the probabilities of all the configurations $\{\psi_m(k;\rho,\mathbf{J})\}$ consistent with the configuration $\{\psi_{m+1}(k;\rho,\mathbf{J})\}$ of the block variables.

As usual, this allows us to define the effective Hamiltonian for the matter variables after m iterations of the block transformation, according to

$$P^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})\equiv\frac{1}{Z^{(m)}}\exp[-H^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})], \tag{3.71}$$

where

$$Z^{(m)}\equiv\sum_{\{\psi_m(k;\rho,\mathbf{J})\}}\exp[-H^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})]. \tag{3.72}$$

Such $H^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})$ are defined up to an additive constant term (e.g. [31]). If we also stipulate, as is standard usage in the Renormalization Group approach, that the effective Hamiltonian $H^{(m+1)}(\{\psi_{m+1}(k;\rho,\mathbf{J})\})$ takes the same functional form as $H^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})$, i.e., if

$$P^{m+1}(\{\psi_{m+1}(k;\rho,\mathbf{J})\}_{1,\ldots,N})=\frac{\exp\left[-<\alpha\rho^{(m+1)}+\alpha^h J_h^{(m+1)}>_{\epsilon_{m+1}}\right]}{\sum_{\rho,J}\exp\left[-<\alpha\rho^{(m+1)}+\alpha^h J_h^{(m+1)}>_{\epsilon_{m+1}}\right]}, \tag{3.73}$$

then this indeterminacy can be transferred to the *effective* matter variables $\rho^{(m+1)}$ and $\mathbf{J}^{(m+1)}$, in terms of which $H^{(m+1)}(\{\psi_{m+1}(k;\rho,\mathbf{J})\})$ is defined.

It is immediately checked that such effective matter variables are defined by (3.71) up to the transformations

$$\begin{aligned} \rho^{(m+1)} &\to \rho^{(m+1)} + f, \\ \mathbf{J}^{(m+1)} &\to \mathbf{J}^{(m+1)} + \mathbf{v}, \end{aligned} \tag{3.74}$$

where f and $\mathbf{v}$ are, respectively, a scalar function (sufficiently regular) and a vector field defined on Σ.

Among the various normalization conditions that we may adopt in order to avoid the indeterminacy connected with (3.74), the natural one comes about *by requiring the divergence and the Hamiltonian constraints to hold at each stage of the renormalization.* This being the case, upon coupling matter to the geometry, the lapse α and the shift α^i maintain their rôle of Lagrange multipliers enforcing the (four-dimensional) diffeomorphism invariance of the theory.

Such requirements imply that the full effective Hamiltonian (matter plus geometry), takes on the standard ADM form pertaining to gravity interacting with a barotropic fluid at every stage of the renormalization.

According to the remarks above, we can consider as an independent parameter in the (effective) fluid Hamiltonian $H^{(m)}(\{\psi_m(k;\rho,\mathbf{J})\})$ the three-metric $g_{ab}^{(m)}$ of the three-manifold Σ, whereas the matter density $\rho^{(m)}$ and the current density $\mathbf{J}^{(m)}$ are at each stage connected to $g_{ab}^{(m)}$ and $K_{ab}^{(m)}$ by the Hamiltonian and divergence constraints that hold at each stage. Then the effect of the renormalization induced by the blocking procedure $\{\psi_m(k;\rho,\mathbf{J})\} \to \{\psi_{m+1}(k;\rho,\mathbf{J})\}$ and the corresponding rescaling, can be symbolized as a non-linear operation acting on the metric $(g_{ab}^{(m)})$ so as to produce the metric $(g_{ab}^{(m+1)})$, i.e.

$$(g_{ab}^{(m+1)}) = \mathcal{R}(g_{ab}^{(m)}), \tag{3.75}$$

whereas the renormalization of the second fundamental form K_{ab} is generated by the *linearization* of (3.75), as is discussed in section 5.

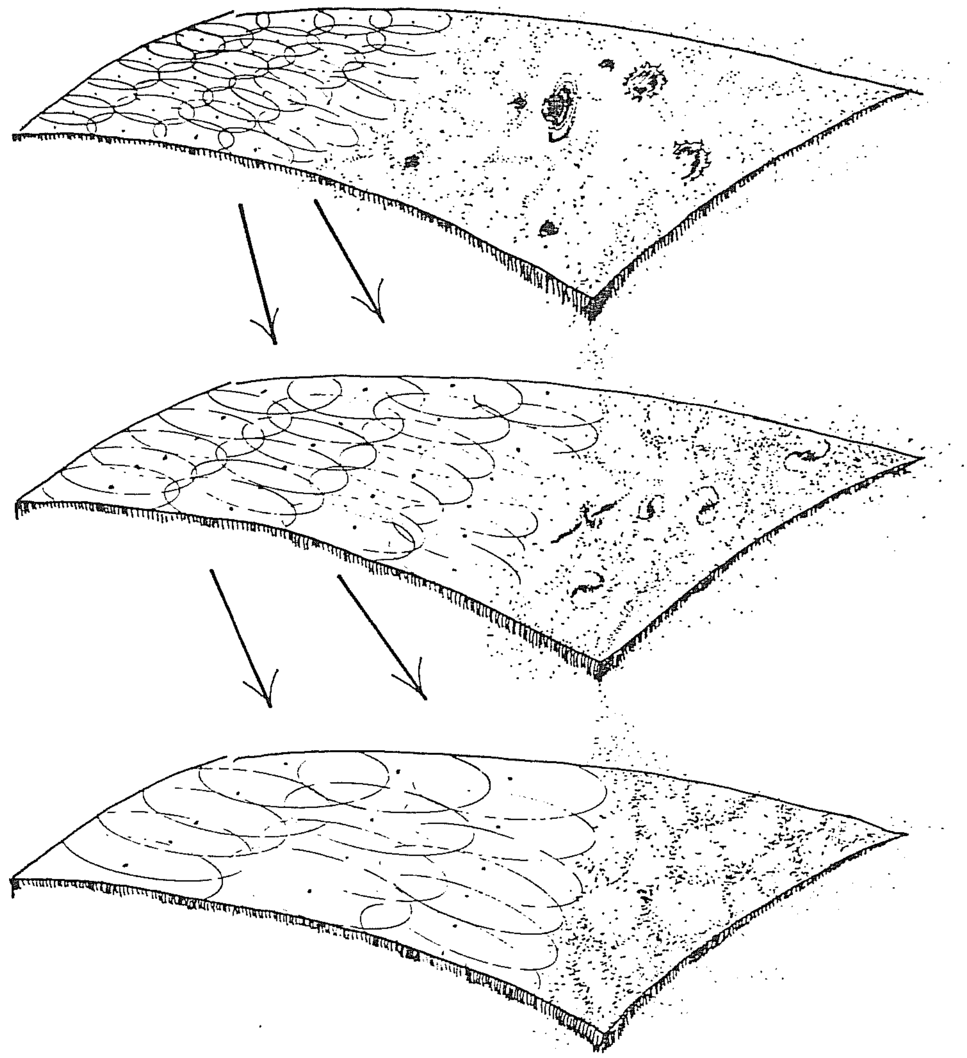

Figure 3.8: Given a probability law according to which matter is distributed at a given length scale (say, the planetary scale), we can get the corresponding probability distributions obtained by averaging the matter variables over regions of ever increasing scales. The resulting *effective* Hamiltonians are defined up to additive constants which affect the renormalized mass density and the renormalized momentum density. Such indeterminacy can be naturally removed by enforcing the Hamiltonian and diverence constraints at each step of the renormalization procedure.

This way, this renormalization transformation $\mathcal{R}$ defines a trajectory in the space of riemannian metrics of Σ. One moves on such trajectory by discrete jumps. However, in what follows we shall replace such discrete dynamical system by a smooth dynamical system, describing renormalizations of the parameters (g_{ab}), and (K_{ab}).

The deformation of the initial data set for the Einstein field equations, symbolically denoted by $\mathcal{R}$ above in (3.75), realizes in fact a *formal*, (at least at this stage), mapping between the initial data sets for the field equations. As we have seen above, this renormalization acts in such a way that at each its step the deformed data satisfy the constraints. The time evolution, in turn, of any such data set generates a one-parameter family of solutions to the field equations which interpolates between the initial inhomogeneous space-time to be averaged-out and its renormalized (deformed) counterpart.

3.6.3.1 The Ricci–Hamilton flow

In order to replace the discrete operator $\mathcal{R}$, describing the effect of renormalization of g_{ab} with a continuous flow, we can start by discussing some geometrical implications of equation (3.40). According to the Renormalization Group analysis of the previous section, they follow by considering the average $< f >_\epsilon$ as a functional of the metric and thinking of the metric g_{ab} as a running coupling constant, depending on the cut-off. In this connection, it can be verified that we can equivalently interpret (3.40) as obtained by considering the variation of $< f >_\epsilon$ under a suitable smooth deformation of the background metric g_{ab}, rather than by deforming the (euclidean) radius of the balls $\{B(x_i, \epsilon)\}$.

As a matter of fact, we can equivalently rewrite the second term on the right hand side of (3.40) as

$$< R >_\epsilon < f >_\epsilon - < Rf >_\epsilon = -D < f >_\epsilon \cdot \frac{\partial g_{ab}}{\partial \eta}, \tag{3.76}$$

where, $D < f >_\epsilon \cdot \partial g_{ab}/\partial \eta$ denotes the formal linearization of the functional $< f >_\epsilon$ in the direction of the symmetric 2−tensor $\partial g_{ab}/\partial \eta$, and where

$$\frac{\partial g_{ab}(\eta)}{\partial \eta} = \frac{2}{3} < R(\eta) > g_{ab}(\eta) - 2R_{ab}(\eta), \tag{3.77}$$

$R_{ab}(\eta)$ being the components of the Ricci tensor $Ric(g(\eta))$, and $< R(\eta) >$ is the average scalar curvature given by

$$< R(\eta) >= \frac{1}{vol(\Sigma)} \int_{\Sigma} R(\eta) d\mu_{\eta}. \tag{3.78}$$

Indeed, the linearization of $< f >_{\epsilon}$ in the direction of the generic 2−tensor $\partial g_{ab}/\partial\eta$ is provided by

$$D < f >_{\epsilon} \cdot \frac{\partial g_{ab}}{\partial\eta} = \frac{1}{2} < f g^{ab} \frac{\partial}{\partial\eta} g_{ab} >_{\epsilon} - \frac{1}{2} < f >_{\epsilon} < g^{ab} \frac{\partial}{\partial\eta} g_{ab} >_{\epsilon}, \tag{3.79}$$

so (3.76) follows, given the expression (3.77) for $\partial g_{ab}/\partial\eta$.

According to what has been said in the previous paragraphs, the effective distribution of matter sources, according to a set of instantaneous observers $\{B(x_i, \epsilon)\}$, is characterized by the underlying three-geometry thought of as an effective parameter depending on the cutoff ϵ. Since the Hamiltonian and the divergence constraints hold at each stage of the renormalization procedure, (they fix the effective Hamiltonians which otherwise are undetermined up to a constant factor), the renormalization of the matter fields is intrinsically tied with the renormalization of the three-metric.

The invariance of the long distance properties of the matter distribution, under simultaneous change of the cutoff ϵ and the parameter g_{ab}, can be expressed as a differential equation for the effective Hamiltonian $\mathcal{H}(\rho, \mathbf{J})$, (actually for the partition function associated with this effective Hamiltonian), namely,

$$[-\epsilon\frac{\partial}{\partial\epsilon} + \beta_{ab}(g)\frac{\partial}{\partial g_{ab}}] \sum_{\rho, J} \exp[-\mathcal{H}(\rho, \mathbf{J})] = 0. \tag{3.80}$$

Recall that $\mathcal{H}(\rho, \mathbf{J})$ is explicitly provided by (in the given approximation)

$$\mathcal{H}(\rho, \mathbf{J}) =< \alpha p(\rho) + \alpha^i J_i >_{\epsilon} \sum_h vol(B(x_h, \epsilon)), \tag{3.81}$$

where the average $< \alpha p(\rho) + \alpha^i J_i >_{\epsilon}$ is a functional of the three-metric g_{ab} here thought of as the running coupling constant.

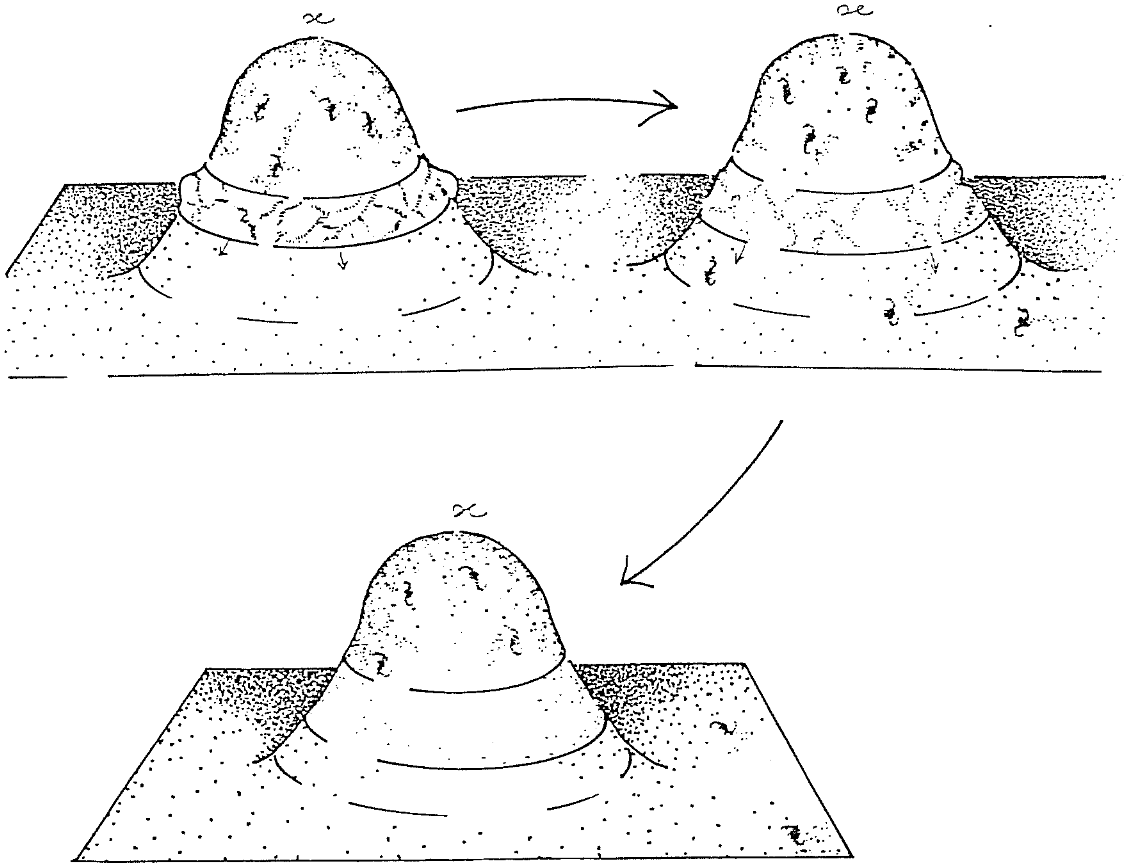

Figure 3.9: If we enlarge the ball from ϵ to $e^{\eta}\epsilon$, while deforming the metric according to the Ricci-Hamilton flow (3.77), (for a *parameter time* $\eta = \gamma$), then the average value remains *as constant as it can be* since (3.77) smoothes the curvature fluctuations in the annulus $B(x, e^{\gamma}\epsilon)\backslash B(x,\epsilon)$.

Thus, in order that the equation (3.80) is satisfied, it is *sufficient* that

$$[-\epsilon\frac{\partial}{\partial\epsilon} + \beta_{ab}(g)\frac{\partial}{\partial g_{ab}}] < \alpha p(\rho) + \alpha^i J_i >_{\epsilon}= 0. \tag{3.82}$$

The Renormalization Group equation (3.82) states that increasing the cutoff length (i.e., the radius of the averaging balls) from ϵ to $e^{\gamma}\epsilon$, while deforming the metric g_{ab} by flowing along the beta function $\beta_{ab}(g)$ for a parameter-time γ, has no net effect on the long distance properties of the considered system.

According to the above remarks and equation (20), it can be verified that the beta function yielding for (3.82), defined by

$$\epsilon\frac{\partial}{\partial\epsilon}g_{ab} = \beta_{ab}(g), \tag{3.83}$$

is exactly provided by (3.77), namely,

$$\beta_{ab}(g) = \frac{\partial g_{ab}(\eta)}{\partial \eta} = \frac{2}{3} < R(\eta) > g_{ab}(\eta) - 2R_{ab}(\eta), \tag{3.84}$$

where the parameter η is the logarithmic change of the cutoff length ϵ.

Since the manifold Σ is compact, (3.77) has to be interpreted as a Renormalization Group equation in a finite geometry, and thus the relevant phenomena are here related to finite size scaling (see e.g. [50]). A continuous theory, describing the (universal) properties of the cosmological sources and of the corresponding geometry, may arise when the correlation length associated with the distribution of cosmological matter is of the order of the size of the underlying manifold.

The metric flow (3.77) is known as the Ricci-Hamilton flow [124], studied in connection with the quasi-parabolic flows on manifolds; quite independently it has been discussed in investigating the Renormalization Group flow for general σ-models (see e.g. [175] and references quoted therein). The Ricci-Hamilton flow is always solvable [124] for sufficiently small η and has a number of useful properties, (apart from being volume preserving which is simply a consequence of the normalization chosen), namely, any symmetries of $g_{ab}(\eta_o)$ are preserved along the $g_{ab}(\eta)$ flow for all $\eta > \eta_o$, and the limiting metric (if attained) $\overline{g}_{ab} = \lim_{\eta \to \infty} g_{ab}(\eta)$ has constant sectional curvature. Thus equation (3.77), with the initial condition $g_{ab}(0) = g_{ab}$, defines (when globally solvable) a smooth family of deformations of the initial three-manifold, deforming it into a three-space of constant curvature.

3.6.3.2 Fixed points and basins of attraction

Thus the point of the above discussion is that in order to arrive at a *fixed point* of the RG equation (3.40), the geometry has to be deformed according to the Ricci-Hamilton flow (3.77). In this setting of the problem, the Ricci-Hamilton equations appear naturally and in fact the approach proposed, enables us to attach a physical meaning to them within the coarse-graining picture. This element was lacking in [51] where Hamilton's theorem appeared rather *ad hoc*. On the other hand, our approach demonstrates that the smoothing issue is deeply connected with the geometry and exhibits how this relationship works. First some general remarks.

A fixed point is a point in the coupling constants space that satisfies

$$g^{\star}_{ab} = \mathcal{R}(g^{\star}_{ab}), \tag{3.85}$$

i.e. it is mapped onto itself by RG transformation.

Under RG transformation length scales are reduced by a factor $m+1$. Namely, for the block variables, the correlation length measured in units appropriate for them, $L_m(\epsilon_m)$, is smaller than the correlation length L_o of the initial system measured in units of ϵ_o. The actual physical value of the correlation length L is of course unchanged by the process of blocking, thus $L = L_m(\epsilon_m) = L_o\epsilon_o$, so

$$L_m = \frac{L_o}{m+1}. \tag{3.86}$$

Since $L_m < L_o$, the system with Hamiltonian $\mathcal{H}^{(m)}$ must be further from criticality than the original system. Thus, we conclude that the system is at a new effective reduced "temperature" $g^{(m)}_{ab}$. At a fixed point (3.85), $L^{\star}$ can be only zero or infinity since then $L^{\star} = L^{\star}/(m+1)$.

As is standard in RG analysis, we will refer to a fixed point with $L = \infty$ as a *critical fixed point*, when $L = 0$ we will call it *trivial*. Each fixed point has its domain or basin of attraction, namely, the points in the coupling constants space in such a basin necessarily flow towards and end up at the fixed point, after an infinite number of iterations of RG transformation.

Let us, for the purpose of clarity, employ for a moment the ferromagnetic analogy. In this case, for a system exhibiting a phase transition, there are two attractive fixed points. One is the high-temperature fixed point which attracts each point with $T > T_{crit}$ in the coupling constants space, and it corresponds to the effective Hamiltonian for the system as $T \to \infty$. In this phase the variables assume random values and are uncorrelated. Upon a sensible blocking of such a system the probability distribution of the block variables remains unchanged.

The second fixed point is the low-temperature fixed point which is the effective Hamiltonian for the system when $T \to 0$. This corresponds to a system in a complete spins alinement and the block variables are ordered then. Every point corresponding to $T < T_{crit}$ is eventually attracted to this fixed point.

Across the Hamiltonian space there should exist then a surface, the so-called *critical surface*, which separates the effective Hamiltonians flowing to the high-temperature fixed point from those flowing to the low-temperature fixed point. Notice that the word *surface* here has rather a heuristic meaning. Indeed, the *set* of metrics, separating those flowing towards a "low-temperature" fixed points from those flowing towards a "high-temperature" fixed point, has quite a complex structure whose understanding is deeply connected with some, yet unsolved, conjectures in $3D$-manifold topology, (see below for more details).

If we now choose to start with a point on the critical surface, then upon RG transformation it will stay within the critical surface. There is a possibility that, as the number of iterations of RG goes to infinity, the Hamiltonian will tend to a finite limit $\mathcal{H}^{\star}$. This point is the critical fixed point and within the critical surface it is attractive (this is roughly speaking the basic mechanism for universality), along the direction out of the critical surface it is repulsive. This fixed point is related to the singular critical behaviour of the system due to the fact that all points in its basin of attraction have infinite correlation length. The simplest case is when the fixed points are isolated points, but it is also possible to have lines or surfaces of fixed points.

With these preliminary remarks along the way, let us discuss how some of the above general characteristic of Renormalization Group flows find their proper counterpart in the particular situation we are analysing.

In the previous paragraph we showed that it was possible to replace the discrete operation of increasing the scale size of the observational averaging region by a "transformation", smoothly deforming the background metric g_{ab}, which turned out to be the Ricci-Hamilton flow. In this setting, following the example above, we would like to adopt the fundamental hypothesis linking RG to the critical phenomena, namely, the existence of a "critical" metric (on the critical surface) g_{ab}^{crit}, and of a "fixed point" metric $g_{ab}^{\star}$, such that

$$\lim_{m\to\infty} \mathcal{R}^{(m)}(g_{ab}^{crit}) = g_{ab}^{\star}. \tag{3.87}$$

In (3.87) $g_{ab}^{\star}$ is a mathematical object invariant under RG and we assume that g_{ab}^{crit} represents the physics (in a sense to be clarified further) of curvature fluctuations of a manifold at its critical point ("critical manifold"). Since the Ricci-Hamilton flow can be interpreted as a dynamical system on a set of closed riemannian manifolds, we can adopt the following interpretation of (3.87). Suppose, that we look at our manifold through a "microscope" and are able to discern the curvature fluctuations down to a size ϵ_m. $\mathcal{R}^{(m)}$ then represents the operation of decreasing the magnification factor, by m say, i.e. the sample seen appears to shrink by this factor. We have to assume that the system is sufficiently large so that the edges of the sample will not appear in the view. The hypothesis (3.87) states that if we decrease the magnification by a sufficiently large amount, we shall not see any change if we decrease it even further.

We are going to describe to what extent such hypothesis holds.

We already said that the Ricci-Hamilton flow (3.77), while always solvable for sufficiently small η [124], may not yield for a non-singular solution as η increases. Hamilton noticed that there are patterns in the kind of singularities that may develop. Typically the curvature blows up, but in a very regular way (e.g. for $S^1 \times S^2$ with the standard symmetric metric). This has led him to a research program which, roughly speaking, amounts to saying that any three-manifold can be decomposed into pieces on which the Ricci-Hamilton flow is global and thereby, each of these pieces can be smoothly deformed into a locally homogeneous three-manifold. Singularities may develop in the regions connecting the smooth-able pieces, but such singularities should be of a finite number of types and all of a rather symmetric nature (namely, if they are blown up, they should be associated with symmetric manifolds as e.g. $S^1 \times S^2$ [56, 142]).

It may be said that Hamilton's program is an analytic approach to prove Thurston's conjecture, which claims that any closed three-manifold can be cut into pieces such that each of them admits a locally homogeneous geometry (e.g. [246]). The rationale, underlying this *analytical* program towards Thurston's geometrization conjecture, lies in the above nice structural properties of the Ricci-Hamilton flow. Several steps are involved in this program. Let us briefly recall them since, even if some of them are yet unproven, they shed light on the assumption (3.87) and on (3.77) when interpreted as a Renormalization Group flow.

The *first step* is the assignment of an *arbitrary* metric g on the three-manifold Σ. In the Renormalization Group approach, this corresponds to picking up a vastly inhomogeneous and anisotropic geometry for the physical space, (here equivalent to the *high-temperature phase*). Such a choice may not conform to the actual *quasi-homogeneous* three-geometry of the physical space as is experienced now. However this quasi-homogeneity, in our opinion, may be related to the possibility that the actual universe is near criticality, a circumstance that we want to discuss rather than assume from the outset.

The *second step* is to deform this metric g via the Ricci-Hamilton flow (3.77). In general, this flow develops local singularities which should be related to the manifold decomposition in Thurston's conjecture. Away from each of the local singularities it is conjectured (but not yet proved) that the Ricci-Hamilton flow approaches that of a locally homogeneous geometry in each disconnected piece.

This picture may be consistent with our blocking procedure yielding for (3.77) as the Renormalization Group flow. According to the analysis carried out in section 3.1, a highly inhomogeneous and anisotropic geometry (Σ, g) can be characterized by minimal geodesic ball coverings $\{B(x_i, \epsilon_o)\}$, whose balls are to a good approximation largely uncorrelated when seen on a suitable scale. This means that the values of the scalar curvatures, $R(i)$, evaluated at the centers of the balls of the covering, are random variables. Upon enlarging the balls and rescaling, certain regions of the manifold might be such that correlations develop among the corresponding $R(i)$, whereas in other regions, no matter how we enlarge the balls and rescale, the $R(i)$ will remain independently distributed random variables. The former regions, exhibiting a persistence length, are those that should approach a locally homogeneous geometry under blocking and rescaling (i.e., under the flow (3.77)). The latter regions should rather develop singularities under (3.77), since no matter how much we block and rescale, the curvature will maintain its *white noise* character.

The *third step* is to study the behavior of (3.77) for the locally homogeneous geometries, for this accounts for the structure of the critical set of (3.77). This has been accomplished [142], and one finds that, depending on the initial locally homogeneous geometry, the Ricci-Hamilton flows either, *(i)* converges to a constant curvature metric, *(ii)* asymptotically approaches (as $\eta \to \infty$) a flat degenerate geometry, of either two or one dimensions (*pancake* or *cigar* degeneracy), with the curvature decaying at the rate $1/\eta$, or, *(iii)* hits a curvature singularity in finite time, with this singularity being that of the Ricci-Hamilton flow for the standard metric on $S^2 \times S^1$. Note that constant curvature geometries always occur whenever the manifold can support them, (in dimension three, constant curvature manifolds and Einstein manifolds are synonyms). It is also quite interesting to note that the Ricci-Hamilton flow of homogeneous metrics usually (with a few exceptions) tends to approach or converge to the maximally symmetric homogeneous metric in the class considered (see [142] for details).

Interpreting (3.77) as the Renormalization Group flow, it follows that locally homogeneous geometries evolving under (3.77) towards an isolated constant curvature manifold, are sinks describing a stable phase of the corresponding cosmological model. For instance a FLRW model (with closed spatial sections) characterizes such a phase. Non-isolated constant curvature manifolds, (e.g. flat three-tori) provide less trivial examples of limiting behaviour of (3.77), (see e.g. [125, 56]). The locally homogeneous manifolds non-admitting Einstein manifolds (i.e. there are no left-invariant Einstein metrics on the group $SL(2, \mathbb{R})$), provide even more interesting behavior. In this case, a metric renormalized under the action of (3.77) develops degeneracies, and one gets, in our setting an *effective* cosmological model with spatial sections of lower dimensionality.

All these limit points of (3.77), either fixed or not, have their own basins of attraction. Rigorously speaking, they are the sets of three-metrics flowing under (3.77) to the respective limit points, briefly discussed above. There are nine such basins of attraction, corresponding to the nine classes of homogeneous geometries that can be used to model (*by passing to the universal cover*) the local inhomogeneous geometries on closed three-manifolds. By labelling these classes according to the minimal isometry group of the geometries considered we distinguish the following basins (here we follow closely the exposition in [142]):

(i): The $\mathbb{R}^3$*-Basin.* It contains all three-metrics flowing towards the homogeneous flat $\mathbb{R}^3$ metrics. This basin is eventually attracted by flat space, (flat tori, when reverting to the original manifold rather than to its universal cover).

(ii): The $SU(2)$*-Basin.* It contains all three-metrics flowing towards the three-parameter family of homogeneous $SU(2)$ metrics. This class admits Einstein metrics, in particular the round metrics on the three sphere. This basin is exponentially attracted to the round three-sphere, (modulo identifications). It is the basin of attraction yielding for closed FLRW cosmological models.

(iii): The $SL(2, \mathbb{R})$*-Basin.* It contains all three-metrics flowing towards the three-parameter family of homogeneous $SL(2, \mathbb{R})$ metrics. This class does not admit Einstein metrics. This basin goes degenerate, yielding for a *pancake* degeneracy whereby a two-dimensional geometry survives: two of the components of the metric increase without bound while the other shrinks to zero.

(iv): The Heisenberg-Basin. It contains all three-metrics flowing towards the three-parameter family of homogeneous Nil-metrics. Again, this class does not contain any Einstein metrics. This basin too undergoes a pancake degeneracy.

(v): The $E(1,1)$*-Basin*, where $E(1,1)$ is the group of isometries of the plane with flat Lorentz metric. It contains all three-metrics flowing towards the three-parameter family of homogeneous Solv-metrics. Also this basin fails to contain Einstein metrics. This basin eventually exhibits a *cigar* degeneracy: the curvature dies away, and while one diameter expands without bound, the other two diameters shrink to zero.

(vi): The $E(2)$-Basin, where $E(2)$ is the group of isometries of the euclidean plane. It contains all three-metrics flowing towards the three-parameters family of homogeneous Solv-metrics containing the flat geometry. This basin is eventually attracted by flat metrics.

(vii): The $H(3)$-Basin, where $H(3)$ is the group of isometries of hyperbolic three-space. It contains all three-metrics flowing towards the one-parameter family of homogeneous metrics constant multiples of the standard hyperbolic metric. This basin is attracted to hyperbolic space.

(viii): The $SO(3) \times \mathbb{R}^1$-Basin. It contains all three-metrics flowing towards the two-parameter family of homogeneous metrics obtained by rescaling the standard product metric on $S^2 \times \mathbb{R}^1$. It does not contain Einstein metrics. This is a singular basin, it is attracted towards a curvature singularity: the round two-sphere shrinks, while the scale on $\mathbb{R}^1$, (or if you prefer, the S^1 factor in the original manifold), expands.

(ix): The $H^2 \times \mathbb{R}^1$-Basin, where $H(2)$ is the group of isometries of the hyperbolic plane. It contains all three-metrics flowing towards the two-parameter family of homogeneous metrics obtained by rescaling the product metric on the product manifold, $\mathbb{R}^1 \times H^2$. Again, this basin does not contain Einstein manifolds, and it is attracted towards a pancake degeneracy.

The basins of attraction just described, and in particular those yielding for fixed points (Einstein manifolds), are relatively uninteresting in connection with the Renormalization Group interpretation of (3.77). Such fixed points, e.g. the round three-sphere, or the flat three-tori, are all totally attractive. As already recalled they can be thought of as distinct stable phases yielding for distinct cosmological models, characterized by the nature of the metric (and of its infinitesimal deformations) at the fixed point. For instance, the $SU(2)$-Basin characterizes FLRW models with closed spatial sections, and the related homogeneous anisotropic models, (see section 6 for more details).

3.6.3.3 Critical fixed points: an example

The existence of critical fixed points, (critical in the sense of (3.87)), characterizing the (universal) properties of (continuous) phase transitions between two different cosmological regimes, cannot be immediately read off from the above analysis. It is rather the consequence of the (conjectured) existence of the decomposition of a manifold into pieces that is associated, according to Hamilton, to the local singularities of the flow (3.77), (see the *second step* of Hamilton-Thurston geometrization program). The origin of this connection between critical behavior of (3.77) in the sense of Renormalization Group, and Hamilton's program, can be seen by considering the following detailed example.

Let us assume that topologically Σ is a three-sphere, $\Sigma \simeq S^3$. We shall consider on Σ a metric g_1 obtained by glueing through a smooth connected sum two copies of a round three-sphere

$$\Sigma \simeq S^3_{(1)} \sharp S^3_{(2)}, \tag{3.88}$$

and endowing each $S^3_{(i)}$, $i = 1, 2$, factor with a round metric of volume $v_1 = 1$, and the joining tube

$$S^2 \times ([0,1] \subset \mathbb{R}^1), \tag{3.89}$$

with the standard product metric of volume const.$e^{-3\tau}$, (τ is a suitable parameter, see below), for $\tau \gg 1$.

To explicitly construct this latter metric we can proceed as follows:

Let y_1 and y_2 respectively denote two chosen points in both factor copies, $S^3_{(1)}$ and $S^3_{(2)}$. Let $h_{(i)}: \mathbb{R}^3 \to S^3_{(i)}$, $i = 1, 2$, be two imbeddings given by the exponential mappings

$$\begin{aligned} \exp_{y_1} &: T_{y_1} S^3_{(1)} \simeq \mathbb{R}^3 \to S^3_{(1)}, \\ \exp_{y_2} &: T_{y_2} S^3_{(2)} \simeq \mathbb{R}^3 \to S^3_{(2)}. \end{aligned} \tag{3.90}$$

We assume that $h_{(1)}$ preserves the orientation while $h_{(2)}$ reverses it.

Let $\alpha: (0, \infty) \to (0, \infty)$ denote an orientation reversing diffeomorphism, and define $\alpha_3: \mathbb{R}^3/\{0\} \to \mathbb{R}^3/\{0\}$ by the map $\alpha_3(v) = \alpha(|v|)\frac{v}{|v|}$ for every vector $v \in \mathbb{R}^3/\{0\}$. For every point $x_1 = h_1(v)$ in the geodesic ball $B(y_1, r)/\{0\} \subset S^3_{(1)}$, with $0 < r \leq \frac{\pi}{2}$, (with $\pi/2$ the injectivity radius of the unit three-sphere $S^3_{(i)}$), we identify $h_1(v)$ with $h_2(\alpha_3(v)) \in B(y_2, r)/\{0\} \subset S^3_{(2)}$.

The space obtained in this way

$$S^3_{(1)} \# S^3_{(2)} = (S^3_{(1)} - \{y_1\}) \cup_{h_{(2)}\alpha_3 h^{-1}_{(1)}} (S^3_{(2)} - \{y_2\}) \tag{3.91}$$

is a particular realization of the connected sum [159] of two copies of unit three-spheres.

In order to give to the neck, joining the two $S^3_{(i)}$, a cylindrical shape we blow up [103] the metrics of the three-spheres in the neighborhoods of the points y_1 and y_2. Consider, for simplicity, only the $S^3_{(1)}$ factor since the argument goes in an analogous way also for the $S^3_{(2)}$ factor.

Exploiting the exponential mapping the round metric of $S^3_{(1)}$, (actually *any* sufficiently smooth metric), can be written in a neighborhood of y_1 in *geodesic polar coordinates* as

$$g(x) = dr^2 + r^2 h_{ij} d\theta^i d\theta^j + \mathcal{O}(r^4), \tag{3.92}$$

where $r = \mathrm{dist}(x, y_1)$ is the distance between y_1 and the point considered, $h_{ij} d\theta^i d\theta^j$ is the metric on the two-dimensional unit sphere S^2, and as usual, the higher order correction terms involve the curvature.

Now, if we blow up this metric by rescaling it through r^2 we get, up to curvature corrections, a distorted cylindrical metric

$$\tilde{g}(x) = \frac{g(x)}{r^2} = \frac{dr^2}{r^2} + h_{ij} d\theta^i d\theta^j + \mathcal{O}(r^2). \tag{3.93}$$

In order to eliminate the axial distortion due to $\frac{dr^2}{r^2}$, we substitute for the radial variable r, introducing a new coordinate τ defined as

$$r = \exp[-\tau]. \tag{3.94}$$

When expressed in terms of τ we get that the blown up metric $\tilde{g}(x)$ reduces to

$$\tilde{g}(x) = d\tau^2 + h_{ij} d\theta^i d\theta^j + \mathcal{O}(e^{-2\tau}). \tag{3.95}$$

Thus, the blown up metric approaches the cylindrical metric exponentially fast as $\tau \to \infty$. In order to introduce smoothly such a metric on the neck of $S^3_{(1)} \sharp S^3_{(2)}$ we can proceed as follows.

Choose smooth functions $\delta_{(i)}(r)$, $i = 1, 2$, satisfying

$$\begin{aligned} \delta_{(i)}(r_i) &= \frac{1}{r_i^2}, \ 0 < r_i \leq \frac{\pi}{4}, \\ \delta_{(i)}(r_i) &= C^\infty, \text{ decreasing}, \ \frac{\pi}{4} \leq r_i \leq \frac{\pi}{2}, \\ \delta_{(i)}(r_i) &= 1, \ r_i \geq \frac{\pi}{2}, \end{aligned} \tag{3.96}$$

where $r_i = \mathrm{dist}(x, y_{(i)})$. Also, let us introduce the following inclusion maps

$$\begin{aligned} \chi_{(1)} &: S^3_{(1)} - \{y_1\} \hookrightarrow (S^3_{(1)} - \{y_1\}) \cup_{h_{(2)}\alpha_3 h^{-1}_{(1)}} (S^3_{(2)} - \{y_2\}), \\ \chi_{(2)} &: S^3_{(2)} - \{y_2\} \hookrightarrow (S^3_{(1)} - \{y_1\}) \cup_{h_{(2)}\alpha_3 h^{-1}_{(1)}} (S^3_{(2)} - \{y_2\}), \\ \chi_{(3)} &: h_1(T_{y_1} S^3_{(1)} - \{0\}) \cup h_2(T_{y_2} S^3_{(2)} - \{0\}) \hookrightarrow (S^3_{(1)} - \{y_1\}) \cup_{h_{(2)}\alpha_3 h^{-1}_{(1)}} (S^3_{(2)} - \{y_2\}), \end{aligned} \tag{3.97}$$

($\chi_{(1)}$ and $\chi_{(2)}$ denote the inclusions of the spheres, minus the points $y_{(i)}$, into the connected sum; $\chi_{(3)}$ is the inclusion of the neck).

If $\{\xi_\alpha\}$ is a partition of unity associated with the covering corresponding to the above inclusions $\chi_{(\alpha)}$, with $\alpha = 1, 2, 3$, we can then define the following metric on $S^3_{(1)} \sharp S^3_{(2)}$,

$$\tilde{g}(x) = \sum_\alpha \xi_\alpha \chi^*_{(\alpha)} [g(x) \cdot \delta_{(i)}(\mathrm{dist}(x, y_{(i)}))]. \tag{3.98}$$

For x not in the geodesic balls (of radius $\pi/2$), centered on y_1 and y_2, this is the standard round metric on the unit three-sphere; for x in the geodesic balls of radius $\pi/4$, centered on y_1 and y_2, this is, up to curvature corrections, the cylindrical metric introduced above. For x in the annuli between such balls $\tilde{g}$ is a smooth interpolating metric joining the spheres to the cylindrical neck.

By expressing (3.98) in terms of the variable τ, we get the metric on $S^3_{(1)} \sharp S^3_{(2)}$ which is the round metric on each $S^3_{(i)}/B(y_i, e^{-\tau})$ factor, and these factors are connected, for τ large enough, (i.e., nearby y_i), by a thin flat cylinder.

The Ricci-Hamilton evolution of $(S^3_{(1)}\sharp S^3_{(2)}, \tilde{g})$ can be explicitly constructed as follows.

According to [142] let us write the metric on $S^2 \times \mathbb{R}^1$ as

$$\tilde{g}|_{\text{neck of } S^3_{(1)}\sharp S^3_{(2)}} = D g_{\mathbb{R}^1} + E g_{S^2}, \tag{3.99}$$

where $g_{\mathbb{R}^1}$ is the metric on $\mathbb{R}^1$, g_{S^2} is the round metric on the unit two-sphere, and D and E are constants. The Ricci-Hamilton flow (3.77) preserves the structure of this metric and reduces to the coupled system of ordinary differential equations

$$\begin{aligned} \frac{d}{d\eta}E &= -\frac{2}{3}, \\ \frac{d}{d\eta}D &= \frac{4}{3}\left(\frac{D}{E}\right), \end{aligned} \tag{3.100}$$

which immediately integrate to

$$\begin{aligned} E &= E_0 - \frac{2}{3}\eta, \\ D &= \frac{D_0 E_0^2}{[E_0 - (2/3)\eta]^2}, \end{aligned} \tag{3.101}$$

where $E_0^2 = E^2(\eta = 0)$ is the initial radius of the round two-sphere, whereas $D_0 = D(\eta = 0)$ is the scale on $\mathbb{R}^1$.

Given this solution of (3.77), we can construct the Ricci-Hamilton evolution of $(S^3_{(1)}\sharp S^3_{(2)}, \tilde{g})$ by looking for a solution of (3.77) in the form

$$\tilde{g}(x;\eta) = \sum_{\alpha} \xi_\alpha \chi^*_{(\alpha)}(\eta)[g(x) \cdot \delta_{(i)}(\mathrm{dist}_\eta(x, y_{(i)}))], \tag{3.102}$$

where the inclusion maps $\chi_{(\alpha)}$ depend now on the deformation parameter η. We assume that the metric $g(x)$, (the round metric on the spheres $S^3_{(i)}$ and the standard product metric on the neck$\simeq S^2 \times \mathbb{R}^1$) being locally homogeneous, is preserved by (3.77), since the Ricci-Hamilton flow preserves isometries. As the *plumbing* between the spheres and the neck shrinks as η increases, (as the S^2 factor in the neck), the inclusions $\chi_{(\alpha)}$ are necessarily η-dependent. The dynamics of $\chi_{(\alpha)}$ can be obtained as follows.

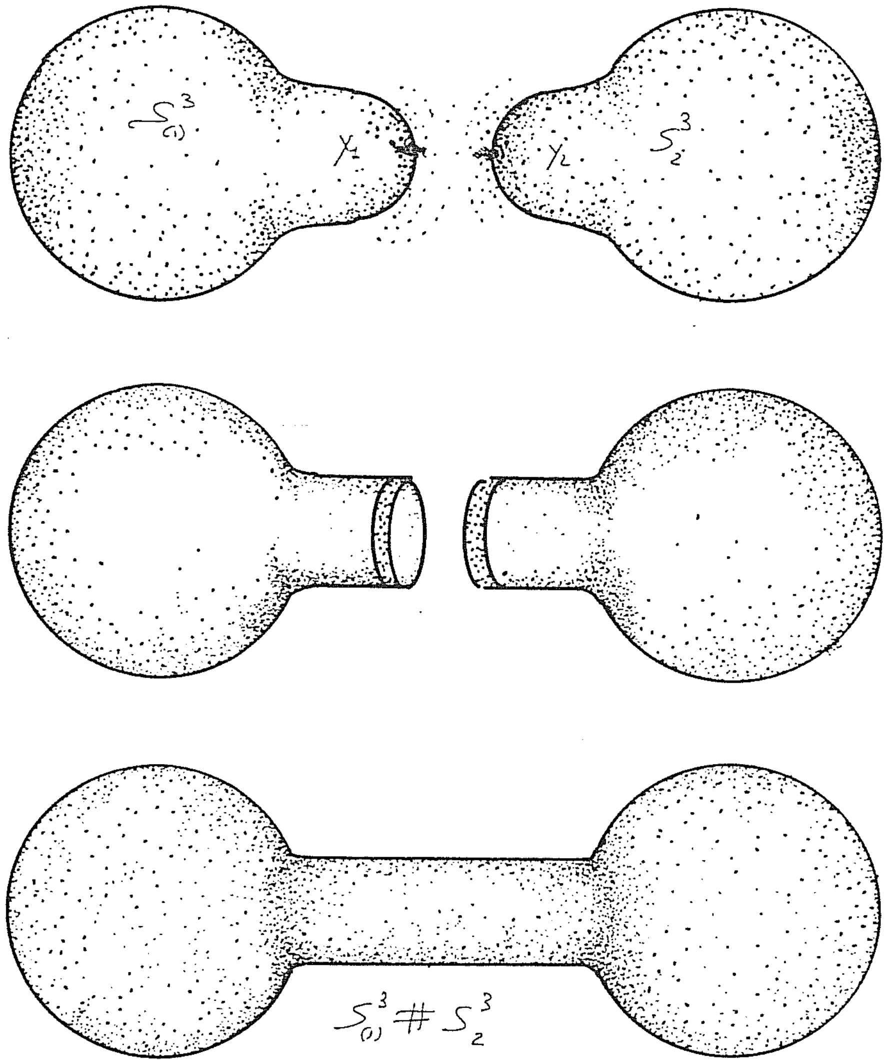

Figure 3.10: If we blow up the metric of two three-spheres $S^3_{(1)}$ and $S^3_{(2)}$, (here represented as two-spheres), around the points y_1 and y_2 and then join the resulting manifolds, we obtain the connected sum $S^3_{(1)} \sharp S^3_{(2)}$. According to this procedure, the neck $S^2 \times \mathbb{R}^1$ inherits a cylindrical metric up to exponentially small correction terms.

The Ricci-Hamilton flow (3.77) for $\tilde{g}(x;\eta)$ is given by

$$\begin{cases} \frac{\partial \tilde{g}_{ab}(\eta)}{\partial \eta} & = \frac{2}{3} < \tilde{R}(\eta) > \tilde{g}_{ab}(\eta) - 2\tilde{R}_{ab}(\eta), \\ \tilde{g}_{ab}(\eta = 0) & = \tilde{g}_{ab}, \end{cases}$$

and, in terms of the pulled-back metric it is

$$\begin{aligned} \frac{\partial}{\partial \eta}[\chi^*_{(\alpha)}(\eta) g(x) \cdot \delta_{(i)}]_{ab} &= \frac{2}{3} < \tilde{R}(\eta) > [\chi^*_{(\alpha)}(\eta) g(x) \cdot \delta_{(i)}]_{ab} \\ &- 2\tilde{R}(\chi^*_{(\alpha)}(\eta) g(x) \cdot \delta_{(i)})_{ab}. \end{aligned} \tag{3.103}$$

The left hand side of (3.103) can be evaluated according to a suggestion first exploited by D. DeTurck, *viz.*,

$$\begin{aligned} \frac{\partial}{\partial \eta}[\chi^*_{(\alpha)}(\eta) g \cdot \delta_{(i)}]_{ab}(x) &= \chi^*_{(\alpha)}(\eta) \left[\frac{\partial}{\partial \eta}[g \cdot \delta_{(i)}]_{ab}(\chi(\eta, x))\right] \\ &+ \chi^*_{(\alpha)}(\eta) \left[L_{w(\eta)}[g \cdot \delta_{(i)}]_{ab}(\chi(\eta, x))\right], \end{aligned} \tag{3.104}$$

where the Lie derivative $L_{w(\eta)}[g \cdot \delta_{(i)}]_{ab}(\chi(\eta, x))$ is along the vector field $w(\eta;\alpha)$ which generates the η-evolution of the inclusions $\chi_{(\alpha)}$, *viz.*,

$$\begin{aligned} \frac{\partial \chi_{(\alpha)}(\eta)}{\partial \eta} &= w(\eta; \chi_{(\alpha)}(\eta)), \\ \chi_{(\alpha)}(\eta = 0) &= \chi_{(\alpha)}. \end{aligned} \tag{3.105}$$

Let us denote by $< R >_{(\alpha)}$ the average of the scalar curvature $R(\chi_{(\alpha)}(x))$ over the images of the inclusions $\chi_{(\alpha)}$, (*viz.*, for $\alpha = 1, 2$, $< R >_{(\alpha)}$ is the average over $S^3_{(i)}$, while for $\alpha = 3$, the average is over the neck). In terms of these averages, the Ricci-Hamilton flow (3.103) can be written as

$$\begin{aligned} \frac{\partial}{\partial \eta}[\delta_{(i)} \cdot g_{ab}(\chi_{(\alpha)}(x))] &= \frac{2}{3} < R(\eta) >_{(\alpha)} \delta_{(i)} \cdot g_{ab}(\chi_{(\alpha)}(x)) - 2R_{ab}(\chi_{(\alpha)}(x)) \\ &+ \frac{2}{3}\delta_{(i)} \cdot g_{ab}(\chi_{(\alpha)}(x)) \left[< R(\eta) > - < R(\eta) >_{(\alpha)}\right] \\ &- L_{w(\eta)}[g \cdot \delta_{(i)}]_{ab}(\chi(\eta, x)). \end{aligned} \tag{3.106}$$

As can be checked, the Ricci-Hamilton flow preserves the local homogeneous geometries over the spheres $S^3_{(i)}$, and on the neck, if and only if the vector field $w(\eta;\alpha)$ satisfies

$$\frac{2}{3}\delta_{(i)}\cdot g_{ab}(\chi_{(\alpha)}(x))\left[< R(\eta) > - < R(\eta) >_{(\alpha)}\right] = L_{w(\eta)}[g\cdot\delta_{(i)}]_{ab}(\chi(\eta,x)). \qquad (3.107)$$

By taking the trace of (3.107) with respect to $g\cdot\delta_{(i)}$, we get

$$< R > - < R >_{(\alpha)} = \bar{\nabla}^k w_k(\eta;\alpha), \qquad (3.108)$$

where $\bar{\nabla}$ denotes the riemannian connection with respect to $g\cdot\delta_{(i)}$.

For $\alpha = 1,2$, i.e. for the punctured spheres $S^3_{(i)} - \{y_i\}$, the above relation yields upon integration over $S^3_{(i)} - B(y_i, r)$

$$\left[< R > - < R >_{(\alpha)}\right] vol\left(S^3_{(i)} - B(y_i,r)\right) = \int_{S^2(\eta)} w_k(\eta;\alpha)d\sigma^k, \qquad (3.109)$$

where $S^2(\eta)$ is the η-dependent boundary of $S^3_{(i)} - B(y_i, r)$. Since $< R > - < R >_{(\alpha)}$, for $\alpha = 1,2$, is proportional to the average scalar curvature of the neck, the term on the left hand side of (3.109) blows up as $(E_0 - \frac{2}{3}\eta)^{-1}$ as η increases. On the other hand, by the $S^2(\eta)$- rotational symmetry, the term $w_k(\eta;\alpha)d\sigma^k$ is spatially constant on the two-sphere boundary $S^2(\eta)$ of the punctured three-spheres. Thus, (3.109) implies that as η increases, the surface area of $S^2(\eta)$ shrinks to zero, as $(E_0 - \frac{2}{3}\eta)^{-1}$, by moving along the outer normal direction of $S^2(\eta) \subset (S^3_{(i)} - B(y_i, r(\eta)))$.

For $\alpha = 3$, i.e., for the neck, (3.109) yields

$$\left[< R > - < R >_{(\alpha=3)}\right] vol\left(S^2(\eta)\times\mathbb{R}^1\right) = \int_{S^2_{(2)}(\eta)} w_k(\eta;\alpha)d\sigma^k - \int_{S^2_{(1)}(\eta)} w_k(\eta;\alpha)d\sigma^k, \qquad (3.110)$$

where $S^2_{(2)}(\eta) - S^2_{(1)}(\eta)$ is the oriented boundary of the cylinder $S^2(\eta)\times\mathbb{R}^1$. Since $< R >$ $- < R >_{(\alpha=3)}$ is, up to small correction terms (coming from the collars joining the spheres to the neck), the average curvature over the spheres $S^3_{(i)}$, $i = 1,2$, (3.110) simply tells us that as the $S^2(\eta)$ factor in the neck shrinks, the scale-length of the $\mathbb{R}^1$ factor grows, so that the volume $vol\,(S^2(\eta)\times\mathbb{E}^1)$ remains constant during the Ricci-Hamilton evolution, namely,

$$vol\left(S^2(\eta)\times\mathbb{E}^1\right)=\frac{\int_{S^2_{(2)}(\eta)}w_k(\eta;\alpha)d\sigma^k-\int_{S^2_{(1)}(\eta)}w_k(\eta;\alpha)d\sigma^k}{<R>-<R>_{(\alpha=3)}}. \tag{3.111}$$

Notice that by introducing, as above, a new variable

$$\mathrm{dist}_\eta(x,y_{(i)})=\exp[-\tau(\eta)], \tag{3.112}$$

with

$$\tau(\eta)=\frac{\tau}{[1-\frac{2}{3}\eta]}, \tag{3.113}$$

and by rescaling the unit two-sphere metric h_{ij} according to

$$h_{ij}(\eta)=[1-\frac{2}{3}\eta]h_{ij}, \tag{3.114}$$

the above analysis of the Ricci-Hamilton flow (3.103) provides on the neck the metric

$$\tilde{g}(x;\eta)=d\tau^2(\eta)+h_{ij}(\eta)d\theta^i d\theta^j+\mathcal{O}(e^{-2\tau(\eta)}), \tag{3.115}$$

and on $S^3_{(i)}$ the standard round metric.

Notice also that (3.115) goes cylindrical exponentially fast in η. As η increases $0\leq\eta<\frac{3}{2}$, the neck becomes longer and longer while getting thinner and thinner. In the limit, $\eta\to\frac{3}{2}$, we get

$$(S^3_{(1)}\sharp S^3_{(2)},\tilde{g})\to S^3_{(1)}\coprod S^3_{(2)}, \tag{3.116}$$

where the three-spheres $S^3_{(i)}$ carry the round metric of volume 1, while the smooth joining regions shrink exponentially fast around the points $y_{(i)}$, $i=1,2$.

3.6.3.4 Critical surfaces and topological crossover

Let g_o be a metric on Σ with positive Ricci curvature,

$$Ric(g_o) > 0 \tag{3.117}$$

and with volume $vol(\Sigma, g_o) = 2$.

By means of g_o and the metric $\tilde{g}_1$ introduced in the previous section, we can construct on Σ a smooth one-parameter, $(0 \leq t \leq 1)$, family of metrics g_t, with $g_{t=0} = g_o$ and $g_{t=1} = \tilde{g}_1$ by setting

$$\begin{aligned} &g_t \equiv (1-t)g_o + t\tilde{g}_1, \\ &0 \leq t \leq 1. \end{aligned} \tag{3.118}$$

According to Hamilton's theorem [124], there is a right-open neighborhood of $t = 0$ such that all metrics g_t in this neighborhood are in $SU(2)$-Basins and are attracted, under the action of the Ricci-Hamilton flow (3.77), towards the round metrics on S^3 with volumes v_t, (since the volume of g_t changes with t, we have a family of fixed points, namely, round three-spheres parameterized by the corresponding volumes v_t which are kept constant under the Ricci-Hamilton flow; thus $v_{t=0} = 2$).

On the other hand, according to the remarks above, we have an open neighborhood of $t = 1$ such that all metrics g_t in this neighborhood go *singular* under the Ricci-Hamilton flow. Indeed, for $t = 1$, the Ricci-Hamilton flow of $\tilde{g}$ fixes the round S^3 factors while the joining tube, $S^2 \times \mathbb{E}^1$, is driven towards a curvature singularity. By continuity, this behavior extends to a left-open neighborhood of $t = 1$, and a three-sphere in this neighborhood splits apart into two round spheres.

It follows that there is a neighborhood of the $\tilde{g}$ metric such that some of the metrics in this neighborhood are driven towards attractive $SU(2)$-Basins, while others are driven towards a *singular* $(S^3_{(1)} \sharp S^3_{(2)})$-Basin.

The set of metrics driven towards this basin defines a *critical surface.* Below we shall define a critical fixed point. Since the Ricci-Hamilton flow preserves the metric $\tilde{g}$, *up to* *a trivial rescaling* of the neck in $S^3_{(1)}\sharp S^3_{(2)}$, (rigorously speaking this is true only up to exponentially small correction terms which arise in the regions joining the three-spheres with the $S^2\times\mathbb{E}^1$ neck), we can characterize the *critical fixed point* for (3.77), in the example considered, as the metric on $S^3_{(1)}\sharp S^3_{(2)}$ given by

$$\mathcal{G}_{crit} \equiv \{\bar{g}(x;\eta)|\ \eta \to \frac{3}{2}\}, \tag{3.119}$$

where $\bar{g}(x;\eta)$ is the one parameter η-family of metric solution of the Ricci-Hamilton initial value problem (3.103). This is a consistent characterization of a critical fixed point for (3.77), since (3.77) fixes $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$, and as required for (3.87), the correlation length associated with $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$ is $L=\infty$. This is so since, roughly speaking, in order to describe $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$ we fix any representative $(S^3_{(1)}\sharp S^3_{(2)})$ which is characterized by a correlation length $L_0 \simeq$ *the length of the neck*, and then rescale with the Ricci-Hamilton flow (3.103) from L_o to $L=\infty$.

Notice however that in this particular case, the critical point is not related to a phase transition. As mentioned before, we are here in presence of significant finite-size effects which are concerned with a dimensional crossover and they show up, as usual, as an *effective reduction of dimensionality* [50]. In this case, the three-dimensional neck of $(S^3_{(1)}\sharp S^3_{(2)})$ goes one-dimensional. This crossover to (quasi-)one-dimensional behavior is not accompanied by a singular behavior in thermodynamical quantities, such as correlation functions, but nonetheless an anomalous behavior is present. Indeed, geometrically speaking, the critical fixed point $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$, and the corresponding *critical surface*, separates two stable phases under the renormalization generated by (3.77). One is given by the manifolds eventually evolving towards the round three-sphere of volume 2. The other is generated by those manifolds which pinch off and eventually evolve towards two round three-spheres, each of volume 1. The *pinching off* through thinner and thinner necking is necessary for such a topological crossover.

From a physical point of view, and as we shall see in section 6, the fixed point described above separates two possible different closed FLRW regimes. One with closed spatial sections which, at some particular instant, are a three-sphere of volume 2, while in the other regime, we have *two* distinct closed FLRW universes having spatial sections, (at a given instant), of volume 1. We may have also many different necks corresponding to a regime whereby the spatial section Σ yields for many closed FLRW universes.

It is clear that the above explicit construction of a critical fixed point for (3.77) can in principle be generalized to more general situations. The strategy is to take two or more trivial fixed points, such as those associated with the $SU(2)$ and $H(3)$-Basins, and connect them through the degenerate basins (such as $SO(3)\times\mathbb{R}^1$, as in the above example, or through the $H(2)\times\mathbb{R}^1$-Basin, etc.). In this connection notice that the connected sum mechanism yielding for the $(S^3_{(1)}\sharp S^3_{(2)})$-critical fixed point, can be generalized to an operation of joining two (or more) manifolds (corresponding to stable attractors) along tubular neighborhoods of surfaces (rather than points, as in the case for the standard connected sum).

A particularly interesting connecting geometry would be $H(2)\times\mathbb{R}^1$, (by compactifying the hyperbolic plane in a closed surface). In this latter case the scale of the hyperbolic geometry goes to infinity under the Ricci-Hamilton flow (*pancake degeneracy* [142]). Finite size effects are again at work, but this time the effective reduction of dimensionality is more interesting than the one in the $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$ case. Indeed, two out of three dimensions are infinite, and there will be a crossover to a critical behavior with critical exponents characteristic of a two-dimensional system.

One can consider the analysis presented above as a physically non-trivial application of Hamilton-Thurston's geometrization program (cf. [247]). In some of its aspects it is rather conjectural and speculative, but in our opinion, it is quite intriguing that motivations coming from geometry and a physical problem, like the one addressed here, namely, the construction of cosmological models out of a local gravitational theory, go hand in hand in such a way.

3.6.4 Linearized RG flow

The relative slopes of $< f >_{\epsilon_m} (g_1)$ and $< f >_{\epsilon_m} (g_2)$, with f given by equation (3.61), as $m \to \infty$, and for g_1 in a neighborhood of g_2, are of some relevance to our discussion. In a standard RG analysis such relative slopes are related to critical exponents. Given

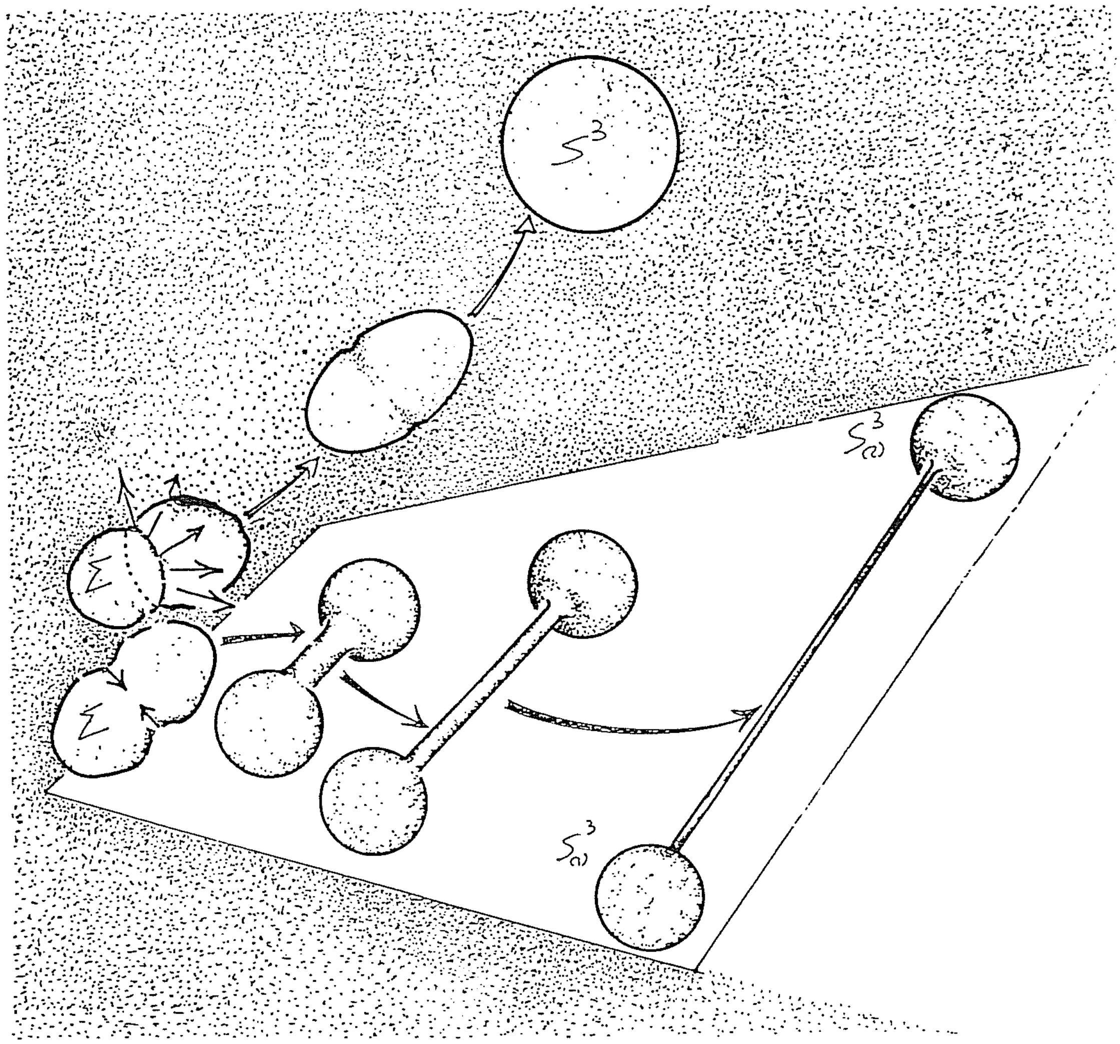

Figure 3.11: Different deformations of the initial three-manifold Σ may have quite different fates. If the waist of Σ is increased enough by the deformation, (i.e., if it is rounded), the Ricci-Hamilton flow will evolve Σ toward the round three-sphere. If the waist is shrinked enough by the initial deformation, the Ricci-Hamilton flow will evolve Σ toward $S^3_{(1)} \coprod S^3_{(2)}$. The three-manifolds attracted towards $S^3_{(1)} \coprod S^3_{(2)}$ define a *critical surface*.

the blocking procedure for $< f >_{\epsilon_m} (g_1)$ as $m \to \infty$, yielding for a g_1 "renormalized" according to the Ricci- Hamilton flow (3.77), one can sensibly ask what happens if g_1 is slightly perturbed, namely, if we replace it by

$$g_{ab} \to g_{ab} + \delta K_{ab}, \tag{3.120}$$

where K_{ab} is a symmetric bilinear form (a choice of the symbol is quite intentional, since later the above consideration will be applied to the second fundamental form). It can be shown that if g_1 is scaled, according to the Ricci-Hamilton flow (3.77), then K_{ab} gets renormalized according to the linearized Ricci-Hamilton flow, namely (η in the brackets suppressed),

$$\begin{aligned} \frac{\partial}{\partial \eta} K_{ab} &= \frac{2}{3} < R > K_{ab} + \frac{2}{3} g_{ab} [\frac{1}{2} < R g^{ab} K_{ab} > - \frac{1}{2} < R >< g^{ab} K_{ab} > - \\ & < R^{ab} K_{ab} >] - \Delta_L K_{ab} + 2[div^*(div(K - \frac{1}{2}(Tr\, K)\, g))]_{ab}, \end{aligned} \tag{3.121}$$

with the initial data $K_{ab}(\eta = 0) = K_{ab}$, where $K \in S^2\Sigma$ is a given symmetric bilinear form, Δ_L is the Lichnerowicz-DeRham Laplacian on bilinear forms

$$\Delta_L K_{ab} \equiv -\nabla^s \nabla_s K_{ab} + R_{as} K^s_{\ b} + R_{bs} K^s_{\ a} - 2 R_{asbt} K^{st}, \tag{3.122}$$

and the operators Δ_L, div^*, div and Tr are considered with respect to the flow of metric $(g, \eta) \to g(\eta)$, solution of (3.77). The div (here, minus the usual divergence) is the divergence operator on $S^2\Sigma$, div^* is the L^2- adjoint of div, acting from the space of vector fields on Σ to $S^2\Sigma$ (it can be identified with $\frac{1}{2}$[Lie derivative] of the metric tensor along a vector field).

Note that $K(\eta)$ solution of the linear (weakly) parabolic initial value problem (3.121) always exists and is unique [124], and represents an infinitesimal deformation of metrics connecting the two neighbouring flows of metrics, $g(\eta)$ and $g'(\eta)$, (obtained as solutions of problem (3.77) with initial data $g(\eta = 0) = g$ and $g'(\eta = 0) = g(\eta = 0) + \epsilon K(\eta = 0) + \mathcal{O}(\epsilon^2)$, respectively). For what concerns the structure of this solution, one can verify that corresponding to the "trivial" initial datum $K(\eta = 0) = L_X g$ (where $X : \Sigma \to T\Sigma$ is a smooth vector field on Σ), the solution of (3.121) is provided by

$$K_{ab}(\eta) = L_X g_{ab}(\eta). \tag{3.123}$$

This property expresses the $Diff(\Sigma)$ equivariance of the Ricci-Hamilton flow. (Notice that X is η-independent).

The above fact follows by noticing that along the trajectories of the flow $(\eta, g) \to g(\eta)$, solution of (3.77), we have

$$\frac{\partial}{\partial\eta}L_X g_{ab}(\eta) = L_X[\frac{\partial}{\partial\eta}g_{ab}(\eta)] = \frac{2}{3} < R(\eta) >_\eta L_X g_{ab}(\eta) - 2L_X R_{ab}(\eta). \tag{3.124}$$

But the $Diff(\Sigma)$-equivariance of the Ricci tensor, i.e. the fact that $Ric(\varphi^* g) = \varphi^* Ric(g)$ for any smooth diffeomorphism $\varphi : \Sigma \to \Sigma$, implies that

$$L_X R_{ab} = D\,Ric(g) \cdot L_X g_{ab}, \tag{3.125}$$

where $D\,Ric(g)K$ is the formal linearization of $Ric(g)$, around g, in the direction K:

$$\begin{aligned} D\,Ric(g) \cdot K &\equiv \frac{d}{dt}[Ric(g + tK)]_{t=0} \\ &= \frac{1}{2}\Delta_L K - div^*[div(K - \frac{1}{2}(Tr\,K)g)]. \end{aligned} \tag{3.126}$$

Upon introducing (3.125) in (3.124) we get

$$\frac{\partial}{\partial\eta}L_X g_{ab}(\eta) = \frac{2}{3} < R(\eta) >_\eta L_X g_{ab}(\eta) - 2D\,Ric(g(\eta)) \cdot L_X g_{ab}(\eta). \tag{3.127}$$

One can check that the right hand side of the above expression coincides with the right hand side of (3.121), when this latter is evaluated for $K_{ab}(\eta) = L_X g_{ab}(\eta)$. Hence $L_X g_{ab}(\eta)$ solves the partial differential equation (3.121) and, since for $\eta = 0$, $K_{ab} = L_X g_{ab}$, the uniqueness of any solution of the initial value problem (3.121) implies that $K_{ab}(\eta) = L_X g_{ab}(\eta)$, whenever $K_{ab}(\eta = 0) = L_X g_{ab}$, as stated.

Moreover, if $K(\eta)$ is a solution of (3.121), with initial datum $K(\eta = 0) = K$, then the space average of $Tr\,K(\eta)$ over $(\Sigma, g(\eta))$ is preserved along the flow $(\eta, g) \to g(\eta)$, namely,

$$< Tr\,K(\eta) >_\eta = < Tr\,K >_o, \quad 0 \leq \eta < \infty. \tag{3.128}$$

This property of the solutions of (3.121) is an immediate consequence of the volume-preserving character of the Ricci-Hamilton flow.

Finally, another relevant property of the initial value problem (3.121) can be stated as follows. If $(\eta, K_{ab}) \to K_{ab}(\eta)$ is the flow solution of (3.121), with initial datum $K_{ab}(\eta =$

$0) = K_{ab}$, then it can always be written as [175]

$$K_{ab}(\eta) = \hat{K}_{ab}(\eta) + L_{v(\eta)} g_{ab}(\eta), \tag{3.129}$$

where the bilinear form $\hat{K}_{ab}(\eta)$ and the η−dependent vector field $v(\eta)$, respectively, are the solutions of the initial value problems:

$$\begin{aligned}
&\frac{\partial}{\partial \eta} \hat{K}_{ab} = \frac{2}{3} < R > \hat{K}_{ab} + \frac{2}{3} g_{ab} [\frac{1}{2} < R g^{ab} \hat{K}_{ab} > - \frac{1}{2} < R >< g^{ab} \hat{K}_{ab} > - \\
&< R^{ab} \hat{K}_{ab} >] - \Delta_L \hat{K}_{ab}, \\
&\hat{K}_{ab}(\eta = 0) = K_{ab},
\end{aligned} \tag{3.130}$$

and

$$\frac{\partial}{\partial \eta} v_a(\eta) = -\nabla^c (\hat{K}_{ca} - \frac{1}{2} \hat{K}^{rs} g_{rs} g_{ca}), \quad v(\eta = 0) = 0. \tag{3.131}$$

To summarize, as $\eta \to \infty$, $K_{ab}(\eta)$ may either approach a Lie derivative term, such as $L_{v(\eta)} g_{ab}(\eta)$, or a non-trivial deformation $\hat{K}_{ab}(\eta)$ [175]. The non-trivial deformation is present only if the corresponding Ricci-Hamilton flow for $g_{ab}(\eta)$ approach an Einstein metric on Σ which is not isolated. In such a case, (e.g. flat tori), there is a finite dimensional set of such Einstein metrics, and the non-trivial $\hat{K}_{ab}$ simply are the infinitesimal deformations connecting one Einstein metric $\bar{g}_1$ in Σ and the infinitesimally non-equivalent one. Also in this case the Lie derivative term may be present. What it represents is a reparameterization of the metric $\bar{g}_1$ (under the action of the infinitesimal group of diffeomorphisms generated by v (see equation (3.129)), ("gauge artifact"). This latter Lie derivative term is the only surviving term when $\bar{g}$ is isolated (like e.g. in the case of the round three-sphere. As is known, the round metric $\bar{g}$ on the three-sphere S^3 is isolated, in the sense that there are not volume-preserving infinitesimal deformations of $\bar{g}$ mapping it to another inequivalent constant curvature metric $\bar{g}'$.)

To summarize, the flow (3.130), (3.131), tells us how to renormalize the second fundamental form in such a way that the blocking prescription $< f >_{\epsilon_m} \to < f >_{\epsilon_{m+1}}$ works both for the initial metric g as well as for the *perturbed* metric $g + \delta K_{ab}$.

3.6.4.1 Scaling and critical exponents

From the characterization of critical fixed points $g^{\star}_{ab}$ for (3.77), (see section 4.2), we can get information on the critical exponent characterizing critical behavior of the metrics nearby $g^{\star}_{ab}$. This is the content of this section. In particular, we shall discuss the critical exponents related to the critical fixed point $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$. Even if this point is not a thermodynamically interesting critical point, it exhibits many of the general features of the more interesting type of singularities.

In order to characterize these critical exponents we can use the linearized Ricci-Hamilton flow associated with the one-parameter family of metrics $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$. In the following we shall however proceed more directly and examine the properties of the two-point correlation function associated with $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$ and the probability law of relevance to our analysis.

Let $y_i \in S^3_{(i)}$, $i = 1, 2$, be the two points in $(S^3_{(1)}\sharp S^3_{(2)}, \mathcal{G}_{crit})$ around which the round metrics of $S^3_{(i)}$ have been blown up. Let f denote a non-negative scalar field on $(S^3_{(1)}\sharp S^3_{(2)})$, distributed according to the probability law formally defined by

$$dP \equiv \frac{\exp[-H(f)]\prod_x df(x)}{\int \exp[-H(f)]\prod_x df(x)}, \tag{3.132}$$

where the functional integration is over the space of fields $f\colon (S^3_{(1)}\sharp S^3_{(2)}) \to \mathbb{R}$, equipped with the L^2-inner product

$$(f|f') \equiv \int_{(S^3_{(1)}\sharp S^3_{(2)})} f(x)f'(x)d\mu_g(x), \tag{3.133}$$

and where

$$H(f) = \int_{(S^3_{(1)}\sharp S^3_{(2)})} f(x)d\mu_g(x). \tag{3.134}$$

(We could consider f as related to the matter fields, as in $f = \alpha\rho + \alpha^i J_i$, but the following analysis is quite independent from a particular meaning of f). Notice also that the measure $d\mu_g(x)$ in the above formulae is the riemannian measure on $(S^3_{(1)}\sharp S^3_{(2)})$ associated with the metric $\bar{g}(x)$ defined by (3.98).

Let us now concentrate on the behavior of (3.132) when the neck of $(S^3_{(1)} \sharp S^3_{(2)})$ gets thinner and longer under the action of (3.77).

It can then be checked that along the Ricci-Hamilton flow (3.101) associated with $\bar{g}|_{\mathrm{neck}}$, both the L^2-inner product (3.133) and the Gibbs factor $\exp[-H(f)]$ corresponding to (3.134), are invariant. Thus it follows that the probability measure (3.132) is invariant under (3.77), and the correlation function defined by

$$E_{dP}[f(y_1)f(y_2)], \tag{3.135}$$

where E_{dP} denotes the expectation with respect to dP, is well defined over the *critical* fixed point $\mathcal{G}_{crit}$.

In section 4.1 we have interpreted the Ricci-Hamilton flow (3.77) as the RG flow in a finite geometry, characterized by the length scale L, which is large compared to the microscopic scale (in particular, it is much larger than the radius of the geodesic ball coverings used to discretize the theory). The correlation function depends on such a dimensional parameter. If

$$L(y_1, y_2) \equiv \int_{y_1}^{y_2} \sqrt{D} dz \tag{3.136}$$

denotes the distance between the points y_1 and y_2 along the cylindrical neck of $(S^3_{(1)} \sharp S^3_{(2)})$, and if this distance is large (as compared with the radius of geodesic ball coverings), then the *correlation length* ξ associated with the two-point (connected) correlation function can be read off from

$$\xi \simeq_{L(y_1,y_2)\gg 1} \frac{-L(y_1, y_2)}{\ln E_{dP}(f(y_1)f(y_2))_{conn}}. \tag{3.137}$$

Since the correlation function remains invariant under the Ricci-Hamilton deformation of the cylindrical neck, we get that along (3.101), $\xi/L(y_1, y_2)$ remains constant which implies that on $(S^3_{(1)} \sharp S^3_{(2)}, \mathcal{G}_{crit})$ the correlation length ξ behaves as

$$\xi \simeq \sqrt{D_0} E_0[E_0 - \frac{2}{3}\eta]^{-(1)}. \tag{3.138}$$

According to standard usage, we can define the *critical exponent* ν associated with the correlation length of a finite size system (with typical size L) by the condition

$$\frac{\partial}{\partial p}\xi(L,p)|_{p=p_c} \simeq L^{1+\frac{1}{\nu}}, \tag{3.139}$$

where p is a parameter driving the system to criticality, and p_c is its corresponding critical value. In our case, it is natural to set $p = \eta$, with $p_c = \eta_c = \frac{3}{2}$. This immediately yields for the critical exponent ν the value

$$\nu = 1. \tag{3.140}$$

A similar computation for the critical exponent associated with the correlation length can be carried out when the connecting geometry is $H^2 \times \mathbb{R}^1$. This takes place when two riemannian manifolds M_1 and M_2, which are supposed to evolve nicely under the Ricci-Hamilton flow, are connected through a tubular neighborhood of a surface S_h of genus h, *viz.*, $\sigma = M_1 \sharp_{S_h} M_2$.

In this case, the metric on the neck can be written as [142]

$$g_{neck} = Dg_{\mathbb{R}^1} + Eg_{H^2}, \tag{3.141}$$

where g_{H^2} is the metric on the hyperbolic plane. The Ricci-Hamilton flow equations take a form similar to the $S^2 \times \mathbb{R}^1$ case, up to a minus (important!) sign, namely,

$$\begin{aligned} \frac{d}{d\eta}E &= \frac{2}{3}, \\ \frac{d}{d\eta}D &= -\frac{4}{3}\left(\frac{D}{E}\right), \end{aligned} \tag{3.142}$$

which upon integration provide

$$\begin{aligned} E &= E_0 + \frac{2}{3}\eta, \\ D &= \frac{D_0 E_0^2}{[E_0 + (2/3)\eta]^2}, \end{aligned} \tag{3.143}$$

where $E_0^2 = E^2(\eta = 0)$ is the initial scale of the hyperbolic geometry, whereas $D_0 = D(\eta = 0)$ is the scale on $\mathbb{E}^1$.

The scale E of the hyperbolic geometry increases linearly with η, while the scale factor D on the line $\mathbb{E}^1$ shrinks. It can be checked [142], that corresponding to this scale dynamics, the curvature decays according to $||Ric|| = \sqrt{2}/(E_0 + \frac{2}{3}\eta)$, and we get in the limit $\eta \to \infty$ a *pancake* degeneracy.

The relevant correlation function is now $E_{dP}(f(y_1)f(y_2))_{conn}$, with y_1 and y_2 fixed points in the H^2 factor, (i.e., on the surface S_h). Again, owing to the symmetries of the geometry involved, it immediately follows that the ratio between the correlation length ξ and the distance $L(y_1, y_2)$ must remain constant under the Ricci-Hamilton flow. This implies that the correlation length behaves as

$$\xi \simeq \sqrt{E_0 + \frac{2}{3}\eta}, \tag{3.144}$$

to which corresponds again a critical exponent $\nu = 1$, (in this case one cannot apply (3.139) since the system is actually going to an infinite size).

A striking feature of these topological crossover phenomena associated with the renormalization of the cosmological matter distribution, is that their pattern resembles the linear sheet-like (or honeycomb-like) structure in the distribution of galaxies on large scales (e.g. [111]). It is evident that if $(\Sigma, g, K, \rho, \mathbf{J})$, the initial data set for the real universe (see the next section), is close to *criticality*, in the sense discussed in the previous section, then the corresponding averaged model exhibits a tendency to topological crossover in various regions (the ones where the inhomogeneities are larger). Filament-like and sheet-like structures would emerge, and the overall situation would be the one where such structures appear together with regions of high homogeneity and isotropy in some sort of hierarchy. This situation is akin to that of a ferromagnet nearby its critical temperature, whereby we have islands of spins up and down in some sort of nested pattern. Even if there is a tendency to homogeneity at very large scales, the picture just sketched, of the "hierarchy of structures", might be qualitatively valid in a good part of the universe; indeed recent observational data seem to suggest the existence of still larger and larger structures (e.g.

[61]). Notice that this picture bears some resemblance to the "cascade of fluctuations" in critical phenomena [255]. Droplet fluctuations nucleated at the lattice scale in the critical state can grow to the size of the correlation length where the details of the lattice structure become lost and the scale invariant distributions of the large "droplets" are universal.

A possible connection of this whole picture with the self-organized criticality (SOC) [16] can be envisaged, whereby the problem of structure formation, in terms of growth phenomena, could be tackled in the framework of avalanche activity used in SOC [232]. It would be particularly interesting to estimate the scales in the universe, where it is necessarily critical and trapped into self-organized (critical) states (similar suggestion was recently posed in [212]).

3.6.5 Effective cosmological models

The results of the previous sections have interesting consequences when applied in a cosmological setting.

Let us assume that at the scale over which General Relativity is experimentally verified, a cosmological model of our universe is provided by evolving a set of consistent initial data $(\Sigma, g, K, \rho, \mathbf{J})$, according to the evolutive part of Einstein's equations.

The data $(\Sigma, g, K, \rho, \mathbf{J})$ describing the interaction between the actual distribution of sources and the inhomogeneous geometry of the physical space, (Σ, g, K), are required to satisfy the Hamiltonian and the divergence constraints. Respectively,

$$8\pi G\rho = R(g) - K^{ab}K_{ab} + k^2, \tag{3.145}$$

$$\nabla^i K_{ih} - \nabla_h k = 16\pi G\, J_h, \tag{3.146}$$

where, $k \equiv K^a{}_a$, (see also section 4, where they were written in terms of the three-metric g_{ab} and its associated conjugate momentum π_{ab}).

According to the blocking and renormalization procedure, discussed at length in the previous sections, we can implement a coarse-graining transformation on this actual data set by suitably renormalizing the metric and the second fundamental form, while retaining the functional form and the validity of the constraints.

The metric is renormalized according to the Ricci-Hamilton flow (3.77),

$$\begin{cases} \frac{\partial g_{ab}(\eta)}{\partial \eta} & = \frac{2}{3} < R(\eta) > g_{ab}(\eta) - 2R_{ab}(\eta), \\ g_{ab}(\eta = 0) & = g_{ab}. \end{cases}$$

The second fundamental form K_{ab} is being renormalized too according to (3.121), *viz.*,

$$\begin{aligned} \frac{\partial}{\partial \eta} K_{ab} &= \frac{2}{3} < R > K_{ab} + \frac{2}{3} g_{ab} [\frac{1}{2} < R g^{ab} K_{ab} > - \frac{1}{2} < R >< g^{ab} K_{ab} > - \\ & < R^{ab} K_{ab} >] - \Delta_L K_{ab} + 2[div^*(div(K - \frac{1}{2}(Tr\, K)\, g))]_{ab}, \end{aligned} \tag{3.147}$$

with the initial data $K_{ab}(\eta = 0) = K_{ab} + L_\alpha g_{ab}$, and where α^i is the shift vector. Notice that we renormalize the second fundamental form by using as initial condition, not the actual second fundamental form but rather, the deformation tensor $H_{ab} \equiv K_{ab} + L_\alpha g_{ab}$. This way we can get rid of the possible $Diff$-induced shear which may develop in $\lim_{\eta \to \infty} K_{ab}(\eta)$. Indeed, according to equation (3.129), as $\eta \to \infty$ the solution of this initial value problem, $K(\eta)$, approaches a non-trivial deformation $\lim_{\eta \to \infty} \hat{K}(\eta)$ plus a Lie-derivative term, $\lim_{\eta \to \infty} L_{v(\eta)} g_{ab}(\eta)$. The former is present only if the Ricci-deformed metric $\bar{g}_{ab} = \lim_{\eta \to \infty} g_{ab}(\eta)$ is not isolated.

Since at this stage we are mainly interested in FLRW space-times, let us assume that $\bar{g}_{ab}$ is isolated, while in order to take care of the $Diff$-induced shear, $\lim_{\eta \to \infty} L_{v(\eta)} g_{ab}(\eta)$, we can choose the shift vector field α^i in such a way that $\lim_{\eta \to \infty} L_{v(\eta)} g_{ab}(\eta)$ is compensated by $L_\alpha g_{ab}$. Since $L_\alpha g_{ab}$ is a trivial datum for (3.121), it is thus sufficient to choose

$$\alpha^i = \lim_{\eta \to \infty} v(\eta). \tag{3.148}$$

Notice that α^i is the three-velocity vector of the chosen instantaneous observers on Σ, thus (3.148) provides a map identifying corresponding points between the initial *actual* manifold $(\Sigma, g, K, \rho, \mathbf{J})$ and its renormalized counterpart $\lim_{\eta \to \infty} (\Sigma, g, K, \rho, \mathbf{J})(\eta)$.

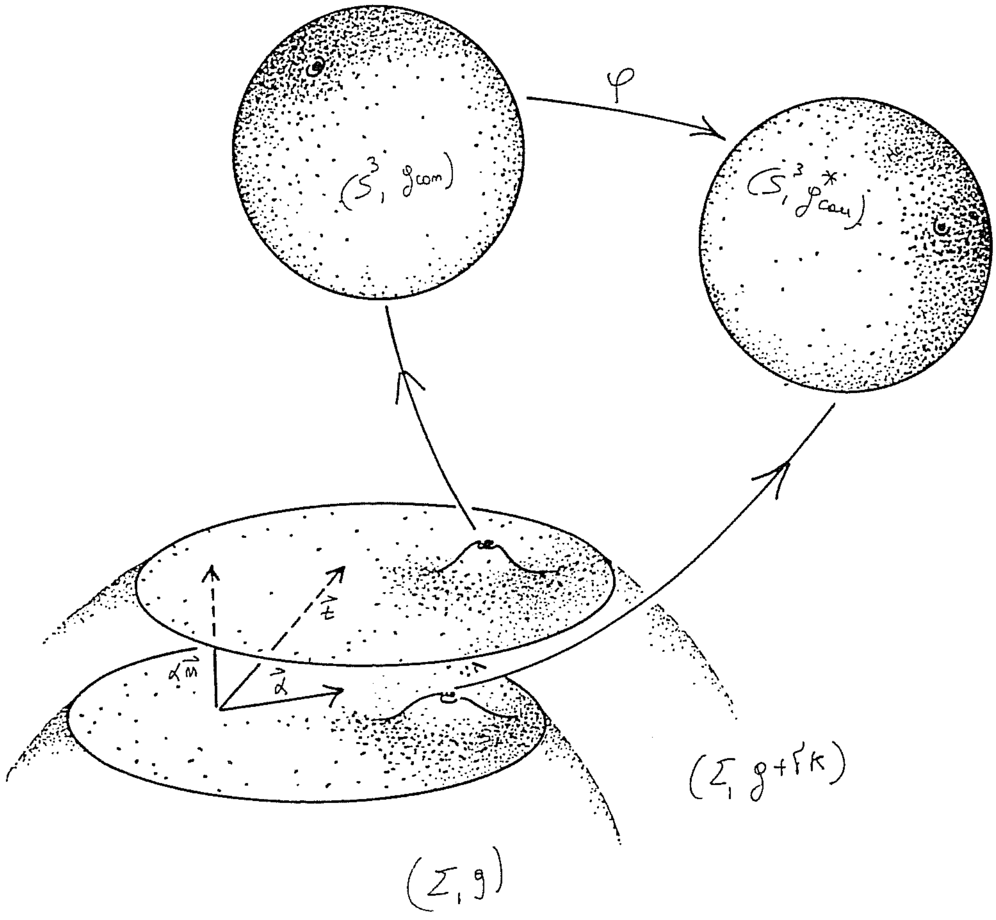

Figure 3.12: Two neighbooring three-manifolds (Σ, g) and $(\Sigma, g+\delta K)$, with the same volume, evolving under the Ricci-Hamilton flow toward isometric round three-spheres (S^3, g_{can}) and (S^3, g^*_{can}). In general, g_{can} and g^*_{can} differ by an infinitesimal diffeomorphism ϕ generated by a vector field v, i.e. $g^*_{can} = g_{can} + \delta(L_v g_{can} - \frac{2}{3} g_{can} \nabla_i v^i)$. By exploiting the freedom in choosing the shift vector field $\vec{\alpha}$, we can get rid of ϕ.

In this section, and mainly for actual computational purposes, we assume that the original inhomogeneous initial data set is such that the Ricci-Hamilton flow is global. As already mentioned, we are interested in connecting an inhomogeneous cosmological spacetime to its corresponding FLRW model. This is the case, in particular, if we assume that the original manifold (Σ, g) has a positive Ricci tensor, (this case is obviously quite similar to the analysis in [51], there are however important differences that we are going to emphasize). Or more generally, if we assume that the original manifold (Σ, g) is in the $SU(2)$-Basin of attraction or *nearby* the critical point $\sharp_i S^3_{(i)}$, with $i = 1, 2, \ldots$, yielding for a manifold (Σ, g), nucleating under the Ricci-Hamilton flow (3.103), (extended to many connected sums), to disjoint three-spheres $S^3_{(i)}$.

Given this setting, we see that due to the properties of the Ricci-Hamilton flow we have $\lim_{\eta\to\infty} K_{ab}(\eta) = \frac{1}{3} < k >_o \cdot \bar{g}_{ab}$. The given K_{ab} is deformed by gradual elimination of its shear $K_{ab} - \frac{1}{3}kg_{ab}$ and the original (position dependent) rate of volume expansion k is being replaced with its corresponding average value.

Since the constraints, (3.145) and (3.146), are required to hold at each step of the renormalization procedure, we get

$$8\pi G(\eta)\rho(\eta) = R(g(\eta)) - K^{ab}(\eta)K_{ab}(\eta) + k^2(\eta), \tag{3.149}$$

$$\nabla(\eta)^i K_{ih}(\eta) - \nabla(\eta)_h k = 16\pi G(\eta)\, J_h(\eta), \tag{3.150}$$

where we have explicitly introduced a possible η-dependence into the gravitational coupling G.

Let us now explore the consequences of (3.149) and (3.150). From the stated hypothesis on the Ricci-Hamilton flow it follows immediately that, $K_{ab}(\eta) \to \frac{1}{3}\bar{k}\bar{g}_{ab}$ as $\eta \to \infty$, thus (3.150) implies that

$$\lim_{\eta\to\infty} J_h(\eta) = 0. \tag{3.151}$$

In order to analyze (3.149), we will make use of a property of the Ricci-Hamilton flow, namely that the flow $K(\eta)$, solution of (3.121), is such that $\frac{\partial}{\partial\eta} < k(\eta) >_\eta = 0$, i.e. the space average of the trace of the second fundamental form remains constant during the deformation. This allows us to write

$$< k >_o^2 = \bar{k}^2, \tag{3.152}$$

since in the limit, the volume expansion is simply a constant, and where $< \ldots >_o$ denotes the full average of the enclosed quantity with respect to the initial metric.

Equation (3.152) provides the Hubble constant on the FLRW time slice associated with the smoothed data corresponding to $(\Sigma, g, K, \rho, \mathbf{J})$.

More explicitly, let us write the FLRW metric in the standard form (units $c = 1$)

$$ds^2 = -dt^2 + S^2(t) d\sigma^2, \tag{3.153}$$

where $d\sigma^2$ is the metric of a three-space of constant curvature and it is time independent. As we are interested in the three-sphere case, the metric $d\sigma^2$ can be written as

$$d\sigma^2 = d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\phi^2). \tag{3.154}$$

Since the volume $vol(\Sigma, g)$ of the original inhomogeneous manifold is preserved by the Ricci-Hamilton flow, we can relate the factor $S^2(t)$, providing the inverse (sectional) curvature of the FLRW slice $t = t_o$, to $vol(\Sigma, g)$

$$S^2(t_o) = \left(\frac{vol(\Sigma, g)}{2\pi^2}\right)^{\frac{2}{3}}. \tag{3.155}$$

Notice in particular that the scalar curvature, towards which $R(g(\eta))$ evolves under the Ricci-Hamilton flow, is given by

$$\bar{R} \equiv \lim_{\eta \to \infty} R(g(\eta)) = \frac{3}{S^2} = 3\left(\frac{vol(\Sigma, g)}{2\pi^2}\right)^{-\frac{2}{3}}. \tag{3.156}$$

Having said this, the equation (3.149) becomes in the limit, after extracting a trace free part of K_{ab},

$$8\pi\bar{G}\bar{\rho} = \bar{R} + \frac{2}{3}\bar{k}^2, \tag{3.157}$$

since no residual shear survives, and where we have introduced the renormalized gravitational coupling

$$\bar{G} \equiv \lim_{\eta \to \infty} G(\eta). \tag{3.158}$$

On the other hand (3.149) gives

$$\frac{2}{3}k^2(\eta) = 8\pi G\rho(\eta) - R(g(\eta)) + \tilde{K}^{ab}(\eta)\tilde{K}_{ab}(\eta), \tag{3.159}$$

where, the shear $\tilde{K}_{ab} \equiv K_{ab} - \frac{1}{3}kg_{ab}$ has been explicitly introduced.

We can rewrite the last equation upon taking the average with respect to the initial metric, as

$$\frac{2}{3} < k^2 >_o = 8\pi G < \rho >_o - < R >_o + < \tilde{K}^{ab}\tilde{K}_{ab} >_o . \tag{3.160}$$

Taking into account (3.152), and noticing that the Hubble constant on the FLRW slice, corresponding via the Ricci-Hamilton flow to (Σ, g, K), is

$$H_o^2 = \left(\frac{1}{S}\frac{dS}{dt}\right)^2 = \frac{2}{9}\bar{k}^2, \tag{3.161}$$

we get

$$H_o^2 = \frac{8\pi G}{3} < \rho >_o - \frac{1}{3} < R >_o + \frac{1}{3} < \tilde{K}^{ab}\tilde{K}_{ab} >_o - \frac{2}{9}(< k^2 >_o - \bar{k}^2), \tag{3.162}$$

which corresponds, through (3.157), to the matter density distribution

$$8\pi\bar{G}\bar{\rho} = 8\pi G < \rho >_o + \bar{R} - < R >_o + < \tilde{K}^{ab}\tilde{K}_{ab} >_o - \frac{2}{3}(< k^2 >_o - \bar{k}^2). \tag{3.163}$$

We have to consider this expression for $\bar{G}\bar{\rho}$ as the *renormalized* effective sources entering into the Friedmann equation, if we want to describe the real locally inhomogeneous universe through a corresponding idealized FLRW model. As expected, the renormalized matter density shows contributions of geometric origin, either coming from the shear anisotropies or from the local fluctuations in curvature and volume expansion.

At this stage, few comments are in order, concerning in particular the *effective Hubble constant* H_o (3.162). First, we wish to emphasize that this is the theoretical expression for the Hubble constant if *one wishes to model the real locally anisotropic and inhomogeneous universe with a corresponding FLRW one.* Expression (3.162) clearly shows that apart from the expected contributions to the expansion rate coming from matter and curvature, there is a negative contribution coming from the local fluctuations in the expansion rate, and a *positive* contribution coming from the *shear term* $\frac{1}{3} < \tilde{K}^{ab}\tilde{K}_{ab} >_o$.

This term has its origin both in the presence of the gravitational radiation

$$\frac{1}{3} < \tilde{K}_{\perp}^{ab}\tilde{K}_{\perp ab} >_o, \tag{3.164}$$

where $\tilde{K}_{\perp}$ is the divergence-free part of the shear $\tilde{K}$, and in the anisotropies generated by the motion of matter

$$\frac{1}{3} < \tilde{K}_{||}^{ab}\tilde{K}_{||ab} >_o, \tag{3.165}$$

where $\tilde{K}_{||}$ is the longitudinal part of the shear, obtained as a solution to the equation

$$\nabla^i(\tilde{K}_{||})_{ih} - \frac{2}{3}\nabla_h k = 16\pi G\, J_h. \tag{3.166}$$

Notice that

$$< \tilde{K}^{ab}\tilde{K}_{ab} >_o = < \tilde{K}_{\perp}^{ab}\tilde{K}_{\perp ab} >_o + < \tilde{K}_{||}^{ab}\tilde{K}_{||ab} >_o. \tag{3.167}$$

The shear $< \tilde{K}^{ab}\tilde{K}_{ab} >_o$ yields a contribution to H_o which can be roughly estimated by exploiting the anisotropy measurements in the cosmic microwave background (CMB) radiation, as long as the frame used in averaging (i.e., the lapse α and the shift α^i) is, on the average, comoving with the cosmological fluid; this is quite a non-trivial requirement, since the Ricci-Hamilton flow also prescribes the shift. We also require that the original locally inhomogeneous manifold (Σ, g, K) does not differ too much from a standard FLRW $t =$ const. slice. In such case one can apply the analysis of [178] to conclude that, at the present epoch, the ratio between shear and expansion is of the order

$$\left(\frac{|\tilde{K}_{ab}|}{H}\right)_o < 4\epsilon, \tag{3.168}$$

where $\epsilon \equiv \max(\epsilon_1, \epsilon_2, \epsilon_3)$ denotes the upper limit of currently observed anisotropy in the CMB radiation temperature variation, and where ϵ_1, ϵ_2, ϵ_3, respectively denote the dipole, quadrupole, and octopole temperature anisotropies. On choosing $\epsilon \simeq 10^{-4}$, as indicated by the recent CMB radiation anisotropy measurements, one gets that the shear is at most about 10^{-3} of the expansion [178].

Thus, as long as we assume that the original locally inhomogeneous manifold (Σ, g, K) does not differ too much from a standard FLRW $t =$ const. slice, the contribution to H_o from $< \tilde{K}^{ab}\tilde{K}_{ab} >_o$ is certainly quite small.

However, if we take seriously the possibility that the real universe may be close to the critical phase, as argued in the previous sections, then the contribution from the shear is not just of a conceptual value. For example, the original data set (Σ, g, K) may be near the critical surface associated with the $\sharp_i S^3_{(i)}$ critical point. In this case we may generate, upon smoothing, a whole family of disconnected $t =$ const. FLRW slices, each one with its own Hubble constant $H_o(i)$, and these Hubble constants will be quite dominated by the large anisotropies $< \tilde{K}^{ab}\tilde{K}_{ab} >_o$ of the original manifold. We would not notice the contribution from the local shear, (since we would have been looking at a rather homogeneous and isotropic island), the large contribution would come from the regions of large inhomogeneities and anisotropies which, under the Ricci-Hamilton renormalization, undergo the topological crossover.

In the above analysis, we also introduced a renormalized gravitational coupling. In a sense this is superfluous since the three-dimensional metric g of Σ is acting as the running coupling constant, and we can always reabsorb G in the definition of g. Nevertheless, the use of the renormalized coupling $\bar{G}$ may be helpful if one wishes to use the standard average of matter $< \rho >_o$ in the Friedmann equation, rather than the effective matter distribution $\bar{\rho}$. The explicit expression for $\bar{G}$ can be easily obtained by setting $\bar{\rho} =< \rho >_o$

in (3.163)

$$8\pi\bar{G} = 8\pi G + \frac{\bar{R} - < R >_o + < \tilde{K}^{ab}\tilde{K}_{ab} >_o - \frac{2}{3}(< k^2 >_o - \bar{k}^2)}{< \rho >_o}. \tag{3.169}$$

Notice however, that it is $G(\eta)\rho(\eta)$ which is inferred from measurements for different scales, and thus the use of $\bar{G}$ is not particularly remarkable.

More interestingly, it is more important in this connection to discuss the dependence of $G(\eta)\rho(\eta)$ as the local scale is varied, namely as η increases, (recall that η is the logarithmic change of the cutoff length associated with the geodesic ball coverings). For simplicity, we do this only for the case in which no shear is present ($\tilde{K}_{ab} = 0$), and the rate of volume expansion is spatially constant (k = const.). Under such hypothesis, we get for the scale-dependence of the average $< G\rho >$

$$\begin{aligned} \frac{\partial}{\partial\eta}[8\pi < G(\eta)\rho(\eta) >] &= \frac{\partial}{\partial\eta} < R(g(\eta)) >= \\ &= 2 < \tilde{R}^{ik}\tilde{R}_{ik} > + \frac{1}{3}\left(< R >^2 - < R^2 >\right), \end{aligned} \tag{3.170}$$

where $\tilde{R}_{ik} = R_{ik} - \frac{1}{3}g_{ik}R$ is the trace-free part of the Ricci tensor. From this expression we see that not only shear anisotropies but also metric anisotropies favour an increasing in $G(\eta)\rho(\eta)$ with the scale. To give an explicit example, let us consider as initial metric to be smoothed-out a locally homogeneous and anisotropic $SU(2)$-metric g. Following the notation and the analysis in the paper of Isenberg and Jackson [142], we can write such a metric and its Ricci-Hamilton evolution, in terms of a left-invariant one-form basis $\{\theta^a\}$, $a = 1, 2, 3$, on $SU(2)$ as

$$g = A(\eta)(\theta^1)^2 + B(\eta)(\theta^2)^2 + C(\eta)(\theta^3)^2, \tag{3.171}$$

where A, B, C are scale (η)-dependent variables. With respect to this parameterization, the scalar curvature is given by

$$R(\eta) = \frac{1}{2}\left[[A^2 - (B - C)^2] + [B^2 - (A - C)^2] + [C^2 - (A - B)^2]\right]. \tag{3.172}$$

While the squared trace-free part of the Ricci tensor is given by

$$\begin{aligned} ||\tilde{Ric}||^2 &= \frac{1}{6}\left[[A^2-(B-C)^2]^2+[B^2-(A-C)^2]^2+[C^2-(A-B)^2]^2\right] \\ &- \frac{1}{6}\left[[A^2-(B-C)^2][B^2-(A-C)^2]+[A^2-(B-C)^2][C^2-(A-B)^2]\right] \\ &+ \frac{1}{6}[B^2-(A-C)^2][C^2-(A-B)^2] \end{aligned} \tag{3.173}$$

The Ricci-Hamilton flow for this metric g exponentially converges to the fixed point $A = B = C = 1$, with the normalization $ABC = 1$, and with $A(\eta) \geq B(\eta) \geq C(\eta)$ for all η [142]. From the above expression for $R(\eta)$ it follows that $R(\eta)$ monotonically increases from its initial value $R(\eta = 0)$ and exponentially approaches

$$\lim_{\eta\to\infty} R(\eta) = \frac{3}{2}. \tag{3.174}$$

This increase is generated by the exponential damping of the anisotropic part of the Ricci tensor (3.174). This part of the Ricci tensor is smoothed by the Ricci-Hamilton flow and, roughly speaking, is redistributed uniformly in the form of scalar curvature.

Thus, if we smooth out the initial data set $(\Sigma \simeq S^3, g, K, \rho, \mathbf{J})$, with g the above $SU(2)$-metric, $K_{ab} = \frac{1}{3}g_{ab}k$, with $k = const.$, $\vec{J} = 0$, and with the matter density ρ such that the Hamiltonian constraint holds, we get that $G(\eta)\rho(\eta)$ monotonically increases as $\eta \to \infty$, exponentially approaching a fixed value.

As we shall see in the next section, this increase with the scale of the product $G(\eta)\rho(\eta)$ may be of significance in the correct interpretation of recent cosmological data.

At this stage, we are going to comment on an important point in the interpretation of the formalism, before turning to a discussion of its physical consequences.

The general picture arising from the above analysis is that we pick up an appropriate initial data set which, when evolved, gives rise to the *real* space-time. The description of this data set and of the resulting space-time is too detailed for being of relevance to cosmology. Intuitively, one would like to eliminate somehow all the unwanted (coupled) fluctuations of matter and space-time geometry on small scales, and thus extract the effective dynamics capturing the global dynamics of the original space-time. The possibility of actually implementing such an approach is strongly limited by the fact that we do not know *a priori* the structure of the space-time we are dealing with. But we may alternatively decide to handle the unwanted fluctuations at the level of data sets, since the time evolution of the initial data set for the Einstein equations is actually determined by the very constraints which that data set has to satisfy. As we have seen above, this can be done quite effectively in the framework of the Renormalization Group which is naturally well suited to this purpose. However, and here we come to the point we wish to make clear, *different initial data sets giving rise to the same inhomogeneous and anisotropic space-time, may yield smoothed data set giving rise to different FLRW space-times.* Obviously, one may have privileged initial data set with respect to which the Renormalization Group smoothing can be implemented. For instance a suitable slice of a frame comoving with matter, or of a frame minimizing the anisotropies in the CMB radiation. However optimal is this choice, it must be stressed that the time evolution of such data does not commute with the Ricci-Hamilton evolution. In other words, **the renormalization procedure and the dynamics do not commute.** The dynamics gives rise to the crossover between different FLRW space-times, or more generally, between different renormalized models of the same original irregular space-time.

This situation is in fact not so paradoxical as it may seem. From a thermodynamical point of view, we have seen that one of the members of the initial data set, namely the three-metric g, plays the rôle of a temperature. Thus, by varying the metric, one can move through the possible *pure phases* of the thermodynamical system considered. In this sense, a real, locally inhomogeneous universe, is to be considered as akin to a generic complex thermodynamical system. The possible locally homogeneous cosmological models, arising from it by suitable choices of initial data set to smooth-out, correspond to its *distinct pure phases.* The resulting dynamics yields, in an analogy with common statistical system, a *dynamical crossover between different pure phases.* This makes accessible in cosmology too, the whole subject of *critical phenomena* with a plethora of interesting consequences. Critical phenomena are always manifested macroscopically, as phase transitions are *collective* phenomena in their nature. This aspect may turn out to be of importance for the study of structure formation and clustering in the universe (cf. comments at the end of the previous section).

3.6.6 Cosmological implications

According to the contents of the previous sections the key idea on which our whole analysis rests is that of RG, namely, the involved physics is that of the running (scale dependence, be it energy or momentum scales) of the couplings and the relevant quantities accordingly. This philosophy is recognized and established in particle physics, e.g. it is well known that the fine structure constant α measured at low energies is different from the one measured at the LEP energy scale. In each case it is the presence of "fluctuations" (of any kind) that requires a scale dependent redefinition of the physical parameters which can, in turn, modify them, as well as the very structure of the theory, in a non-trivial way. Applications of the RG to particle physics have usually been in the ultra-violet limit (e.g. in QED, QCD, GUT) whereas in condensed matter physics they have been in the infra-red limit, in the study of critical phenomena and phase transitions.

We have taken this infra-red direction in cosmology. The application of the concept of running of the physical quantities, motivated by RG, appears to be a new important feature in a cosmological setting, providing (at least) partial explanation of some controversies of standard cosmology, which we are going to discuss below. Let us first point out that generally running, also, of cosmological quantities is as such motivated by the asymptotically free higher derivative quantum gravity, according to which the gravitational constant is asymptotically free [14]. Taking into account this fact, of running G, one can explore its consequences in the standard FLRW cosmology (cf. [114, 25]).

A word of caution is in order - since we do not have an ultimate theory of quantum gravity (QG), such approaches to cosmology are not on a rigorous basis from a theoretical point of view and should rather be taken as phenomenological. In principle, once we have a valid QG theory will one be able to directly derive a RG equations for various cosmological quantities. Although we have taken here a more standard view in which a split exists between the background (associated with infra-red effects) and renormalization of fluctuations (usually taken to be associated with ultra-violet effects), there may well be scales where such a split is not sensible at all. However, a RG capable of interpolating between the qualitatively different degrees of freedom in a parameter space of gravity, remains to be developed which may, after all, as well be possible after QG theory is within our reach.

One of the major issues in modern cosmology is concerned with the value of the Hubble constant and the apparent conflict between the observed age of the universe and the predicted one, in the standard FLRW model, based on the recent measurements of the Hubble constant. Namely, the recent measurements of the Hubble constant using the Virgo cluster (distance $\simeq$ 15 Mpc) strongly support that the Hubble constant H_o has somewhat larger value $h = 0.87 \pm .07$[15] [204]. At the same time, other distance indicators yield a systematically smaller value of H_o, e.g. around $0.55 \pm .08$ using NGC 5253 at a distance of 4 Mpc [222], while an analysis of the gravitational lensing of QSO 0957+561 indicates $h = 0.50 \pm 0.17$ [210]. When one calculates now the age of the universe, using the larger value of H_o, in the FLRW model one runs into a serious problem, as the predicted age turns out to be too small to accomodate the measured ages (~14-18 Gyr) of the globular clusters in our galaxy [221, 165].

Moreover, typical inferred values of the density parameter $\Omega_o = \rho_o/\rho_o^{crit}$ (ρ_o is the present value of the total energy density of the universe and ρ_o^{crit} its present critical energy density, defined as $\rho_o^{crit} = 3H_o^2/(8\pi G)$ where G is Newton's gravitational constant) increase correspondingly with the increasing size of various structures (e.g. [158]). These measurements can at most account for a fraction of Ω_o which according to the inflationary paradigm should be equal to 1. This in turn is one of the reasons for postulating the existence of non-baryonic Dark Matter (DM) which is also required to explain the structure formation.

Various people have since then looked at possible theoretical alternatives to these DM scenarios, such as e.g. introducing a cosmological constant in the Einstein equations or *ad hoc* modifications of the usual theory of gravity. The important point in this respect may as well be the one addressed here. It is usually taken for granted that, on large scales, the universe is described by the FLRW solution. There is no alternative really since we do not know any solution of Einstein's equations capable of describing a clumpy universe.

[15] h is H_o measured in units of 100 $km/sec/Mpc$

Nevertheless, even in the absence of explicit fine-grained models, we would like to know how in principle and when one could extract a background model from an inhomogeneous one, such that (*i*) they both obey, "approximately", the Einstein equations despite the averaging or smoothing involved, and (*ii*) observational determinations of cosmological parameters (H_o, Ω_o,...) correspond in a sensible way to that mathematical averaging procedure. Thus an issue of importance for cosmology [187], is the question on what scale is the FLRW model supposed to describe the universe? Likewise, what averaging scale are we referring to when we give the value of Ω_o, whose definition necessarily refers to an idealised, i.e. smoothed background model. There has been recently an increased effort in this direction with some interesting results, as e.g. that the coarse-graining effects could be non-negligible in the context of affecting the age of the universe. For example, [242] considered a model, of locally open (underdense) universe embedded in the spatially flat universe, in which the expansion rate in our local universe is larger than the global average. Similar model was considered in [193] where a local void in the global FLRW model was studied and the inhomogeneity described by the Lemaître-Tolman-Bondi solution. The results indicate that if we happened to live in such a void, but insisted on interpreting the observations by the FLRW model, the Hubble constant measurements could give results depending on the separation of the source and the observer, providing a possible explanation for the wide range of their reported values and capable of resolving the age-of-universe problem.

On the other hand [114] studied the QG effects at cosmological scales (the phenomenon in question is exactly that of quantum coherence, known from a laboratory to happen on macroscopic scales of the order of cm), assuming asymptotic freedom of the gravitational constant and incorporating running G, according to the appropriate RG equations, into FLRW model. Such G takes the value of Newton's constant G_N at short distances but then slowly rises as distance increases. However, as mentioned in [114], the RG equations used there might not be applicable in the infra-red regime studied, but these concerns were put aside having in mind absence of any other available beta functions for QG in the infra-red regime.

One can also approach the averaging problem, modifying the FLRW metric (or equivalently Cosmological Principle) and the Einstein equations by an introduction of a generalized scale factor which depends both on t and the scale r [155]. This introduces the scale dependence of $(G\rho)$ and running of other cosmological quantities such as H_o, Ω_o and the age of the universe t_o as functions of distance scales.

Let us notice that this picture of running of cosmological quantities comes about naturally in our approach which is physically motivated by RG, namely, due to increasing of the gravitational constant with scale (and possibly increasing amount of DM), as shown in the previous section, Ω_o has effectively increased too. Moreover, since the scale factor is governed by the scale dependent $(G\rho)$ it now depends on the scale r as well, i.e. it increases at any fixed time with the increasing distance. Consequently, the value of H_o is not the same everywhere in the observable universe and depends on what scale it is measured. Moreover, since $H_o \sim 1/t_o$ the universe becomes older when its age is estimated on a smaller scale. This is not to be taken as implying that the age depends on where one calculates (every observer using the same scale r at some time t_1 will obtain the same age). The key point to emphasize is that, having in mind the RG arguments and interpretation, a direct comparison of cosmological quantities makes sense only when they are measured (or calculated) with respect to the same scale, since the same quantity can take different values at different scales.

Notice that independently, also Quantum Cosmology advocates, though in a different context of bubble universes, a possibility that we may live in a universe in which the value of Hubble constant and the measured density are different in different places and in our local neighborhood Ω_o may well be less than 1 [173].

Thus the modified Friedmann equations suggested [114, 25, 154] can be written in the following form (locally $k = -1$)

$$\dot{R}^2 = \frac{8\pi}{3}\rho G(d) R^2 + 1, \tag{3.175}$$

$$G\rho = G\rho_o \left[\frac{R_o}{R}\right]^3, \tag{3.176}$$

$$\Omega_o(d_o) = \frac{G \cdot \rho_o}{G \cdot \rho_o^{crit}} = \frac{8\pi G(d_o)\rho_o}{3H_o^2(d_o)}, \tag{3.177}$$

where $G(d)$ is the value of G at (proper) distance d which can be interpreted as an inverse of momentum.

(Strictly speaking, one should define the Hubble constant as $H(t,r) = \dot{D}(t,r)/D(t,r)$ where $D(t,r) = \int_0^r R(t,r')dr'$ is the proper distance, but as mentioned in [155] one can confidently use the usual one, $H = \dot{R}/R$, for small redshifts.)

Some consequences of the above equations were investigated in [114, 25, 154]. Introducing

$$G(d) = G_N(1 + \delta_G(d)), \tag{3.178}$$

with $d(t,r) = R(t,r)10^7 r/(R(t_o,r)3)$ expressed in units of kpc [25, 154] ($r = 1$ corresponds $\sim$ to the size of the horizon and $r = 3 \cdot 10^{-7}$ to 1 kpc at present), we encode the variation of G in some growing function δ_G.

An interpretation of (3.177), in [114], enables to arrive at the conclusion of less or no DM needed in order to explain the observational data on Ω. No distance dependence of H_o was assumed there. An analysis of (3.178) in [25] leads in turn to the conclusion of increased power of the 2−point correlation function at large distances without a help of DM with $\Omega_o = 1$.

It is worth emphasizing that the running of G basically does not affect the early universe. If the relevant G calculated at the horizon scale at the time of, say, nucleosynthesis is close to G_N (basically the same as d is very small), (this is the case in [114]), the nucleosynthesis proceeds as usual. Likewise, other features of inflationary scenarios are not affected. This makes this proposal even more promising.

3.7 Concluding remarks

Whenever the macroscopic characteristics of the system which evolves dynamically are studied, coarse-graining is necessary[16] and relevant, when we are interested in its behaviour over rather short time scales. It basically means some kind of averaging over small but

[16] Another possibility can be long-time averaging, which is usually equivalent to statistical averaging over an *ensemble* of all possible micro-states that can produce given macroscopic features. This is appropriate for systems in a quasi-static state.

finite volumes of phase space. Coarse-graining does not change the original system, but the coarse grained quantities may behave qualitatively differently from the microscopic ones.

A particular framework to carry this out is in the Renormalization Group formalism. RG calculations are concerned with the asymptotic properties of critical systems, in the sense of (semi-)infinite size and proximity to the critical point and thus, they predict singular power law behaviour of the systems with universal exponents and scaling functions.

To put things in perspective, what we demonstrated here is that the same theoretical techniques can be applied to problems in Condensed Matter, Particle Physics and Gravitation theory. Among these, we have concentrated on the applications of Renormalization Group. On the one hand, there is a typical application of the RG to Condensed Matter Physics, in the infra-red limit in the description of critical phenomena and second-order phase transitions, and possibly the quantum Hall effect. On the other hand, in Particle Physics RG is usually applied in the ultra-violet regime, e.g. in QED to find the Landau pole, or in QCD - asymptotic freedom, while in the electro-weak sector it is used to extrapolate gauge couplings, measured at LEP, to higher energies to test consistency with some GUT.

We have put forward an application of the RG in classical relativistic cosmology to tackle the smoothing problem, which belongs to the various-length-scales class of problems and due to this, according to K.Wilson [256] can be grasped using RG methods. As such our theory can be looked at in parallel with Condensed Matter applications of RG since, also in our case, we were interested in the macroscopic effective characteristics of our system upon some coarse-graining.

Let us add that another application of RG is of relevance in Quantum Gravity, within the framework of string theory, to 2−dimensional theories on the world-sheet, where Zamolodchikov's c−theorem [264] may give insight into the effective string theory that hopefully describes this part of the universe.

Further, RG already found its way in the membrane approach to black holes, where its horizon is described in terms of a dynamical surface. Within path integral approach in the presence of this dynamical boundary, RG arguments yield a description of the dynamics of the horizon by the action of the relativistic bosonic membrane [180]. Also here a classical description of the state of a black hole (non-rotating) is characterized only by the "macroscopic" parameters, such as its mass and charge and there is an implicit coarse-graining over the membrane excitation levels. This membrane approach derived from a RG procedure gives an effective description valid at distances longer than l_{Planck}; the thermodynamic properties derived are in agreement with the standard results.

4 Dynamical systems approach in Cosmology

A dynamical system consists of a phase space which provides us with a description of the system's allowed states and a rule which defines the temporal evolution of those states. For differential equations the evolution is continuous, it can also be discrete as for a mapping. Virtually every model of physical phenomena is a dynamical system and, in fact, most of the models are Hamiltonian dynamical systems; GR belongs to the class of *constrained* Hamiltonian systems [127]. Hamiltonian dynamical systems give rise to symplectic mappings. Also, it is worth stressing the fact that the motion of a fluid particle in an incompressible fluid is Hamiltonian, no matter if the fluid motion itself is viscous, or not. Let us add that mappings are in fact more general than differential equations, but also they are easier to study than differential equations.

Typical questions of physical interest are concerned with the long-time stability of orbits and the determination of the regions accessible to the motion. Also problems concerning transport, i.e. the determination of the time for a bunch of trajectories to move from one region of phase space to another are of interest in physics. Even if the system were not strictly stable, it could be stable in practice if the transport times were longer than the lifetime of the system – this is probably the case with the planetary motions in the solar system, though certainly not so for asteroids.

In practice, the situation is idealized by considering (for positive times) the induced flow on the tangent bundle of the phase space, which is given by the structure of a smooth manifold, and by studying the growth of distances between nearby trajectories[1].

[1] This is normally done with respect to some riemannian metric in the tangent space as it moves along the trajectory.

A number of definitions is in order which will be given while we proceed.

4.1 Stability analysis of dust–radiation universe

The material presented below in this chapter is based on [47].

We consider flat and open universe models containing a mixture of cold matter (dust) and radiation interacting only through gravity and study their stability with respect to linear scalar perturbations. To this end the perturbed universe is considered as a dynamical system, described by coupled differential equations for a gauge-invariant perturbation variable and a relevant background variable.

4.1.1 Introduction

One of the main goals of present - day cosmology is to understand the formation of the structures (galaxies, clusters, superclusters) observed in the universe, while trying to explain why, on large enough scales, this seems to be so well described by the FLRW isotropic models. Given these latter, most theories of structure formation are based on the gravitational instability scenario. At any given epoch, there are perturbations larger than a certain characteristic - time dependent - scale[2]; while perturbations much smaller than this scale oscillate as sound waves, density perturbations on larger scales grow, eventually entering a non - linear regime during the matter dominated epoch, thus forming the observed structures. The mathematical basis for such scenario is the theory of perturbations of FLRW models.

In the inflationary scenario, the perturbations were generated from quantum fluctuations within the Hubble horizon H^{-1}, have evolved classically outside the horizon, and have re-entered it during the radiation or matter dominated epochs (perturbations with larger wavelength re-entering later). During the last decade, observations of the distribution of matter have shown that the scale at which the background homogeneity seems to be reached is larger than what was thought before, being of the order of hundreds of

[2]This can be the Jeans scale, or the Hubble radius (sometimes loosely referred to as horizon); a more important scale is actually the sound horizon, see [18].

Mpc[3]. Consequently, both the inflationary scenario and the large scale observations motivate the use of a fully relativistic theory of perturbations in FLRW models in order to study the formations of the larger structures.[4] Moreover, while the inflationary scenario seems to favour a flat universe, there is as yet no convincing empirical evidence for a critical density parameter $\Omega_0 = 1$, while a low density $\Omega_0 < 1$ universe is in fact favoured by many observations [62]. For this reason, we shall consider perturbations of flat and open models. The density contrast of the larger structures appears to be small enough that a linear perturbation analysis still suffices to describe the evolution of perturbations of the corresponding scale. Even if the density contrast is mildly non-linear, the curvature perturbations are still in the linear regime [19], [100], thus one can imagine the present universe as well described by a linearly perturbed FLRW model at large scales, while non-linearities at smaller scales can be considered smoothed-out in this picture.[5] Then the question arises if in this respect we live in a special epoch - an epoch in which large scale perturbations are still in the linear regime - or if this is a natural output in the theoretical context of perturbed FLRW models. This question is of the same sort as that posed by the "flatness problem": if the geometry of the universe is non–flat, then we live in a special epoch in which the density parameter is still close to unity: $\Omega_0 \sim 1$. It is well known that in a flat dust model there is a perturbation mode that grows unbounded, while in an open dust universe there is a mode that freezes in at an epoch $z \approx \Omega^{-1}$. In both cases the perturbation equation (4.1) is the same for all modes, because for $c_s^2 = w = 0$ (dust) the coefficient β (4.23) does not depend on the wavenumber k. It is usually said that in a dust universe each perturbation evolves as a separate FLRW universe. On the other hand, perturbation scales in a pure radiation model always come within the "sound

[3]The answer to the question "At which scale in the universe there is a transition to homogeneity?" depends on how the question is posed, i.e. how such a scale is defined. In a very loose sense we can say that this scale is measured by the size of the largest structures we see (of the order of 10^2 Mpc), and in a strict sense such a scale does not exist at all if there is not a cut-off in the perturbation spectrum. Various reasonable definitions of such a scale can be given (e.g. specifying a level for the fluctuations, such that we can say that there is a transition to homogeneity at the scale where the fluctuations fall below the specified threshold): cf. [224] and [79].

[4]Newtonian perturbation theory is applicable only for vanishing pressure and at scales much smaller than the horizon.

[5]In general, this again raises an issue of defining a proper averaging procedure – problem considered at length in this thesis.

horizon" [18] and oscillate as sound waves.

This simple picture gives however rather little information about the generality of this behaviour for the perturbations, and on the stability properties of the perturbed FLRW model. We aim therefore at studying the problem of the stability of FLRW models following an alternative approach: instead of looking for analytic solutions (either exact or approximate, see e.g. [203]) of the perturbation equations, we consider the perturbed universe as a dynamical system, described by coupled differential equations for a gauge-invariant perturbation variable and a relevant background variable. In this approach, the evolution of perturbations is represented by the trajectories in the phase space of the dynamical system, and their final fate is linked to the presence and the nature of the fixed (or stationary) points of the system.

In a certain sense the present work extends that of [78] to include perturbations, although we restrict our analysis to open and flat models only, and we take a vanishing cosmological constant $\Lambda = 0$. Previous works have followed the approach to stability used here [261], [43], but they only considered either a pure radiation or a pure dust model. However, since the analysis here is based on the study of the dynamical system at time-infinity, the pure radiation model does not appear physically significant in this limit, while the stability properties of the dust model are affected by the simplifying assumption of the vanishing of the speed of sound. Therefore here we consider a class of perturbed models containing both dust and radiation as a more realistic description of the real universe. The dust component can be taken to represent Cold Dark Matter (massive weekly interacting particles), while the radiation component represents photons and/or other massless particles such as massless neutrinos, therefore these two fluids interact only through gravity.

Universe models containing a mixture of dust and radiation interacting only through gravity were considered before (e.g. in [129]), as well as perturbations of these models (see e.g. [121] and [203]).

Here, our aim is to study the stability properties of these simple models, considering only the total density perturbation (the single component perturbations are not directly relevant to the evolution of curvature perturbations), assuming adiabatic perturbations.

For the case of flat universe models, as will see, we find that there exists a critical wavenumber k_{EC}, which is an invariant characteristic of the model and is related to the only scale entering the flat models, i.e. the Hubble radius at equi-density of matter and radiation H_E^{-1}. The corresponding critical scale λ_{EC} (λ_E is the perturbation wavelength at equi-density), remarks the transition from stability to instability, but in a way which is more rigorous - from the point of view of the stability analysis - than the Jeans or the Hubble scale. We find that there are actually three regimes for the evolution of fluctuations:

(i) growing large scale perturbations (unstable modes);

(ii) overdamped intermediate scale perturbations;

(iii) damped small scale wave perturbations;

where the transition scale from one regime to another is always of the order of λ_{EC}. Also, we show that λ_{EC} is of the order of the Jeans scale λ_{JE} at equi-density in the same model: however, since $\lambda_{EC} \simeq 2.2\lambda_{JE}$ (today $\lambda_{0C} \simeq 67Mpc$), our analysis shows that there are perturbation modes that decay, despite that their scale λ_E is larger than λ_{JE} at equi-density. Thus the evolution of perturbations in these models depends on their scale, in such a way that smaller scales evolve like in a pure radiation model (case (iii)) and larger scales like in a pure dust model (case (i)), while we found a small intermediate range of scales (case(ii)) for which perturbations are overdamped (critical damping occurs for the transition scale between cases (ii) and (iii)), which is an original feature of the dust–radiation models, and to our knowledge was not known before.

For the case of an open dust–radiation model instead, the evolution of perturbations appears to be dominated by the curvature of the background, and their final state is similar to that they have in a pure dust model, i.e. all the perturbation scales are frozen–in to a constant value. These models appear to be marginally stable [13] with respect to perturbations of any wavelength.

In the following we take $c = 1$, $\kappa = 8\pi G$ (G is the gravitational constant), and we assume a vanishing cosmological constant $\Lambda = 0$.

4.1.2 The dynamical system

We give below a brief outline of the method used to study the stability of FLRW models.

In a recent series of papers, Woszczyna and colleagues considered the dynamics of Newtonian [263] and relativistic [261] linearly perturbed universe models, studying the stability of these dynamical systems. The analysis of the relativistic case was however affected by a certain assumption on the allowed range for a scale parameter k (q in [261] and in [43]) and by a misinterpretation of the perturbation variables, as was shown in [43].

We shall focus now on relativistic perturbations of flat and open FLRW models containing two fluids coupled only through gravity: dust and radiation. Open and flat FLRW models expand for an infinite amount of time, and therefore one can apply standard stability criteria (see e.g. [12] and [13], and [8]). Specifically, one can establish if a cosmic dynamical system, i.e. a perturbed cosmological model, is: *stable* (in the sense of Lyapunov) around the unperturbed state, which roughly speaking means that a small change in the initial conditions do not produce a big change in the following evolution, i.e. in this context an initial small inhomogeneity do not grow; *asymptotically stable*, i.e. the unperturbed state is stable *and* the perturbation decay so that the cosmic system returns to the original homogeneous state; *marginally stable*, when is stable but not asymptotically stable; *unstable*, when the perturbations grow with time.

In general, in the gauge- invariant approach to cosmological perturbations the scalar[6] density perturbation variable (we shall consider specific perturbation measures later) satisfy a second order (in some time variable) differential equation. If one restricts the attention to the harmonic component X of the perturbation, and assumes that this is adiabatic (see section 4.1.3.2), its evolution is given by a homogeneous ordinary differential equation

$$\ddot{X} + \alpha(t)\dot{X} + \beta(t)X = 0 , \tag{4.1}$$

where t here is proper time, and the dot indicates a derivative with respect to t. When it is not possible to find a simple solution to (4.1), a qualitative analysis of its properties

[6]It is standard to call scalar perturbations those related with density perturbations describing the clumping of matter (e.g. see [236]). Here we shall consider only these perturbations, as the only relevant to the problem of stability of the universe.

is useful in order to determine the late - time behaviour of the perturbations[7]. The coefficients α and β are functions given by the background dynamics, but in general their time dependence cannot be explicitly determined. Therefore, it is useful to think of α and β as known functions of one or more parameters, and add to (4.1) the evolution equation for the parameters in order to have an autonomous system. A sensible choice followed in [261] is given by the density parameter Ω, which in a FLRW model satisfies the equation

$$\dot{\Omega} = \Omega(\Omega - 1)(\tfrac{1}{3} + w)\Theta \,, \tag{4.2}$$

where $\Theta = 3H = 3\dot{a}/a$ is the expansion of the cosmic fluid, a is the FLRW scale factor, H is the Hubble parameter, and $w = p/\mu$ is the ratio of the pressure to the energy density. It is useful to change the independent variable from the proper time t to a function of the scale factor a; with the choice of [8] $\tau = \ln a^3$, $(\frac{d\tau}{dt} = \Theta)$ equations (4.1) and (4.2) give

$$X'' + \psi X' + \xi X = 0 \,, \tag{4.3}$$

$$\Omega' = \Omega(\Omega - 1)(\tfrac{1}{3} + w) \,, \tag{4.4}$$

and in general

$$w' = -(1+w)\left[c_s^2(w) - w\right] \tag{4.5}$$

gives the evolution of w, and in the previous equations the prime refers to the derivative with respect to $\tau = \ln a^3$.

In a single fluid FLRW model the dynamics is fixed by an equation of state $p = w\mu$, with $w = const.$ (e.g. $w = 0$ dust, $w = 1/3$ radiation), and in this case $c_s^2 = \dot{p}/\dot{\mu}$ is the speed of sound, and $c_s^2 = w$. In dealing with two or more fluids however, $w' \neq 0$, and $c_s^2 \neq w$ is no longer a proper speed of sound unless the fluids are coupled, and ψ and ξ are given by

$$\psi(w,\Omega) = \frac{\dot{\Theta}}{\Theta^2} + \frac{1}{\Theta}\alpha(w,\Omega) = -\frac{1}{3} - \frac{1}{6}(1+3w)\Omega + \frac{1}{\Theta}\alpha(w,\Omega) \,, \tag{4.6}$$

[7] One can obviously find numerical solutions, but the study of the phase space allows us to obtain general conclusions for a generic set of initial conditions.

[8] There are various typos in [261]: there, the power appearing here in the definition of τ is missed in equation (3), and consequently a factor 1/3 is missed in equation (9).

$$\xi(w,\Omega) = \frac{1}{\Theta^2}\beta(w,\Omega)\,, \tag{4.7}$$

while the second step in (4.6) is given by the Raychaudhuri equation

$$\dot{\Theta} + \tfrac{1}{3}\Theta^2 + \tfrac{1}{2}\kappa\mu(1+3w) = 0\,, \tag{4.8}$$

governing the evolution of Θ (see e.g. [74]). In the simplest case of one single fluid [261], [43] with $w = const.$ we have a third order autonomous system, given by (4.3) and (4.4) (or a corresponding pair of first order equations); also, equation (4.4) forms an autonomous first order subsystem in this case. In the most general case we have a fourth order autonomous system, with (4.5) as autonomous subsystem (as we shall see in section 4.1.3.1, $c_s^2 = c_s^2(w)$ is fixed once the fluid components are specified). The order of the system can however be reduced. First, it turns out that in the practical case (see section 4.1.3.1) either w or Ω can be eliminated; second, here we are not really interested in the evolution-law of X, but rather in its qualitative behavior. Because of this, we can achieve a further dimensional reduction of the phase space passing to the Riccati equation corresponding to (4.3). Introducing $Y = X'$ and

$$\mathcal{U} = Y/X = X'/X\,, \tag{4.9}$$

$$\mathcal{R} = \sqrt{X^2+Y^2}\,, \tag{4.10}$$

we pass in the new phase space $\{\mathcal{R},\mathcal{U},w,\Omega\}$, where

$$\frac{\mathcal{R}'}{\mathcal{R}} = \frac{\mathcal{U}(1-\xi-\psi\mathcal{U})}{\mathcal{U}^2+1}\,, \tag{4.11}$$

$$\mathcal{U}' = -\mathcal{U}^2 - \psi\mathcal{U} - \xi\,, \tag{4.12}$$

while the evolution of Ω and w is still given by (4.4) and (4.5). The variable $\mathcal{R}$ represent an "amplitude" of the perturbation and it is not directly relevant to the present analysis. Since (4.12) and (4.4), (4.5) form an autonomous subsystem, one can restrict the analysis to the phase space $\{\mathcal{U},w,\Omega\}$; moreover, as we said above, in practical cases we can restrict our attention either to $\{\mathcal{U},w\}$ or to $\{\mathcal{U},\Omega\}$. The relevant variable here is $\mathcal{U}$: when it is positive we have either a growing density enhancement or an increasing energy deficit, while $\mathcal{U} < 0$ indicates that the inhomogeneity is decreasing (note that from its definition (4.9) $\mathcal{U}$ is a tangent in the original phase space $\{X, Y\}$, thus $-\infty < \mathcal{U} < \infty$). One is therefore interested in the nature of the fixed (or stationary) points (if any) on the $\Omega = 0$ or the $w = 0$ axis (the final state of the cosmological dynamical system). These correspond to the real roots of the right hand side of (4.12). A stable node on the $\mathcal{U} > 0$ semi-axis will indicate that the given perturbation mode will indefinitely grow, thus indicating instability of the underlying cosmic system; a stable node on the $\mathcal{U} < 0$ semi-axis will indicate that the given perturbation mode will finally decay, i.e. the system is asymptotically stable with respect to that perturbation; a stable node on $\mathcal{U} = 0$ means that the perturbation will asymptotically approach a constant value: the cosmic system is then marginally stable (see [13]). Finally, if there are no stationary points on the $\Omega = 0$ (or the $w = 0$) axis, the perturbation maintains a sound wave character at any time. This situation corresponds to complex roots of the right hand side of (4.12) with negative real part, and we shall see in section 4.1.4.4 that these roots are also the eigenvalues of the system given by (4.12), (4.5): therefore the cosmic system is stable against these perturbations.

A discussion of the system (4.12), (4.4) can be found in [261] (and reference therein); a discussion of the flaws of the application of the analysis given in [261] to the perturbation equations of [18] and [86], [87], [45] (see also [262]) is given in [43].

4.1.3 Dynamics of the dust - radiation models

We shall now apply the general method outlined in the previous section to the case of uncoupled dust and radiation. In the following, we shall normalize the scale factor at equi-density of dust and radiation, introducing $S = a/a_E$. An analysis of the phase–space of FLRW models containing dust and radiation has been recently given in [78] (see also [129] and [203]); here we simply review well known results, that are needed for the perturbation analysis, with emphasis on a useful parameterization.

4.1.3.1 The background

Since the two fluids are uncoupled, we have separate energy conservation, with $\mu_d = \frac{1}{2}\mu_E S^{-3}$ and $\mu_r = \frac{1}{2}\mu_E S^{-4}$, where μ_E is the *total* energy density at equi-density. Then the total energy is also conserved, with density $\mu = \frac{1}{2}\mu_E(S^{-3} + S^{-4})$, while the total pressure is that of the radiation component, $p = p_r = \frac{1}{3}\mu_r = \frac{1}{6}\mu_E S^{-4}$; therefore

$$w = \frac{1}{3(S+1)} . \tag{4.13}$$

For two non-interacting fluids the quantity $c_s^2 = \frac{\dot{p}}{\dot{\mu}}$ is only formally the speed of sound: for dust and radiation we have

$$c_s^2 = \frac{4}{3(4+3S)} . \tag{4.14}$$

At equi-density $S = 1$, $w = 1/6$, and $c_s^2 = 4/21$. Clearly, (4.13) can be inverted to give

$$S = \frac{1-3w}{3w} , \tag{4.15}$$

and then the expansion of the universe model we consider is parameterized by w, with w $(1/3 \geq w \geq 0)$ varying from a pure radiation dominated $(t \to 0)$ to a pure matter dominated $(t \to \infty)$ phase.

Note that (4.13), with (4.14) is in practice an integral for (4.5); also, given (4.13) we can integrate (4.2):

$$\Omega = \frac{\Omega_E(S+1)}{\Omega_E(S+1) + 2(1-\Omega_E)S^2} ; \tag{4.16}$$

since today $S_0 = a_0/a_E = 1 + z_E \gg 1$, the density parameter at equi-density is $\Omega_E \simeq \left(1 + \frac{(1-\Omega_0)}{2S_0\Omega_0}\right)^{-1}$. Clearly, equation (4.16) can also be inverted to parameterize the expanding model with Ω. However in the following we shall find convenient the parameterization in w, which will allow us to give a unified treatment of flat and open models: from (4.15) and (4.14) we have

$$c_s^2 = \frac{4w}{3(1+w)} , \tag{4.17}$$

and this with (4.5) give

$$w' = w\left(w - \tfrac{1}{3}\right) , \tag{4.18}$$

where hereafter the prime stands for the derivative with respect to $\tau = \ln S^3$.

From (4.18) and (4.4) we get

$$\Omega = \frac{3\Omega_E w}{3\Omega_E w + 2(1-\Omega_E)(1-3w)^2} , \tag{4.19}$$

which can be used in expressions for ξ and ψ to obtain these coefficients as functions of w only, thus reducing the effective phase space needed for the stability analysis to

$\{\mathcal{U}, w\}$, i.e. that of the plane autonomous system (4.12), (4.18). We note also that the system (4.18), (4.4) that describes the evolution of the background universe model is a plane autonomous subsystem in the full space $\{\mathcal{U}, \Omega, w\}$: this will always be a property of a dynamical system describing a perturbed universe, since the basic assumption of the perturbation analysis is that of neglecting the backreaction of the perturbations on the dynamics of the background model.

Finally, the total energy density is given by

$$\mu = \frac{27}{2}\mu_E \frac{w^3}{(1-3w)^4} \,. \tag{4.20}$$

This obviously also follows from the conservation equation $\mu' = -\mu(1+w)$, integrated using (4.18); and $\mu \to \infty$ as $w \to 1/3$, $\mu \to 0$ as $w \to 0$.

4.1.3.2 The perturbed model

We shall now consider the dynamics of the perturbed dust - radiation models.

First, we have to specify a perturbation measure X, i.e. the variable appearing in (4.1) and (4.3). Here, we shall focus on density perturbations, and we shall take $\Delta = a^2({}^{(3)}\nabla^2\mu)/\mu$ as our fundamental variable, as was originally defined in [87] (see also [262]) following the covariant approach to perturbation introduced in [86].

This covariant quantity is an exact measure of density inhomogeneity, it is scalar and locally defined. With respect to a FLRW background Δ is a gauge - invariant variable that, once expanded at first order in the perturbations is proportional to the Bardeen variable ε_m (see [45]); thus its components with respect to an orthonormal set of scalar harmonic functions are proportional to the density perturbation in the comoving gauge (see [18]; for a comprehensive treatment of perturbations in this gauge, see [168], see also [46]).

In the following, we shall restrict our attention to the harmonic components of Δ, where the scalar harmonics Q are defined by

$${}^{(3)}\nabla^2 Q = -\frac{k^2}{a^2} Q \,, \tag{4.21}$$

${}^{(3)}\nabla^2$ is the Laplace operator in the 3-surface of constant curvature and $k \geq 0$ for $K = 0$, but $k \geq 1$ for $K = -1$ [128], [172]. Thus in the flat case the wavenumber k is simply related to the physical scale of the perturbation i.e. its wavelength $\lambda = 2\pi a/k$, but this is not the case for the open models. However k can always be taken as invariantly characterizing the scale of the perturbation, and we shall do so.

Since the harmonic component $\Delta^{(k)}$ of Δ and ε_m are just proportional [45], in the following the perturbation measure X appearing in (4.1) and (4.3) can be identified either with $\Delta^{(k)}$ or with ε_m.

In the case of a mixture of dust and radiation, in general the evolution equation for X is coupled to the evolution equation for an entropy perturbation variable (see [156]; [74]; [177] and also [203]). However the coupling is important only at small scales: here we shall only focus on perturbations at scales of the order of or larger than the Hubble radius H_E^{-1} at equi-density, thus we shall restrict to purely adiabatic modes, i.e. solutions of (4.1), neglecting the coupling with the entropy perturbation.

In general the coefficients α and β in (4.1) are given by (see [87]; [45]; [203])

$$\alpha = (2 + 3c_s^2 - 6w)H \ , \tag{4.22}$$

$$\beta = -\left[\left(\tfrac{1}{2} + 4w - \tfrac{3}{2}w^2 - 3c_s^2\right)\kappa\mu + 12(c_s^2 - w)\frac{K}{a^2}\right] + c_s^2\frac{k^2}{a^2} \ ; \tag{4.23}$$

using (4.17) for the dust - radiation background we get

$$\alpha = 2\frac{(1-3w^2)}{(1+w)}H \ , \tag{4.24}$$

$$\beta = -\left[\frac{3}{2}\left(1 + \frac{w^2(5-3w)}{(1+w)}\right)\Omega H^2 + 4\frac{w(1-3w)}{(1+w)}\frac{K}{a^2}\right] + \frac{4w}{3(1+w)}\frac{k^2}{a^2} \ . \tag{4.25}$$

From this, we obtain

$$\psi(w,\Omega) = \frac{1}{6}(1+3w)(1-\Omega) + \frac{1 - 6w - 15w^2}{6(1+w)} \ , \tag{4.26}$$

$$\xi(w,\Omega) = -\frac{1}{6}\left(1 + \frac{w^2(5-3w)}{1+w}\right)\Omega + \frac{4w(1-3w)}{9(1+w)}(1-\Omega) + \Xi_K(w,k) \ , \tag{4.27}$$

where $\Xi_K(w,k) = \frac{4w}{27(1+w)}\frac{k^2}{a^2H^2}$ is a function that takes a different form, as function of w, depending on the curvature K: in an open universe $a^2H^2 = (1-\Omega)^{-1}$, but in the flat case we cannot use the Friedmann equation to substitute for a^2H^2. Rather, we use it to substitute for $3H^2 = \kappa\mu$, and μ is given by (4.20), while here $a = Sa_E$ and S is given by (4.15). Thus

$$\begin{array}{lll} K=0 & \Xi = \frac{(1-3w)^2}{1+w}\Xi_0 k^2\,, & \Xi_0 = \frac{8}{81a_E^2H_E^2}\,, \\ K=-1 & \Xi = \frac{4w}{27(1+w)}(1-\Omega)k^2\,, & \end{array} \tag{4.28}$$

where, $3H_E^2 = \kappa\mu_E$, and for open models we can substitute for $\Omega = \Omega(w)$ from (4.19) in all the previous expressions. It is clear from (4.28) that in the space $\{\mathcal{U}, w, \Omega\}$ the function Ξ_K is not continuous (for $k \neq 0$) on the plane $\Omega = 1$, except on the line $w = 1/3$, $\Omega = 1$; this discontinuity of Ξ_K will play an important rôle in the behaviour of perturbations.

4.1.4 Results

We shall now summarize the results we obtained from the analysis of the dynamical system given by (4.12), (4.18) and (4.4), with coefficients ψ and ξ given by (4.26) and (4.27), first restricting to the flat models and then considering the open ones.

4.1.4.1 The flat models

The flat perturbed models are described by the subsystem (4.12), (4.18) substituting $\Omega = 1$ in ψ (4.26) and ξ (4.27) and using Ξ for $K = 0$ given in (4.28). From this latter we see that for flat models there is a characteristic scale that, as we shall see, is related to the late time behaviour of the perturbations: this is the Hubble radius at equi-density H_E^{-1}. It is therefore convenient to define

$$k_E = \frac{k}{a_E H_E} = 2\pi\frac{H_E^{-1}}{\lambda_E}\,, \tag{4.29}$$

which represents a wavenumber normalized at equi-density: $k_E = 2\pi \Leftrightarrow \lambda_E = H_E^{-1}$ i.e. $k_E = 2\pi$ corresponds to a perturbation wavelength that enters the horizon at equivalence epoch.

Stationary points eventually exist on the $w = 1/3$ and $w = 0$ axis, with

$$\mathcal{U}_\pm = \tfrac{1}{2}(-\psi_w \pm \sqrt{\psi_w^2 - 4\xi_w})\,, \tag{4.30}$$

where ψ_w, ξ_w (with $w = 0, 1/3$) correspond to the stationary values of ψ and ξ.

The stationary points for the system (4.12), (4.18) are given in Table 4.1.4.1.

POINT I: UNSTABLE NODE $w = 1/3, \quad \mathcal{U}_- = -\frac{1}{3}$	POINT II: SADDLE $w = 1/3, \quad \mathcal{U}_+ = \frac{2}{3}$
$\lambda_{\mathcal{U}} = 1$ $\lambda_w = \frac{1}{3}$	$\lambda_{\mathcal{U}} = -1$ $\lambda_w = \frac{1}{3}$
POINT III: SADDLE $w = 0, \quad \mathcal{U}_- = -\frac{1}{12} - \frac{1}{2}\lambda_{\mathcal{U}}$	POINT IV: STABLE NODE $w = 0, \quad \mathcal{U}_+ = -\frac{1}{12} - \frac{1}{2}\lambda_{\mathcal{U}}$
$\lambda_{\mathcal{U}} = \sqrt{\frac{2}{3}\left(\frac{25}{24} - \frac{k_E^2}{k_{EC}^2}\right)}$ $\lambda_w = -\frac{1}{3}$	$\lambda_{\mathcal{U}} = -\sqrt{\frac{2}{3}\left(\frac{25}{24} - \frac{k_E^2}{k_{EC}^2}\right)}$ $\lambda_w = -\frac{1}{3}$

Table 4.1: Flat models: the four stationary points with the corresponding eigenvalues along the axis $\mathcal{U}$ and w and their nature; Point III and IV exist only for $k_E \leq \frac{5}{2\sqrt{6}} k_{EC}$.

As we have explained at the end of section 4.1.2, the condition for having growing perturbations is given by the appearance of a stable node on the positive side of the $\mathcal{U}$ axis. From (4.30) we see that we need $\xi_0 \leq 0$ in order to have $\mathcal{U}_+ \geq 0$, i.e.

$$k_E \leq k_{EC} \,, \qquad k_{EC} = \frac{3\sqrt{3}}{4} \,. \tag{4.31}$$

Moreover, stationary points only exist for $\psi_w^2 - 4\xi_w \geq 0$, which is always satisfied only for $w = 1/3$. For $w = 0$ we have $\psi_0^2 - 4\xi_0 = \frac{2}{3}(\frac{25}{24} - k_E^2/k_{EC}^2)$, i.e. stationary points exist on this axis for perturbations of wavelength at equivalence $\lambda_E \geq \frac{2\sqrt{6}}{5}\lambda_{EC} < \lambda_{EC}$. We see from Table 4.1.4.1 that this is the same condition for the reality of the eigenvalues, a result that follows from the fact that w' does not depend on $\mathcal{U}$. From the values of these eigenvalues we have that: Point I is an unstable node, Point II is a saddle, Point III is a saddle, Point IV is a stable node.

When for this latter we have $\mathcal{U}_+ > 0$, i.e. for k_E smaller than the *critical* wavenumber k_{EC}, generic perturbations with physical wavelength $\lambda_E > \lambda_{EC}$ *grow unbounded*, with the critical perturbation scale λ_{EC} corresponding to k_{EC} in (4.31), given by $\lambda_{EC} = \frac{8\pi}{3\sqrt{3}} H_E^{-1}$, i.e. $\lambda_{EC} > H_E^{-1}$: this situation is illustrated in Fig. 4.1. Note in this figure the special *saddle trajectories*: the one ending in $\mathcal{U}_-$ represents the purely decaying mode, while the other (ending in $\mathcal{U}_+$ as the generic trajectory) represents the purely growing mode[9]

[9] The terminology *growing* and *decaying* is purely conventional: it is adopted here because it is standard in the literature to refer in this way to the two modes, as they are effectively growing and decaying e.g. in a flat pure dust model. We stress again that the "growing" mode is actually growing only when $\mathcal{U}_+ > 0$.

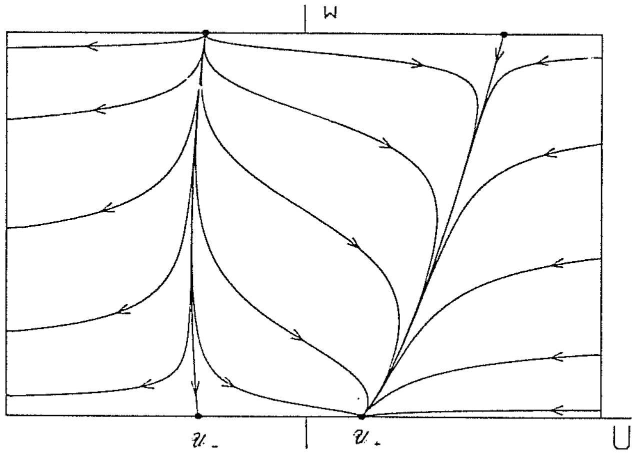

Figure 4.1: Flat models: phase space for the evolution of large scales $\lambda_E > \lambda_{EC}$ that grow unbounded ($\mathcal{U}_+ > 0$). Here and in the other figures dots represent the stationary points, while $0 \leq w \leq 1/3$ and $-1 \leq \mathcal{U} \leq 1$.

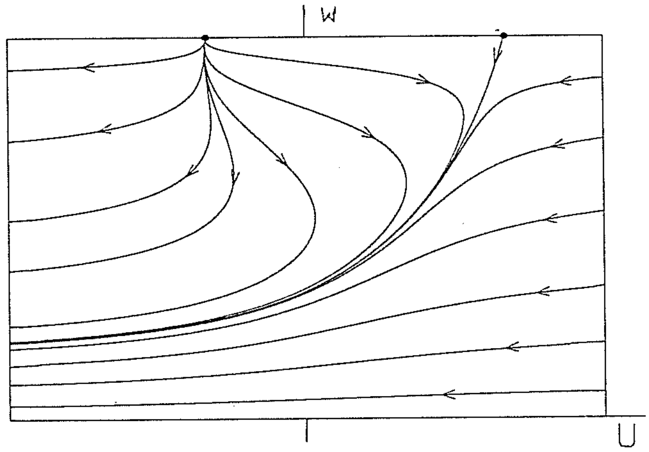

Figure 4.2: Flat models: phase space for the evolution of the damped wave modes on small scales $\lambda_E < \frac{2\sqrt{6}}{5} < \lambda_{EC}$ ($\mathcal{U}_\pm \in \Im$).

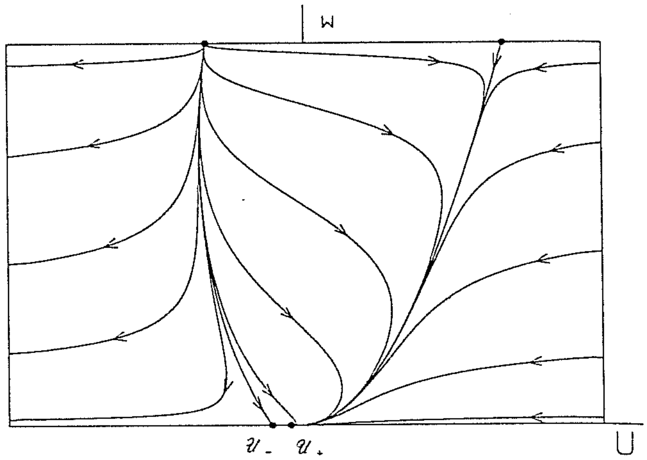

Figure 4.3: Flat models: phase space for the overdamped modes ($\mathcal{U}_+ < 0$) on intermediate scales $\frac{2\sqrt{6}}{5}\lambda_{EC} < \lambda_E < \lambda_{EC}$.

(the attractor); the generic trajectory represents a linear combination of the two modes solutions of (4.3). Also, we remind that the variable $\mathcal{U}$ is a tangent in the original phase space $\{X, X'\}$, so that the trajectories in the figures that start from Point I and go to the left of the saddle trajectory exit the figure on the left boundary, and re-enter it from the right ending in Point IV.

Clearly, when Point IV does not exist, i.e. for perturbations with wavelength at equivalence $\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$, we are considering perturbations that always *oscillate* as "sound waves" (see Fig. 4.2); we have checked through a direct stability analysis of the system (4.37) below (see section 4.1.4.4) that the amplitude of these modes decay. Again, trajectories that exit from the left re-enter from the right: for $\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$ however this happens many times, and indeed it is this fact that characterizes the oscillatory behaviour of these modes in the $\{\mathcal{U}, w\}$ plane.

Also, for scales $k_{EC} < k_E < \frac{5}{2\sqrt{6}}k_{EC}$ we have $\mathcal{U}_+ < 0$ for Point IV, i.e. the stable node is located on the negative side of the $\mathcal{U}$ axis: thus perturbation scales in this range are overdamped: they decay without oscillating, as it appears from Fig. 4.3. Again, the saddle trajectories represent the two modes for (4.3), and the trajectories going out from the left of Fig. 4.3, re-enter from the right, ending in Point IV.

Finally, we point out that for $k_E = \frac{5}{2\sqrt{6}}k_{EC}$ the two stationary points on the $\mathcal{U}$ axis (the saddle and the stable node in the second line of Table 4.1.4.1) coincide, i.e. this is a *fold bifurcation point* (see e.g. [13]). The corresponding scale $\lambda_E = \frac{2\sqrt{6}}{5}\lambda_{EC} = \frac{16\sqrt{2}}{15}\pi H_E^{-1} \simeq 4.7H_E^{-1}$ is quite larger than the Hubble radius at equi-density; thus we can consider scales $\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$, as in Fig. 4.2, such that still $\lambda_E > H_E^{-1}$: this partially justifies our assumption of purely adiabatic perturbations (cf. [196]).

Thus we have three different evolution regimes for the adiabatic perturbation modes of a mixture of uncoupled dust and radiation in a flat universe, depending on their wavelength:

(i) large scale perturbations that grow unbounded, giving instability;

(ii) intermediate scale perturbations that are overdamped, i.e. decaying without oscillating;

(iii) small scale damped perturbations which oscillate like sound waves while their am-

INSTABILITY	STABILITY	
$\lambda > \lambda_{EC}$ grow unbounded	$\frac{2\sqrt{6}}{5}\lambda_{EC} < \lambda_E < \lambda_{EC}$ overdamped	$\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$ damped waves

Table 4.2: Flat models: summary of the three different evolution regimes for different perturbation wavelengths.

plitude decays;

as summarized in Table 4.2.

4.1.4.2 The open models

The open models are in principle described by the full 3-dimensional system given by (4.12), (4.18) and (4.4), with trajectories in the $\{\mathcal{U}, w, \Omega\}$ space. However, given $\Omega = \Omega(w, \Omega_E)$ (4.19), we can substitute for Ω in the expressions for ψ (4.26), ξ (4.27) and Ξ for $K = -1$ (4.28), and restrict our analysis to the 2-dimensional system (4.12), (4.18). In doing this, we are selecting one particular open model in the class parameterized by Ω_E: there is no loss of generality in doing this, as the dynamical properties of the models in this class (as specified by the character of the stationary points in the corresponding phase space) do not depend on Ω_E, i.e. for a given wavenumber k the phase space evolution of all the models in the class is qualitatively the same. From the point of view of the geometry of the phase space $\{\mathcal{U}, w, \Omega\}$ we are looking at the trajectories in the 2-dimensional surface specified by Ω_E in this space: we shall then consider the projection of this surface with its trajectories in the $\{\mathcal{U}, w\}$ plane, as depicted in Fig. 4.4.

It is clear from this figure and Table 4.3 that, contrary to what we have seen for the flat models, the existence, position and the nature of Points I–IV do not depend on the wavenumber k, so that all open models share the same dynamical history. This depends on the vanishing of the function Ξ in the limit $w \to 0$, which also implies $\Omega \to 0$ (because we are moving on the surface specified by Ω_E). Note however that if we do not substitute for Ω from (4.19) in the 3-dimensional system, and we take the limit $\Omega \to 0$, this does not give the 2-dimensional system for flat models: this fact is due to the discontinuity (remarked at the end of section 4.1.4.1) of the function Ξ on the surface $\Omega = 1$, where Ξ

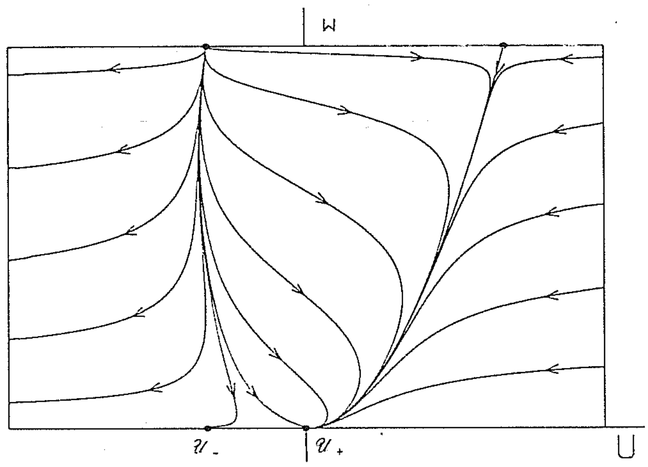

Figure 4.4: Open models: the phase space topology is independent of the wavelength, and all perturbations tend to a constant value ($\mathcal{U}_+ = 0$).

for $K = 0$ do not vanish in general, except for $w = 1/3$ or $k = 0$. Somehow this fact can be seen as an example of "fragility" in cosmology [63].

From table 4.2 and Fig. 4.4 we see that Point IV is a stable node located at $\mathcal{U} = 0$. The generic trajectory ends up in this point, either directly from Point I, or first going out from the left boundary of the picture and then re-entering from the right. Again, the saddle trajectory starting form Point I and ending in Point III represents the evolution of the purely decaying mode, and the saddle trajectory starting from Point III and ending in Point IV represents the purely growing mode (attractor). The fact that Point IV is located in $\mathcal{U} = 0$ means that (by definition of $\mathcal{U}$) all the perturbation modes evolve up to a constant value, and then freeze–in. Then we can say that open models are stable but not asymptotically stable, as the generic perturbation modes do not decay, so they are marginally stable.

A direct analysis of the system (4.37) below (section 4.1.4.4) shows indeed that the appearance of a fixed point on $\mathcal{U} = 0$ for the system (4.12), (4.18) corresponds to the vanishing of one of the eigenvalue for the corresponding fixed point in the phase space $\{X, X', w\}$, while the other two eigenvalues are negative. Thus there is a whole line of fixed marginally stable points.

While this happens only for $k_E = k_{EC}$ in flat models, it is a generic characteristic for any k in open universes. A comparison of Fig. 4.4 and Fig. 1a in [43] shows that open dust-radiation models and open pure dust models share the same stability properties. It appears

POINT I: UNSTABLE NODE $w = \frac{1}{3},\ \mathcal{U}_- = -\frac{1}{3}$	POINT II: SADDLE $w = \frac{1}{3},\ \mathcal{U}_+ = \frac{2}{3}$
$\lambda_{\mathcal{U}} = 1$ $\lambda_w = \frac{1}{3}$	$\lambda_{\mathcal{U}} = -1$ $\lambda_w = \frac{1}{3}$
POINT III: SADDLE $w = 0,\ \mathcal{U}_- = -\frac{1}{3}$ $\lambda_{\mathcal{U}} = \frac{1}{3}$ $\lambda_w = -\frac{1}{3}$	POINT IV: STABLE NODE $w = 0,\ \mathcal{U}_+ = 0$ $\lambda_{\mathcal{U}} = -\frac{1}{3}$ $\lambda_w = -\frac{1}{3}$

Table 4.3: Open models: the four stationary points with the corresponding eigenvalues along the axis $\mathcal{U}$ and w and their nature.

that the curvature of the background dominates also the evolution of the perturbations: for dust radiation models there is a conspiracy between curvature and pressure, such that curvature has an opposite effect on small and large scale: with respect to the behaviour in flat models, it avoids the oscillation of small scales perturbations ($\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$) and damps the growth of those on large scales ($\lambda_E > \frac{2\sqrt{6}}{5}\lambda_{EC}$).

4.1.4.3 Comparison with Jeans instability

We can apply the Jeans instability criterion directly to equation (4.1): this implies that gravitational collapse of a given perturbation mode will occur if $\beta < 0$, i.e. if $k < k_J$, where for the dust–radiation models β is given by (4.25), and the corresponding k_J is

$$\frac{k_J^2}{a^2} = \frac{3(1+w)}{4w} \cdot \left[\frac{3}{2}\left(1 + \frac{w^2(5-3w)}{1+w}\right)\Omega H^2 + \frac{4w(1-3w)}{1+w}\frac{K}{a^2}\right] . \tag{4.32}$$

At equi-density ($w = 1/6$) this gives

$$\frac{k_{JE}^2}{a_E^2} = \frac{279}{32}\Omega_E H_E^2 + \frac{3}{2}(\Omega_E - 1)H_E^2 ; \tag{4.33}$$

in the real universe $\Omega_E \approx 1$ and the contribution to k_{JE} from the curvature term (the last in the equation above) is completely negligible. Thus in a flat universe ($K = 0$) we have

$$\lambda_{JE} \equiv \frac{2\pi a_E}{k_{JE}} = \frac{8\pi}{3}\sqrt{\frac{2}{31}} H_E^{-1} \tag{4.34}$$

for the Jeans scale at equi-density. Then a comparison with the critical scale λ_{EC} defined in section 4.1.4.1 gives

$$\lambda_{JE} = \sqrt{\frac{6}{31}}\lambda_{EC} , \tag{4.35}$$

i.e. for flat models the stability criteria used in section 4.1.4.1 give a critical scale λ_{EC} for instability which is larger than the corresponding Jeans scale at equi-density by a factor of 2, i.e. $\lambda_{EC} \approx 2.3\,\lambda_{JE}$. The fact that the values of these two scales are relatively close appears physically meaningful, and in a way obvious, since both scales are somehow defined through the same differential equation. However, a comparison of (4.35) with the analysis of section 4.1.4.1 shows that there are perturbations with $\lambda_E > \lambda_{JE}$ that decay: those with wavelength $\lambda_{JE} < \lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$ are damped oscillation, while those with $\lambda_{EC} > \lambda_E > \frac{2\sqrt{6}}{5}\lambda_{EC}$ are overdamped. It is immediate to show that today the critical scale corresponding to k_{EC} is

$$\lambda_{0C} = \frac{4\pi}{3}\sqrt{\frac{2}{3}(1+z_E)^{-1}}H_0^{-1}\,, \tag{4.36}$$

i.e. about $67.3\,h^{-1}Mpc$ in a flat universe with $1 + z_E \approx 4.310^{-5}h^{-2}$ (see e.g. [158]). Even if this value may be an artifact of our simplifying assumptions, e.g. the fact that we have neglected isocurvature modes at all times and we have treated radiation as a perfect fluid even at small scales, we believe that the discrepancy between our critical scale and the Jeans scale in the same universe model is a general feature that deserves further investigation in order to consider possible effects for models of structure formation in the universe.

Finally, it is interesting to consider the limit $w \to 0$ of k_J. For flat models, one gets from (4.32) that in this limit $k_J = \frac{3\sqrt{3}}{4}a_E H_E$, i.e. we recover k_{EC} (4.31) in this limit. However the same limit for open models gives a value for k_J that: *a)* is real only for $\Omega_E > 16/25$, a result that, although satisfied in the real universe, appears spurious for the theory; *b)* gives the false impression that there could be growing and oscillating modes also for open models, contrary to what we have shown in the previous section.

4.1.4.4 Metric and curvature perturbations

In the previous sections we have given the results of the analysis of the dynamical system $\{\mathcal{U}, w\}$ (4.12), (4.18) for flat and open models, and inferred conclusions on the evolution of density perturbations, represented by the harmonic component X (see section 4.1.3.2). As we have already pointed out, there is a particular relation between the location and character of fixed points in the phase space $\{\mathcal{U}, w\}$, and the character of the corresponding point in the original phase space $\{X, X', w\}$[10]. Indeed, it is immediate that the original system ($Y \equiv X'$)

$$\left\{ \begin{array}{lcl} X' & = & Y \\ Y' & = & -\psi Y - \xi X \\ w' & = & w\left(w - \frac{1}{3}\right) \end{array} \right. \tag{4.37}$$

admits only two fixed points for $\xi \neq 0$: Point A $\equiv \{X = 0, Y = 0, w = 1/3\}$ and Point B $\equiv \{X = 0, Y = 0, w = 0\}$. The first represents an unperturbed pure radiation model, and the second a pure dust model, while the line connecting them is the mixed background model we are more interested in here. Then it is easy to see that the eigenvalues at these points along the principal directions in the $\{X, Y\}$ planes equate the roots $\mathcal{U}_\pm$ of (4.12), i.e. $\lambda_\pm = \mathcal{U}_\pm$. Point A is the same for flat and open models ($\lambda_- = -1/3$, $\lambda_+ = 2/3$, $\lambda_w = 1/3$), while (4.27), (4.28) and the analysis in section 4.1.4.2 shows that for open models $\xi = 0$ for $w = 0$, and thus Point B degenerates into a line (the $Y = 0$ axis): $\lambda_+ = \mathcal{U}_+ = 0$ in this case, so that open models are marginally stable. For flat models, the same happens for $k_E = k_{EC}$, as already pointed out.

Having clarified the relation between the roots $\mathcal{U}_\pm$ of the system (4.12), (4.18) and the eigenvalues at the fixed points of (4.37), we can now turn to the asymptotic evolution of X. It is clear from the definition of $\mathcal{U} = Y/X$ that around the roots $\mathcal{U}_\pm$ the evolution of X is given by

$$X' = \mathcal{U}_\pm X \ , \quad \Rightarrow \quad X_\pm \sim S^{3\mathcal{U}_\pm} \ . \tag{4.38}$$

When $\mathcal{U}_+ = 0$ we have $X_+ = const.$ (cf. Table 4.3); for $w \to 0$ we recover the well known constant mode for matter dominated open models: here the same mode is found for the critical scale $k_E = k_{EC}$ in flat models.

In Tables 4.4 and 4.5 we give the asymptotic behaviour for X, for a metric perturbation Φ_N and a dimensionless curvature perturbation scalar E/Θ^2. In particular, in Table 4.4 we consider flat models in the limit $w \to 0$, giving the asymptotic behaviour of various variables as functions of the scale factor S and in order of increasing wavelength λ_E. In Table 4.5 we give the asymptotic solutions of open models around Points I–IV: as for $w \to 1/3$ the universe is radiation dominated and also $\Omega \to 1$, then $S \sim t^{\frac{1}{2}}$ in this limit, and the asymptotic solutions around Points I and II are in common with flat models.

[10]For open models we are using the function $\Omega = \Omega(w, \Omega)$ (4.19), so that the further dimension Ω in the phase space is suppressed.

QUANTITY	$\lambda_E < \frac{5}{2\sqrt{6}}\lambda_{EC}$	$\lambda_E = \frac{5}{2\sqrt{6}}\lambda_{EC}$	$\lambda_{EC} > \lambda_E > \frac{5}{2\sqrt{6}}\lambda_{EC}$	$\lambda_E = \lambda_{EC}$	$\lambda_E \gg \lambda_{EC}$
$X_\pm$	$S^{-\frac{1}{4}\pm\imath Q}$	$S^{-\frac{1}{4}}$	$S^{-\frac{1}{4}\pm P}$	$S^{-\frac{1}{2}}$, *const.*	$S^{-\frac{3}{2}}$, S
$\Phi_{N\pm}$	$S^{-\frac{5}{4}\pm\imath Q}$	$S^{-\frac{5}{4}}$	$S^{-\frac{5}{4}\pm P}$	$S^{-\frac{3}{2}}$, S^{-1}	$S^{-\frac{5}{2}}$, *const.*
$E_\pm/\Theta^2$	$S^{-\frac{1}{4}\pm\imath Q}$	$S^{-\frac{1}{4}}$	$S^{-\frac{1}{4}\pm P}$	$S^{-\frac{1}{2}}$, *const.*	$S^{-\frac{3}{2}}$, S

Table 4.4: Flat models: asymptotic behaviour for $w \to 0$ (i.e. around Points III–IV) of X, Φ_N and E/Θ^2, in order of increasing wavelengths, as function of the scale factor S ($S \sim t^{\frac{2}{3}}$ for $w \to 0$). The decaying ($-$, Point III) and growing ($+$, Point IV) modes are distinguished either by the $\pm$ or presented in order: they coincide for the critical damping scale $\lambda_E = \frac{5}{2\sqrt{6}}\lambda_{EC}$. $Q = \left[\frac{3}{2}\left(k_E^2/k_{EC}^2 - \frac{25}{24}\right)\right]^{\frac{1}{2}}$, and $P = \left[\frac{3}{2}\left(\frac{25}{24} - k_E^2/k_{EC}^2\right)\right]^{\frac{1}{2}}$.

QUANTITY	POINT I	POINT II	POINT III	POINT IV
X	S^{-1}	S^2	S^{-1}	*const.*
Φ_N	S^{-3}	*const.*	S^{-2}	S^{-1}
$\frac{E}{\Theta^2}$	S^{-1}	S^2	S^{-2}	S^{-1}

Table 4.5: Open models: asymptotic behaviour around Points I–IV for X, Φ_N and E/Θ^2, as function of the scale factor S. For $w \to 1/3$ $S \sim t^{\frac{1}{2}}$, and Point I and II are in common with flat models, while for $w \to 0$ $S \sim t$. Points I and III correspond to the decaying mode, and II and IV to the growing mode.

In the limit $w \to 0$ the universe models are matter dominated, and then in flat models $S \sim t^{\frac{2}{3}}$, while in open models $S \sim t$.

In the following we shall outline the relation between the gauge–invariant metric potential Φ_N, the curvature variable E, and the density perturbation Δ (see section 4.1.3.2): more details can be found in [45] and references therein. Let E_{ab} be the electric part of the Weyl tensor[11] C_{acbd}: $E_{ab} \equiv C_{acbd}u^c u^d$, where u^a is the 4-velocity of matter; then to first order

$$a^{2\,(3)}\nabla^{b\,(3)}\nabla^a E_{ab} = \frac{\kappa\mu}{3}\Delta\ , \tag{4.39}$$

where ${}^{(3)}\nabla_a$ is a covariant derivative orthogonal to u^a. The analogue of E_{ab} in Newtonian theory is the tidal field $E_{\alpha\beta} = \nabla_{\alpha\beta}\phi$, where ϕ is the Newtonian potential and $\nabla_{\alpha\beta}\phi \equiv \nabla^2\phi - \frac{1}{3}\delta_{\alpha\beta}\phi$. Then it is possible to show that to linear order a similar formula holds in relativistic perturbation theory, i.e. $E_{\alpha\beta} = \nabla_{\alpha\beta}\Phi_N$ for the scalar part of E_{ab} (all the $\{0,0\}$

[11] Latin indices are 4-dimensional $(0,1,2,3)$, and greek indices 3-dimensional $(1,2,3)$.

and $\{0, \alpha\}$ components are second order). Moreover the field Φ_N, which play here the rôle of a gauge–invariant analogue of the Newtonian potential, is just $\Phi_N = \frac{1}{2}(\Phi_A - \Phi_H)$, where Φ_A and Φ_H are the gauge invariant metric perturbations defined by Bardeen [18] ($\Phi_A = -\Phi_H$ for perfect fluids). Then using $E_{\alpha\beta} = \nabla_{\alpha\beta}\Phi_N$ in (4.39), the harmonic decomposition (4.21) and $\Delta^{(k)} = -k^2\varepsilon_m$, we get equation (4.3) of Bardeen [18]:

$$2(3K - k^2)\Phi_N = \kappa a^2 \mu \varepsilon_m \ . \tag{4.40}$$

Then this relation can be used to determine the asymptotic evolution of the gauge–invariant metric potential Φ_N, and also that of amplitude $E = \frac{1}{2}\sqrt{E_a{}^b E_b{}^a}$ of the tidal field E_{ab}: indeed from a comparison of (4.39) and (4.40) the harmonic components of E and Φ_N are related by

$$E = \frac{1}{2} a^{-2} k^2 \Phi_N \ . \tag{4.41}$$

Then it is usual to consider a dimensionless scalar to measure the relative dynamical significance of a given field using Θ to take into account the expansion (see e.g. [116]): thus in the case of E we consider E/Θ^2, for which we have

$$\frac{E}{\Theta^2} \sim \frac{a^{-2}\Phi_N}{\Theta^2} \sim \frac{\kappa\mu}{\Theta^2}\varepsilon_m \sim \Omega\varepsilon_m \ . \tag{4.42}$$

Hence, for example for a flat model $\Phi_N \sim$ *const.* for very large scales, but the relative amplitude of the tidal field $E/\Theta^2 \sim \Delta \sim a$ grows unbounded (cf. [73]).

4.1.5 Conclusions

We have considered the stability properties of FLRW models with uncoupled Cold Matter (dust) and radiation, and the perturbed models were considered as dynamical systems described by an evolution equation for a gauge–invariant density perturbation variable coupled with the equations governing the evolution of relevant background variables: the pressure–energy density ratio $w = p/\mu$ and the density parameter Ω. For the subsystem describing flat models, given by $\Omega = 1$, we deal with a planar autonomous system, and for the open models we have shown that we can also restrict the analysis to a planar system for each particular value of the density parameter at equi-density Ω_E.

The analysis of flat and open models gives different results: flat models admit unstable perturbation modes, while open models are marginally stable with respect to perturbations, irrespective of their scale, ie. the perturbations freeze–in at a constant value. Qualitatively, both results can be expected on the basis of the results of a more traditional study of the behaviour of the perturbations in simple models. The final fate of the perturbations in open models appears to be dominated by the background curvature, which governs the background expansion at late times: therefore the stability properties of the dust–radiation models are the same as those of a pure dust model, irrespective of the size of the perturbation and of the radiation content of the given model.

Instead, we find more interesting features for flat models: in this case each model has a characteristic critical invariant wavenumber k_{EC} which depends on the proportion of matter and radiation (e.g. at present), and the corresponding scale λ_{EC} determines the transition from stable to unstable modes. The scale λ_{EC} is of the order of the Hubble radius H_E^{-1} at the equi-density of matter and radiation, H_E^{-1} being the only scale entering the background model. The present value of this transition scale is $\lambda_{0C} = (1+z_E)\lambda_{EC}$ and depends on the actual present proportion of matter and radiation, for which it is of the order of $67Mpc$. Therefore the stability properties of dust–radiation flat models are a mixture of the properties of pure dust and pure radiation models, being well known that perturbations in a flat dust models grow unbounded irrespective of their size, while all perturbation scales in a radiation model always enter an oscillatory (sound waves) regime after they enter the sound horizon (e.g. see [18]).

However, we actually find a structure in the stability properties of flat dust–radiation models which is more interesting than expected, because we find three different regimes for the evolution of the perturbations. Perturbations on scales $\lambda_E > \lambda_{EC}$ grow unbounded (unstable modes), while perturbations in the range $\frac{2\sqrt{6}}{5}\lambda_{EC} < \lambda_E < \lambda_{EC}$ are overdamped; finally, perturbations on scales $\lambda_E < \frac{2\sqrt{6}}{5}\lambda_{EC}$ are damped, i.e. they oscillate as "sound waves" while decaying.

The stability properties of perturbed FLRW models that we have found, should be taken *cum grano salis* with respect to the problem of structure formation in the universe. For example, it is clear that the marginal stability we have found for open models should not be taken as implying that structures cannot form in an open universe: in fact it can be expected that this stability could be broken by non–linearity. In other words, here we have found that open models (see also [43]) are marginally stable against *linear* perturbations;

in going beyond this level of approximation we can expect that instability will be switched on by non–linearity.

Also, we have assumed adiabatic perturbations at all times, neglecting isocurvature modes that are in principle important at small scales and late times, and we have neglected photon diffusion, treating radiation as a perfect fluid irrespective of the scale of the perturbation (in particular, the oscillatory behaviour of small scales in flat models is probably an artifact of this assumption). However, we believe that these assumptions should not question the validity of our main results for flat models: *a)* in a given flat universe model (including the assumptions about the matter content and how to treat it) there are perturbations that decay despite the fact that their wavelength at equi-density λ_E is larger than the Jeans scale λ_{JE}; *b)* the today size of the critical scale we have found is of the order of $67\,Mpc$, a fact that perhaps deserves further investigation regarding its implication for structure formation in the universe. Also, while in standard flat CDM models the fluctuations stop to oscillate when the universe enters the matter dominated era, we have found that small scale fluctuations continue to be damped even at late times. It will be therefore interesting to make a proper comparison between the power spectrum of standard CDM and the one that could be derived from the analysis given here (normalizing the spectrum at large scales), eventually considering also the isocurvature modes that, as said above, are expected to modify the small scale fluctuations behaviour. This sort of analysis could lead to interesting results, giving less power at small scales in comparison with the standard CDM results. However we cannot go beyond this speculative level at this stage, and we leave this for a future work.

Finally, we remark that a comparison of the stability properties of dust–radiation models with the observed small amplitude of the large scale density fluctuations seems to suggest that if the spatial curvature of the universe vanishes, then we live in a special epoch in which these perturbations are still in a linear regime of growth, while if we live in a universe with negative spatial curvature the smallness of the large scale perturbations is a characteristic of the model at all times and the description of the universe as an open FLRW model is appropriate at any epoch.

Thus open models are special (at a given time) on average, because the density parameter Ω depends on time (this is the flatness problem), but not from the local point of view, because large scale structures are frozen–in, while the reverse is true for flat models, because $\Omega = 1$ at all times, but large scale perturbations grow unbounded in these models.

5 Outlook and directions for future research

Following the conclusions of chapter three we would like to advocate that the application of RG in cosmology is a line of research worth pursuing. It seems at the end to be connected with critical phenomena, chaos, self-organized criticality (SOC), fractals, flicker noise, and in particular, the notion of entropy for gravitational systems and complexity. The interconnections are not yet understood nor appreciated, and many issues are not understood, e.g. the flicker noise ("$1/f$") (e.g. [195]) is still one of the great mysteries of physics and only now have some models displaying SOC been shown to exhibit this kind of noise.

Particularly interesting are, in our opinion, critical phenomena. Critical phenomena are after all always manifested macroscopically since phase transitions are collective phenomena (arising from interactions between quasi-particles of the system), with the presence of non-linearities of great importance. It would therefore be very interesting and valuable to study further critical phenomena and the Renormalization Group in a cosmological setting, and in particular, the relevance of these to the structure formation and clustering in the universe.

This is an important point to understand how things become organized into complex structures in the universe. We can conjecture that some "laws" of organization (self-organization)[1] or information-processing are necessary to describe not only the quantity, but also the organization of the information within a system. As proposed in [58], one could describe complexity of a state by the running-time or entropy production of the

[1] L. Smolin, private communication.

shortest possible program needed to compute the state. These ideas, relevant to a definition of randomness and chaos, in general, have not yet played any important rôle in the understanding of cosmology.

Expanding further these considerations, and having in mind RG, let us notice a useful connection between gravitation and turbulence at high Reynolds number, statistical field theory and critical phenomena. Firstly, there is an analogue between classical gravitation, turbulence and field theories with an infra-red attractive RG fixed point. This is the situation for field theories corresponding to critical lattice spin systems, and in particular the Ising model. The following list can be drawn up:

Turbulence	**Critical Phenomena**	**Gravitation**
space separation	wavenumber	separation (geodesic)
viscosity	temperature	3−metric
energetic length-scale	UV cut-off (or, inverse lattice-spacing)	coarse-graining scale
dissipation wavenumber	correlation length	correlation length
velocity correlation function	spin correlation function	two-point correlation fn, etc.
intermittency exponent	correlation exponent	correlation exponent

The analogies between turbulence and critical phenomena have been pointed out e.g. in [67] (see also [98] recently) and we will not discuss them here. The analogy with gravitation has not been, to our knowledge, spelled out. Above, the rôles of space and wavenumber for turbulence and critical phenomena are interchanged, because turbulence as opposed to critical phenomena, exhibits short distance scaling believed to be generic and essentially independent of the large scale statistics or driving mechanisms.

Secondly, let us point out for completeness, an analogy that can be made between turbulence and field theories with an ultra-violet attractive fixed point [98], and Quantum Gravity:

Turbulence	**Field Theory**	**Quantum Gravity**
space separation	space separation	separation
viscosity (or Kolmogorov scale)	lattice-spacing	"discretization"-spacing[2]
energetic length-scale	correlation length	correlation length $\sim \frac{1}{G}$
Kolmogorov wavenumber	UV cut-off = inverse lattice spacing	UV cut-off
velocity correlation function	field-theoretic Green function	correlation function on the lattice

Let us remark that field theories such as UV asymptotically free QCD exhibits scaling at short distances, just as turbulence does. Likewise, QG models studied so far (mainly numerically studied dynamically triangulated QG models), exhibit similar finite-size scaling, and in dim= 4 a second order phase transition, most probably, takes place from the "smooth" to the "crumpled" phase [126, 4, 5, 2, 3, 54] (the last one is an analytical approach)[3].

Let us also remark that in the first analogy the inverse rôle of large and small scales arises from the different character of "cascade" in the two cases. In the cascade picture, there is a transfer of excitation on the average from the large turbulent eddies to the small ones by a stepwise process, which is chaotic in nature and entails a loss of memory of the large-scale statistics. Wilson has emphasized [255] that there is also a "cascade of fluctuations" in critical phenomena. Droplet fluctuations nucleated at the lattice-scale in the critical state can grow to the size of the correlation length. But now the details of the lattice structure are lost instead and the scale-invariant distributions of the large "droplets" are universal. In this connection, let us notice that interestingly enough there is a "hierarchy of structures" in the universe that is governed by relativistic gravitation field theory, which in certain sense has in it a cascade picture too.

Now the challenge would be to really understand these analogies (and differences). While concrete problems may be highly non-trivial, e.g. how to define a suitable analog of the Migdal decimation or Kadanoff block-spin transformation for curved dynamical lattices (see [145] for an attempt in 2 dimensions). In any case, the overall conceptual picture sketched here could be, we hope, of help for further research. It is after all one

[2] E.g. size of simplicial complexes in triangulations or edge links as in Regge calculus [209].

[3] For a good introduction to this subject see [66] (cf. also [39]).

more example of the great unity and pluralism in theoretical physics [93], and this is what makes it even more interesting.

An alternative way to perform the averaging was recently put forward by Zeeman [265], who adopted the $\epsilon-$smoothing in applying the related Fokker-Planck equations as a way to "stabilize" the resultant dynamical system. It would therefore be interesting to compare the outcomes of this approach with the others.

Appendix A
Some useful notions from Riemannian Geometry

We recall here some basic concepts in Riemannian Geometry [27, 109] (see also [20] for a concise overview).

Let $\mathcal{M}$ be a smooth C^{∞}, Hausdorff, connected, oriented, compact n-dimensional manifold without boundary and let g be a riemannian metric on $\mathcal{M}$, i.e., a smoothly varying family of inner products G_x on the tangent spaces $T_x\mathcal{M}$, $x \in \mathcal{M}$.

A metric g on $\mathcal{M}$ is called an *Einstein* metric if the Ricci curvature $Ric(g) = \lambda g$ for some constant λ. By normalization, one can always assume to be in one of the three cases: $Ric(g) = g$ (when $\lambda > 0$), $Ric(g) = 0$ ($\lambda = 0$) or $Ric(g) = -g$ (when $\lambda < 0$). We use the term "Einstein manifolds" for riemannian manifolds of constant Ricci curvature.

Let $c : [0, a] \to \mathcal{M}$ be a curve, and $0 = a_o < a_1 < \ldots < a_n = a$ be a partition of $[0, a]$ such that $c_{|[a_i, a_{i+1}]}$ is of class C^1. The *length of* c is defined by

$$L(c) = \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} |c'(t)| dt, \qquad (A.1)$$

where, $|c'(t)| = \sqrt{g(c'(t), c'(t))}$.

The length of a curve does not depend on the choice of a regular parameterization.

The *riemannian distance* between two points x and y in $\mathcal{M}$ is defined to be the infimum of the length (with respect to g) of the curves from x to y. The *diameter* D of $(\mathcal{M}, g)$ is the diameter of $\mathcal{M}$ for the riemannian distance.

The *geodesics* are the curves which satisfy the Euler-Lagrange equation of the problem of minimization of the energy of a curve. In particular, given any point x in $\mathcal{M}$ and any unit vector $u \in T_x\mathcal{M}$, there is (locally) one and only one geodesic $c_{x,u}$ parameterized by arc length t, such that $c_{x,u}(0) = x$ and $\dot{c}_{x,u}(0) = u$ (such a geodesic is defined for all values of t when $\mathcal{M}$ is closed).

We define the *exponential map* $\exp_x : T_x\mathcal{M} \to \mathcal{M}$, by $\exp_x(tu) = c_{x,u}(t)$, for any $t \geq 0$ and any unit tangent vector u. The exponential map is a local diffeomorphism from a neighbourhood of 0 in $T_x\mathcal{M}$ to a neighbourhood of x in M, its derivative at 0 is the identity map.

An *isometry* f between two riemannian manifolds $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ is a smooth map $f : \mathcal{M} \to \mathcal{N}$ whose derivative induces isometries between the tangent spaces, with respect to the inner products g and h, respectively. In particular, the two riemannian manifolds are isometric if there exists some diffeomorphism $f : \mathcal{M} \to \mathcal{N}$, which transfers h into g, i.e., $f^*h = g$.

A riemannian *structure* is a class of isometric riemannian manifolds. In other words, if $Riem(\mathcal{M})$ denotes the set of riemannian metrics on $\mathcal{M}$, the set of riemannian structures on $\mathcal{M}$ is the quotient $Riem(\mathcal{M})/Diff(\mathcal{M})$ of $\mathcal{M}$ by the group of diffeomorphisms $Diff(\mathcal{M})$ of $\mathcal{M}$.

The various notions of *curvature* measure how the exponential maps differ from being isometries (at least locally). Let P be a 2-plane in $T_x\mathcal{M}$. Given a small enough r, consider the image under the exponential map $\exp_x$ of a circle of radius r and centre 0 in the plane P. This is a closed curve in $\mathcal{M}$ with length $L(r)$. When $r \to 0$ we have Puiseux' formula:

$$L(r) = 2\pi r(1 - \frac{1}{6}\sigma(x, P)r^2 + \mathcal{O}(r^3)). \tag{A.2}$$

The number $\sigma(x, P)$ is called the *sectional curvature* of the 2-plane P at x (see [102] for a lucid exposition).

An oriented riemannian manifold is also equipped with a natural *riemannian measure* v_g, whose expression in a local coordinate system $\{x_i\}$ is $\det(g_{ij})^{\frac{1}{2}}dx$, where dx is the

Lebesgue measure and where $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$. The volume of $(\mathcal{M}, g)$ is denoted by $V(g) = \int_{\mathcal{M}} dv_g$.

We can write the pull-back $\exp_x^* v_g$ of the riemannian measure v_g by the exponential map in polar coordinates in $T_x\mathcal{M}$ by $\exp_x^* v_g = \Theta_x(t,u)dtdu$, where $t \geq 0$, dt is the Lebesgue measure on $\mathbb{R}_+$, u is a unit vector and du is the canonical measure on the unit sphere.

When $t \to 0$, we have

$$\Theta_x(t,u) = t^{n-1}(1 - \frac{1}{6}\rho_x(u)t^2 + \mathcal{O}(t^3)). \tag{A.3}$$

The number $\rho_x(u)$ is a quadratic form on $T_x\mathcal{M}$ which defines a symmetric bilinear form called the *Ricci curvature*, $Ric(g)$ of $\mathcal{M}$ at the point x.

If $\{u, e_2, \ldots, e_n\}$ is an orthonormal basis in $T_x\mathcal{M}$ and if P_i is the 2-plane spanned by u and e_i, we have the formula

$$Ric(g)(u,u) = \sum_{i=2}^{n} \sigma(x, P_i), \tag{A.4}$$

so the Ricci quadratic form is essentially a sum of sectional curvatures.

A riemannian metric g on a compact 3-manifold $\mathcal{M}$ is defined to be *locally homogeneous*, if and only if for every pair (x,y) of points of $\mathcal{M}$, there exist neighbourhoods U_x of x and V_y of y, such that there is an isometry $\psi : U_x \to V_y$ with $\psi(x) = y$.

Generally, these local isometries do not extend to isometries of the whole space $(\mathcal{M}, g)$. If the local isometries do extend, then the geometry is defined to be homogeneous, i.e., $(\mathcal{M}, g)$ is *homogeneous* if for every pair of points x, y in $\mathcal{M}$, there exists an isometry $\phi : \mathcal{M} \to \mathcal{M}$ with $\phi(x) = y$. In this case the group of isometries of $\mathcal{M}$ acts transitively. For every locally homogeneous geometry the universal cover is homogeneous. We say then that the locally homogeneous geometry is modeled by the homogeneous geometry.

Appendix B
The Ricci-Hamilton Flow

The dimensionality $n = 3$, unless explicitly stated otherwise.

Let $Riem(\mathcal{M})$ denote the space (infinite dimensional) of smooth riemannian metrics on $\mathcal{M}$ (this set has a natural structure of Fréchet manifold); and $S^2\mathcal{M}$ the set of smooth bilinear forms on $\mathcal{M}$.

The diffeomorphisms act by pull-back, i.e., $Diff(\mathcal{M}) : S^2\mathcal{M} \to S^2\mathcal{M}$.

The riemannian structure underlying $(\mathcal{M}, g)$ is described by the orbit O_g of metric g, in $Riem(\mathcal{M})$ under the action of $Diff(\mathcal{M})$, it is defined as

$$O_g \equiv \{g' \in Riem(\mathcal{M}) \,|\, g' = \varphi^* g \text{ for some } \phi \in Diff(\mathcal{M})\}. \tag{B.5}$$

The tangent space to $Riem(\mathcal{M})$ at a given g, i.e., $T_g Riem(\mathcal{M})$ is interpreted as the set of infinitesimal deformations of the given g, and is isomorphic to $S^2\mathcal{M}$. In particular, it contains as the subspace a tangent space to the orbit O_g, which is an image in $S^2\mathcal{M}$ of the linear differential operator $L : \Xi \to S^2\mathcal{M}$, $X \to L_X g$, where L_X is Lie differentiation along the vector field X, and Ξ the space of vector fields on $\mathcal{M}$ (Ξ has an interpretation of the Lie algebra of $Diff(\mathcal{M})$).

One can show (using the decomposition theorems [22], see also [49] for a clear account), that we have the L^2-orthogonal splitting

$$T_g Riem(\mathcal{M}) \simeq Im\, L \oplus Ker\, L^*, \tag{B.6}$$

where $\oplus$ stands for the orthogonal sum with respect to the global scalar product on $\mathcal{M}$, defined by

$$(h, h')_g \equiv \int_{\mathcal{M}} h_{ik} h'_{lm} g^{il} g^{km} dv_g, \tag{B.7}$$

for each $h,\ h' \in S^2\mathcal{M}$.

Formula (B.6) can be rewritten as

$$T_g Riem(\mathcal{M}) \simeq Im\, div^* \oplus Ker\, div, \tag{B.8}$$

where, $(div\, S)_i \equiv -g^{jk}\nabla_k S_{ij}$, $(div^* x)_{ij} \equiv \frac{1}{2}(\nabla_i x_j + \nabla_j x_i)$, because the formal L^2-adjoint of L with respect to $(\ ,\ )_g$, $L^* : S^2\mathcal{M} \to \Xi$ is (minus) twice the divergence operator on $S^2\mathcal{M}$.

The geometrical interpretation of (B.8) tells us that any infinitesimal deformation $h \in S^2\mathcal{M}$, can be decomposed into a longitudinal deformation $h_{lon} \in Im\, div^*$, mapping g into, say, g' within the same orbit, i.e., $g \to g' \in O_g$, and a transversal one $h_{tran} \in Ker\, div$, $g \to g'' \notin O_g$, which takes g to the other orbit and provides thus an infinitesimally deformed new riemannian structure on $\mathcal{M}$.

On $Riem(\mathcal{M})$ there is a naturally defined $Ric(g)-$generated field of non-trivial infinitesimal deformations provided by associating with the given metric g, the tensor field $[Ric(g) - kgR(g)]$, where k is any real number. This follows by noticing that for any k, $[Ric(g) - kgR(g)]$ is never tangent, at g, to the orbit O_g, unless it vanishes[4]. In other words, the deformation $Ric(g) - kgR(g)$ mapping $g \to g'$, such that infinitesimaly close riemannian metric $g' = g - \xi[Ric(g) - kgR(g)] + \mathcal{O}(\xi^2)$, defines a new riemannian structure on $\mathcal{M}$, since such $g' \notin O_g$ (for a proof of this fact see [38]).

Now one can investigate the question of existence and behaviour of the integral curves (if any) of this vector field[5].

The answer was given by R. Hamilton [124], who showed that the flow associated with the $Riem(\mathcal{M})$ vector field $g \to -2Ric(g)$ is the local flow of metrics in $Riem(\mathcal{M})$ and

[4] Ricci tensor $Ric(g) : Riem(\mathcal{M}) \to S^2\mathcal{M}$; $R(g)$ stands for the scalar curvature.

[5] It is not evident *a priori* that this $Ric(g)-$generated deformations patch together to define a local (or possibly global) flow of metrics in $Riem(\mathcal{M})$ due to the Fréchet structure of $Riem(\mathcal{M})$.

moreover it is global, on condition the Ricci tensor associated with the metric is a positive bilinear form.

By deforming or smoothing flow of metrics, we mean a curve $g_{ab}(\beta)$, such that $g_{ab}(0)$ is the original given metric, and $g_{ab}(\beta)$ becomes smooth for $\beta \to \infty$. To see how it comes about, consider the general infinitesimal deformation of a metric

$$g_{ab} \to g_{ab} + \Delta\beta h_{ab}, \tag{B.9}$$

where, h_{ab} is any symmetric rank two tensor. Since, in appropriate coordinates, the leading term in the Ricci curvature $Ric(g)$ is $\nabla^c \nabla_c g_{ab}$, a natural choice for h_{ab} is

$$h_{ab} = -2Ric(g). \tag{B.10}$$

Writing this in the form of a differential equation, and adding a term responsible for preserving the volume of $(\mathcal{M}, g_{ab})$ along the flow, results in the Ricci-Hamilton flow equation (B.11).

Theorem 1 *Let $(\mathcal{M}, g)$ be a closed (compact and without boundary) riemannian 3-manifold, such that its Ricci tensor, $Ric(g)$, is a positive definite bilinear form (i.e., $[Ric(g)]_{ab}v^a v^b > 0 \ \forall v \neq 0$ vector field), then the given metric g can be uniformly deformed into a constant curvature metric $\overline{g}$.*

In this case, the universal simply connected cover of $\mathcal{M}$ is the 3-sphere S^3 and the pull back of $\overline{g}$ to S^3 via the covering map $S^3 \to \mathcal{M}$ is the standard metric[6].

The one-parameter flow of metrics on $\mathcal{M}$ $(g, \beta) \to g(\beta)$ (with $\beta \geq 0$ the deformation parameter) realizing the above deformation is the unique solution to the weakly parabolic initial value problem

$$\frac{\partial}{\partial \beta} g_{ab}(\beta) = \frac{2}{3} < R(\beta) >_\beta g_{ab}(\beta) - 2R_{ab}(\beta), \tag{B.11}$$

with the initial data $g_{ab}(\beta = 0) = g_{ab}$ $(a, b = 1, 2, 3)$, where $R_{ab}(\beta)$ are the components of the Ricci tensor $Ric(g(\beta))$, and $< R(\beta) >_\beta$ denotes the average scalar curvature

$$< R(\beta) >_\beta = \frac{1}{Vol(\mathcal{M}, g(\beta))} \int_{\mathcal{M}} R(\beta) dv_\beta \tag{B.12}$$

[6]The theorem in fact, forces $\mathcal{M}$ to be topologically S^3/Γ, i.e., S^3 possibly quotiented by a discrete group.

The Ricci-Hamilton flow equation is a heat-like equation (weakly parabolic) and results in a smoothing deformation of the initial data $g(x, \beta = 0) = g(x)$.

Before examining some of the properties of the Ricci-Hamilton flow we would like to recall the strategy underlying the proof of Hamilton's theorem.

In fact, one can equivalently deal with a simpler initial value problem than (B.11) for what concerns most of the analysis involved in proving theorem 1. Equation (B.11) is not strictly parabolic since the Ricci tensor (thought of as a second order differential operator) is not elliptic. This is a consequence of its $Diff(\mathcal{M})$ equivariance, i.e. $Ric(\varphi^* g) = \varphi^* Ric(g)$, for any smooth diffeomorphism $\varphi : \mathcal{M} \to \mathcal{M}$. There is also associated with (B.11) a natural integrability condition in the form of contracted Bianchi identities, $div[Ric(g(\beta)) - \frac{1}{2}g(\beta)R(g(\beta))] = 0$, which has to hold true for any β for which the flow $g(\beta)$ exists. Up to this integrability condition (B.11) is parabolic and its local solvability can be handled by means of Nash-Moser implicit function theorem. According to the results of De Turck [68] we can associate with (B.11) a manifestly parabolic initial value problem at the expense of a clever use of the mentioned above $Diff(\mathcal{M})$ equivariance, namely

$$\frac{\partial}{\partial \beta} g_{ab}(\beta) = -2R_{ab}(\beta),$$

with the initial condition $g_{ab}(\beta = 0) = g_{ab}$. This simplifies the proof of local solvability for (B.11). Let us remark that the local part of Hamilton's theorem does not rely either on the positivity of the Ricci tensor or the three-dimensionality of $\mathcal{M}$. But apparently both these requirements are needed in order to globalize the local flow $(g, \beta) \to g(\beta)$. To see this, notice that the positivity of the Ricci tensor if assumed initially is preserved along the flow. This fact allows to get *a priori* estimates showing that the solution of (B.11) exists for all $0 \le \beta < +\infty$. Moreover, the sign restriction on $Ric(g)$ yields a proof that as $\beta \to \infty$, the three eigenvalues of the Ricci tensor, at the generic point $x \in \mathcal{M}$, converge to a common value. It is then proved that this common value is a constant (positive).

As $\beta \to \infty$, $g(\beta)$ approaches the constant curvature metric $\bar{g}$ on S^3 uniformly, in fact the convergence is even exponential [124].

Now we will discuss some of the properties of the Ricci-Hamilton flow.

Firstly, the flow $(g, \beta) \to g(\beta)$ preserves the total volume of $(\mathcal{M}, g)$: $Vol(\mathcal{M}, g(\beta)) =$

$Vol(\mathcal{M}, g)$ for $0 \leq \beta < \infty$.

To prove this, we use the fact that along the trajectories of the flow $g(\beta)$ solution to (B.11) we have

$$\frac{\partial}{\partial\beta} dv_\beta = \frac{1}{2}\left(g^{ab}\frac{\partial}{\partial\beta}g_{ab}\right) dv_\beta = [< R(\beta) >_\beta - R(\beta)] dv_\beta,$$

so that

$$\frac{\partial}{\partial\beta} Vol(\mathcal{M}, g(\beta)) \equiv \frac{\partial}{\partial\beta}\int_{\mathcal{M}} dv_\beta = \int_{\mathcal{M}} [< R(\beta) >_\beta - R(\beta)] dv_\beta = 0.$$

The second group of properties follow upon examining the formal linearization of (B.11) around a given solution $g(\beta)$.

The linearized Ricci-Hamilton flow evolves a given infinitesimal deformation yielding a β-parameterised family of vectors $h(\beta) \in \mathrm{T\,Riem}(\mathcal{M})$, connecting two neighbouring flows of metrics.

Upon the formal linearization of the initial value problem (B.11) around a given solution $g(\beta)$, we obtain (the β's in the brackets suppressed)

$$\begin{aligned} \frac{\partial}{\partial\beta} h_{ab} &= \frac{2}{3} < R > h_{ab} + \frac{2}{3} g_{ab}[\frac{1}{2} < R g^{ab} h_{ab} > - \frac{1}{2} < R >< g^{ab} h_{ab} > - \\ & < R^{ab} h_{ab} >] - \Delta_L h_{ab} + 2[div^*(div(h - \frac{1}{2}(Trh)\, g))]_{ab}, \end{aligned} \tag{B.13}$$

with the initial data $h_{ab}(\beta = 0) = h_{ab}$, where, $h \in S^2\mathcal{M}$ is a given symmetric bilinear form, Δ_L is the Lichnerowicz-DeRham Laplacian on bilinear forms, and the operators Δ_L, div^*, div and Tr are considered with respect to the flow of metric $(g, \beta) \to g(\beta)$, solution of (B.11). The div (here, minus the usual divergence) is the divergence operator on $S^2\mathcal{M}$, div^* is the L^2 adjoint of div, acting from the space of vector fields on $\mathcal{M}$ to $S^2\mathcal{M}$ (it can be identified with $\frac{1}{2}$[Lie derivative] of the metric tensor along a vector field).

Note that a $h(\beta)$ solution of the initial value problem (B.13) always exists and is unique, and evolves a given infinitesimal deformation yielding a β-parameterised family of vectors $h(\beta)$ in $\mathrm{T\,Riem}(\mathcal{M})$ connecting two neighbouring flows of metrics $g(\beta)$ and $g'(\beta)$ (obtained as solutions of problem (B.11) with initial data $g(\beta = 0) = g$ and $g'(\beta = 0) = g(\beta = 0) + \epsilon h(\beta = 0) + \mathcal{O}(\epsilon^2)$, respectively).

The solution to (B.13) has a basic property expressing the $Diff(\mathcal{M})$ equivariance of the Ricci-Hamilton flow, namely, a trivial deformation $h_{ab} = L_X g_{ab}$ (where $X : \mathcal{M} \to T\mathcal{M}$ is a smooth vector field on $\mathcal{M}$) is always mapped by (B.13) into a trivial deformation, in other words, the solution to the linearized Ricci-Hamilton initial value problem is determined up to the infinitesimal diffeomorphism.

This result implies that if X is a Killing vector field for the given $(\mathcal{M}, g)$, then it remains so along the trajectories of the flow $(\beta, g) \to g(\beta)$. In other words, all the symmetries which the original metric g may be endowed with are preserved by the Ricci-Hamilton flow.

The natural problem to address, would be next that of generalizing Hamilton's theorem as much as possible.

At this point we refer the reader to the original literature [124] (see also [52] for a comprehensive review); for the understanding, why there is a positivity requirement on the Ricci tensor, and to what extend it can be weakened - from the point of view of solvability the initial-value problem (B.11), as well as of the topological obstructions to positive Ricci curvature - see [52] (see also [53]).

Let us stress that the positivity condition on the Ricci tensor apparently is not a necessary one for the Ricci-Hamilton flow to be global [141, 56]. As explicitly proved in [56], the 3-torus T^3 provides a non-trivial example of Ricci-Hamilton flow (with the Ricci tensor being non-positive), such that Hamilton's initial-value problem admits a global solution[7].

[7] The T^3 cosmology, with a 3-space in form of a 3-torus is the simplest inhomogeneous empty universe [191].

Appendix C Gromov space of bounded geometries

Consider two riemannian manifolds, and let $i_1(\mathcal{M}_1), i_2(\mathcal{M}_2)$ stand for two isometric embeddings of $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, in some metric space (A, d).

A *Hausdorff distance* in (A, d) between $i_1(\mathcal{M}_1)$, $i_2(\mathcal{M}_2)$ can be introduced as follows:

$$\begin{aligned} d_H^A[i_1(\mathcal{M}_1), i_2(\mathcal{M}_2)] &= \inf\{\epsilon > 0 | \ U_\epsilon(i_1(\mathcal{M}_1)) \supset i_2(\mathcal{M}_2), \\ &\qquad U_\epsilon(i_2(\mathcal{M}_2)) \supset i_1(\mathcal{M}_1)\}, \end{aligned} \tag{C.14}$$

where, the ϵ-neighbourhood $U_\epsilon(i_i(\mathcal{M}_i))$ of $i_i(\mathcal{M}_i)$, $i = 1, 2$ is defined as

$$U_\epsilon(i_i(\mathcal{M}_i)) = \{z \in A | \ d(z, i_i(\mathcal{M}_i)) \leq \epsilon\}. \tag{C.15}$$

The Hausdorff distance thus defined is the lower bound of the ϵ, such that $i_1(\mathcal{M}_1)$ is contained in the ϵ-neighbourhood of $i_2(\mathcal{M}_2)$, and vice versa.

The *Gromov distance* $d_G(\mathcal{M}_1, \mathcal{M}_2)$ provides a natural generalization of the Hausdorff distance, and it is defined as the lower bound of the Hausdorff distances, as A varies in the set of metric spaces, and i_1, i_2 vary in the set of all isometric embeddings of $\mathcal{M}_1$ and $\mathcal{M}_2$ in (A, d).

The Gromov distance provides us with a sense of geometric nearness among riemannian structures, which is related to a classification of riemannian manifolds, according to how they can be covered by small geodesic balls. Coverings with the balls packed in similar configurations are possible for riemannian manifolds that can be considered close to each other in the sense of Gromov distance.

In particular, the Gromov distance between two compact manifolds is always finite, and $d_G(\mathcal{M}_1, \mathcal{M}_2) = 0$, when the two manifolds are isometric.

Let us introduce the following class of riemannian structures:

for $k \in \mathbb{R}$ and $D \in \mathbb{R}_+$, let $Ric[n, k, D]$ denote the space of isometry classes of closed, connected n-dimensional riemannian manifolds $(\mathcal{M}, g)$ (without any pre-assumption on their topology) with Ricci curvature $Ric(g) \geq (n-1)kg$ and diameter $\leq D$.

Recall that if we define

$$k(x) \equiv \inf\{\inf Ric(u, u)|\, u \in T_x\mathcal{M},\, |u_x| = 1\}, \tag{C.16}$$

the lower bound of the Ricci tensor of $\mathcal{M}$ is defined as the lower bound of $k(x)$ as x varies in $\mathcal{M}$.

The best such $k = k(g)$ is just the lowest eigenvalue of the Ricci curvature $Ric(g)$. It is a fundamental numerical invariant of a compact riemannian manifold.

For any manifold $\mathcal{M} \in Ric[n, k, D]$, it is possible to introduce the covering by geodesic balls, providing a coarse classification of riemannian structures occurring in $Ric[n, k, D]$.

For any given $\epsilon > 0$, it is always possible to find an ordered set of points $\{p_1, \ldots, p_N\}$ in $\mathcal{M}$ from the above class, so that:

i) the balls $B_{\mathcal{M}}(p_i, \epsilon) = \{x \in \mathcal{M}|\ d(x, p_i) \leq \epsilon\}$, $i = 1, \ldots, N$ (where, $d(\ ,\ )$ denotes the distance function of $\mathcal{M}$) cover $\mathcal{M}$, i.e., the collection $\{p_1, \ldots, p_N\}$ is an *ϵ-net* in $\mathcal{M}$.

ii) the open balls $B_{\mathcal{M}}(p_i, \epsilon/2)$, $i = 1, \ldots, N$ are disjoint, i.e., $\{p_1, \ldots, p_N\}$ is a *minimal* ϵ-net in $\mathcal{M}$.

A *filling function* $N_\epsilon^{(o)}(\mathcal{M})$ of the covering is defined as the function, which associates with $\mathcal{M}$ the maximum number of geodesic balls realizing a minimal ϵ-net in $\mathcal{M}$.

Any minimal net is characterized by its *intersection pattern*, defined as the set of indices pairs

$$I_\epsilon(\mathcal{M}) \equiv \{(i, j)|\, i, j = 1, \ldots, N|\ B(p_i, \epsilon) \cap B(p_j, \epsilon) \neq \emptyset\} \tag{C.17}$$

Any two manifolds $\mathcal{M}_1, \mathcal{M}_2 \in Ric[n, k, D]$ with minimal ϵ-nets $\{p_1, \ldots, p_N\}$, and $\{q_1, \ldots, q_N\}$, respectively, are said to be equivalent, if and only if $N = L$ and if they have the same intersection pattern, i.e., if the equivalence relations

$$N_\epsilon^{(o)}(\mathcal{M}_1) = N_\epsilon^{(o)}(\mathcal{M}_2) \tag{C.18}$$

$$I_\epsilon(\mathcal{M}_1) = I_\epsilon(\mathcal{M}_2), \tag{C.19}$$

are true (up to combinatorial isomorphism).

In fact, the above relations partition $Ric[n, k, D]$ into disjoint equivalence classes, whose finite number can be estimated in terms of the parameters n, k, D.

Two riemannian manifolds in $Ric[n, k, D]$ get closer and closer to each other in d_G, if we can cover them with finer and finer minimal ϵ-nets of geodesic balls with the same intersection patterns.

In order to have $d_G(\mathcal{M}_1, \mathcal{M}_2) < \epsilon$, for any two compact riemannian manifolds, it is sufficient to show that there exist an $\epsilon/2$ lattice in $\mathcal{M}_1$ and an $\epsilon/2$ lattice in $\mathcal{M}_2$, and two isometric embeddings $i_j : \mathcal{M}_j \to Z$ in some metric space (Z, d), such that the distance between the corresponding points of the embedded lattices is $< \epsilon/2$.

When discussing the convergence of a sequence $\{\mathcal{M}_i\}$ of riemannian manifolds with respect to Gromov distance d_G, there is no need to refer to isometric embeddings in metric spaces. The sequence $\{\mathcal{M}_i\}$ admits a convergent subsequence, if and only if $\forall \epsilon > 0$, $\exists$ a number $N_\epsilon^{(o)}$ providing for each i, an upper bound to the maximum number of disjoint geodesic balls of radius ϵ, filling up each $\mathcal{M}_i$, i.e., $N_\epsilon(\mathcal{M}_i) \leq N_\epsilon^{(o)}$, for each i.

Stated differently, the convergence with respect to d_G is related to a uniform control of the "geometric size" of the manifolds $\mathcal{M}_i$, as indicated by the number of balls of a given radius that is needed to fill up each $\mathcal{M}_i$ in the considered sequence. For example, in the case of a sequence of compact surfaces of bounded curvature converging in d_G to S^2, each of them can be filled up by $N_\epsilon(\mathcal{M}_i)$ maximum number of disjoint geodesic balls of radius ϵ, such that $N_\epsilon(\mathcal{M}_i) \leq N_\epsilon^{(o)}(S^2)\ \forall i$.

A metric space E is said to be *precompact*, if $\forall \epsilon > 0$, $\exists$ a finite (open) covering B_j of E, such that the sets B_j have diameter $< \epsilon$. Equivalently, $\forall \epsilon > 0$, there exist a finite set

$F \subset E$, such that $d(x, F) < \epsilon, \forall x \in E$.
A stronger notion is that of compactness, yielded by the closure of the considered space.

Theorem 2 *The set $Ric[n, k, D]$ of isometry classes of compact manifolds, with the Ricci tensor satisfying $Ric(g) \geq (n-1)kg$ and diameter $\leq D$, $(k \in \mathbb{R}, D \in \mathbb{R}_+)$, is precompact when endowed with the Gromov distance d_G.*

This theorem, due to M.Gromov, states that in the set of closed riemannian manifolds (with Ricci curvature bounded below, diameter bounded above) there is a subset, let us call it $\tilde{Ric}$, containing for each ϵ, a finite number of riemannian manifolds $\tilde{\mathcal{M}}_j$, such that for any $\mathcal{M} \in Ric[n, k, D]$, we have $d_G(\mathcal{M}, \tilde{\mathcal{M}}_i) < \epsilon$ for some $\tilde{\mathcal{M}}_i \in \tilde{Ric}$.

What this means is that for each "length scale ϵ", there exists a finite number of "model" geometries, which describes with an ϵ-approximation any given riemannian geometry. Given a ball of a certain radius $> \epsilon$ in any riemannian manifold $\mathcal{M}$ in $Ric[n, k, D]$, there exists a ball metrically similar (up to an ϵ scale) in one of the "model" geometries, which does not retain the details of the original manifold on scales smaller than ϵ. Roughly speaking, ϵ is a measure of the typical curvature inhomogeneity with respect to the model–background.

Let us stress that this is a highly non-trivial result, in the sense that the metrical properties of the manifolds in the infinite dimensional set $Ric[n, k, D]$[8] are up to an ϵ scale described by the metrical properties of just a finite number of "model" riemannian manifolds.

However, since $Ric[n, k, D]$ is only precompact, and not compact we can have for instance a situation where a sequence of manifolds in $Ric[n, k, D]$ converges under d_G, to a manifold of lower dimension[9], or to a space with singularities[10].

Therefore below we will limit ourselves to the subset of $Ric[n, k, D]$, generated by those riemannian manifolds with sectional curvatures bounded in absolute value.

[8]This set is of infinite dimension because a point $\mathcal{M}$ in its interior, remains on the set under small perturbations of the metric, so locally it is in principle as complicated as the set of all riemannian metrics.

[9]The dimension of a manifold is not continuous for the topology defined by d_G.

[10]Phenomena of this kind are, for example the pinching of a geodesic in a torus.

Theorem 3 *The set* $(\tilde{Ric}[n,k,D], d_G)$ *of riemannian structures having diameter* $\leq D$*, volume* $\geq V$ *and sectional curvatures bounded below in absolute value, is compact.*

In certain sense, one can think of the "model" manifolds $\tilde{\mathcal{M}}_i \in \tilde{Ric}$ as the "smoothed out" counterparts of the manifolds in $\tilde{Ric}[n,k,D]$.

A connection can be made between the above theorems (due to M.Gromov) and the Ricci-Hamilton flow [52].

Hamilton's flow associated with a 3−geometry satisfying the rather weak conditions of theorem 2, evolves in a set $Ric[n,k,D]$ which is precompact when endowed with d_G. As is known, precompactness of a set is not a condition strong enough for yielding a globalization of a local (smooth) flow which evolves in it. But if $Ric(g) > 0$ is required for a closed 3−manifold $(\mathcal{M}, g)$ then the associated Hamilton's initial value problem (B.11) defines a flow $(g,\beta) \to g(\beta)$, $0 \leq \beta < \infty$, in the compact set $\tilde{Ric}[n,k,D]$. The positivity requirement for $Ric(g)$, needed for the global version of Hamilton's theorem is a convenient way of controlling the growth of the diameter of the manifolds $(\mathcal{M}, g(\beta))$, which discriminates between the permanence of the Ricci-Hamilton flow in the compact set of smooth geometries versus the possibility of leaving this set and evolving toward a singular geometry.

More details can be found in [119], [109] and [54, 55].

Bibliography

[1] A.M.Abrahams and C.R.Evans, Phys. Rev. Lett. **70** (1993) 2980.

[2] M.E.Agishtein, A.A.Migdal, preprint PUPT-1311 (March 1992).

[3] M.E.Agishtein, A.A.Migdal, preprint PUPT-1253 (April 1991).

[4] J.Ambjørn, J.Jurkiewicz, Ch. Kristjansen, preprint NBI-HE-92-53 (July 1992).

[5] J.Ambjørn, Z.Burda and J.Jurkiewicz, C.F.Kristjansen, preprint NBI-HE-93-3 (Jan. 93).

[6] P.C.Argyres, V.P.Frolov, astro-ph/9412046.

[7] L.Ya.Arifov and A.V.Shayn, Abstracts of Contributed Papers to GR13, eds P.W.Lamberti, O.E.Ortiz (1992).

[8] V.I.Arnol'd, *Ordinary Differential Equations*, Springer (1992).

[9] R.Arnowitt, S.Deser, C.Misner, in *Gravitation. An Introduction to Current Research*, ed. L.Witten Wiley N.Y. (1962).

[10] H.C.Arp and T. Flandern, Phys. Lett. A**164** (1992) 263.

[11] H.Arp, J. Astrophys. Astron. **8** (1987) 241.

[12] D.K.Arrowsmith, C.M.Place, *Ordinary Differential Equations*, Chapman and Hall (1982).

[13] D.K.Arrowsmith, C.M.Place, *An Introduction to Dynamical Systems*, Cambridge (1990).

[14] E.G.Avramidy, A.O.Barvinsky, Phys. Lett. B**159** (1985) 269.

[15] D.C. Backer, R.H. Helings, Ann. Rev. Astron. Astrophys. **24** (1986) 537.

[16] P.Bak, C.Tang, K.Wiesenfeld, Phys. Rev. Lett. **59** (1987) 381 and Phys. Rev. A**38** (1988) 364; P.Bak, K.Chen, C.Tang, Phys. Lett. A**147** (1990) 297; see also

P.Bak and M.Paczuski, Phys. World (Dec. 1993) 39; P.Bak and K.Chen, Sc. Am. (Jan. 1991) 46.

[17] M.N.Barber, Phys. Rep. **29** No 1 (1977) 1-84.

[18] J.M.Bardeen, Phys. Rev.D **22** (1980) 1882.

[19] J.D.Barrow, Quart. J. R.A.S. **30** (1989) 163.

[20] P.H.Bérard, Bull. Amer. Math. Soc. **19** 2 (1988) 371.

[21] A.Berera and Fang Li–Zhi, Arizona Univ. preprint AZPH-TH/93-5.

[22] M.Berger, D.Ebin, J. Diff. Geom. **3** (1971) 379.

[23] S.Bergh, J. Roy. Astr. Soc. Can. **84** (1990) 275.

[24] P.G.Bergmann, Rev. Mod. Phys. **33** (1961) 510.

[25] O.Bertolami, J.M.Mourão, J.Pérez-Mercader, Phys. Lett. B **311** (1993) 27.

[26] B.Bertotti, Proc. R. Soc. **294** (1966) 195.

[27] A.L.Besse, *Einstein Manifolds*, Springer-Verlag (1987).

[28] A.V.Bialko, ZhETF **65** (1973) 849; Sov. Phys. JETP **38** (1974) 421.

[29] S.Bildhauer, Prog. Theor. Phys. **84** (1990) 444.

[30] S.Bildhauer and T.Futamase, Gen. Rel. Grav. **23** (1991) 1251.

[31] J.J.Binney, N.J.Dowrick, A.J.Fisher and M.E.J.Newman, *The Theory of Critical Phenomena. An Introduction to the Renormalization Group*, Clarendon Press (1993).

[32] R.D. Blandford, R. Narayan, Ann. Rev. Astron. Astrophys. **30** (1992) 311.

[33] H.Bondi, M.N. R.A.S. **107** (1947) 410.

[34] W.B.Bonnor, Ap. J. **316** (1987) 49.

[35] W.B.Bonnor and G.F.R.Ellis, M.N.R.A.S. **218** (1986) 605.

[36] S.Borgani, *Scaling in the Universe*, Phys. Rep. (in press) and references therein.

[37] S.Borgani et al, Ap. J. *accepted* (1994).

[38] J.P.Bourguignon, in *Lecture Notes in Mathematics* **838**, Berlin Springer (1980).

[39] M.J.Bowick, E.Marinari, Gen. Rel. Grav. **24** (1992) 1209.

[40] P.R.Brady, Class. Qu. Grav. **11** (1994) 1255; Newcastle Univ. preprint NCL 94-TP12.

[41] D.R.Brill, J.B.Hartle, Phys. Rev. **135** (1964) B271.

[42] T.J.Broadhurst et al, Nature **343** (1990) 726.

[43] M.Bruni, Phys. Rev. D **47** 2 (1993) 738.

[44] M.Bruni, Ph.D. Thesis, SISSA (1991).

[45] M.Bruni, P.K.S.Dunsby and G.F.R.Ellis, Ap. J. **395** (1992) 34.

[46] M.Bruni, D.H.Lyth, preprint SISSA 112/93/A; Phys. Lett. B **323** (1994) 118.

[47] M.Bruni, K.Piotrkowska, M.N.R.A.S. **270** (1994) 630.

[48] C.G.Callan, Phys.Rev.D**2** 1541 (1970); K.Symanzik, Commun. Math. Phys. **18** (1970) 227.

[49] M.Cantor, Bull. Am. Math. Soc. **5** (1981) 235.

[50] J.L.Cardy (ed.), *Finite size scaling*, North-Holland (1988).

[51] M.Carfora and A.Marzuoli, Phys. Rev. Lett. **53** (1984) 2445.

[52] M.Carfora and A.Marzuoli, Class. Qu. Grav. **5** (1988) 659.

[53] M.Carfora and A.Marzuoli, in *Proceed. of 7-th Italian* Conference on Gen. Rel. and Gravit. Phys., World Sc. (1986).

[54] M.Carfora, A.Marzuoli, Critical Phenomena for Riemannian manifolds: simple homotopy and simplicial gravity, unpublished.

[55] M.Carfora and A.Marzuoli, Class. Qu. Grav. **9** (1992) 595.

[56] M.Carfora, J.Isenberg & M.Jackson, J.Diff. Geom. **31** (1990) 249.

[57] M.Carfora, K.Piotrkowska, to be submitted.

[58] G.J.Chaitin, *Algorithmic Information Theory*, C.U.P. (1987).

[59] M.W.Choptuik, in *Approaches to Numerical Relativity* R. d'Inverno, ed. C.U.P. (1992).

[60] M.W.Choptuik, Phys. Rev. Lett. **70** (1993) 9.

[61] R.G.Clowes and L.E.Campusano, M.N.R.A.S. **249** (1991) 218.

[62] P.Coles and G.F.R.Ellis, submitted to Nature, 1994.

[63] A.A.Coley and R.K.Tavakol, Gen. Rel. Grav. **24** (1992) 835.

[64] G.M.Covarrubias, Int. J. Theor. Phys. **32** (1993) 2135.

[65] T.Damour, J.H.Taylor, Ap. J. **366** (1991) 501.

[66] David F., Lectures given at Les Houches School *Gravitation et Quantifications* (1992).

[67] P.G.De Gennes, in *Fluctuation, Instability and Phase Transition* Proc. NATO Adv. Study Inst., Geilo, Norway, ed. T. Riste, Noordhoff, Leiden, ser. B (1975).

[68] D.DeTurck, J. Diff. Geom. **18** (1983) 157.

[69] P.A.M.Dirac, in *Gauge theories, massive neutrinos and proton decay* Proceed. of Orbis Sci., ed. B.Kursunoglu, A.Perlmutter, Plenum N.Y. and London (1981).

[70] H. Dingle, M.N.R.A.S. **94** (1933) 134.

[71] B.P.Dolan, Int. J. Mod. Phys.A**9** (1994) 1261.

[72] C.Domb, M.S.Green eds, *Phase Transitions and Critical Phenomena* (vol. 6) Academic Press London (1976).

[73] P.K.S.Dunsby, Clas. Qu. Grav. **8** (1991) 1785.

[74] P.K.S.Dunsby, M.Bruni and G.F.R.Ellis, Ap. J. **395** (1992) 54.

[75] A.S.Eddington, *Mathematical Theory of Relativity*, Cambridge (1924).

[76] G.Efstathiou et al, M.N.R.A.S. **247** (1990) 10; W.Saunders et al, Nature **349** (1991) 32.

[77] J.Ehlers, *E.B. Christoffel* ed. P.L.Butzer and F.Feher Basel Birkhauser (1981).

[78] J.Ehlers and W.Rindler, M. N. R.A.S. **238** (1989) 503.

[79] J.Einasto and M.Gramann, Ap. J. **407** (1993) 443.

[80] A.Einstein, *Meaning of Relativity*, London (1951).

[81] G.F.R.Ellis, in *Proc. of 2nd Canadian* Conference on General Relativity and Astrophysics, (eds) A.Coley et al. World Sc. (1988).

[82] G.F.R.Ellis, in *General Relativity and Cosmology* Proceed. of XLVII E.Fermi Summer School, (ed.) R.K.Sachs, Academic Press (1971), p.104.

[83] G.F.R.Ellis, in *General Relativity and Gravitation*, (eds) B.Bertotti et al. D.Reidel Publishing Company (1984).

[84] G.F.R.Ellis, Gen. Rel. Grav. **2** (1971) 7.

[85] G.F.R.Ellis et al, Phys. Rep. **124** (1985) 315.

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

[87] G.F.R.Ellis, M.Bruni and J.Hwang, Phys. Rev. D **42** (1990) 1035.

[88] G.F.R.Ellis, J.Hwang and M.Bruni, Phys. Rev.D **40** (1989) 1819.

[89] G.F.R.Ellis and M.J.Jaklisch, Ap. J. **346** (1989) 601.

[90] G.F.R.Ellis, D.R.Matravers, R.Trzeciokas, Gen. Rel. Grav. **15** (1983) 931.

[91] G.F.R.Ellis and D.W.Sciama, in *General Relativity*, (ed.) L.O.'Raiffeartaigh, O.U.P., Oxford (1972).

[92] G.F.R.Ellis and W.Stoeger Class. Qu. Grav. 4 (1987) 1697.

[93] J.Ellis, CERN-TH 6723/92 (talk presented at the First CERN-Torino Meeting on Current Trends in Particle and Condensed Matter Physics, May 11-13 1992).

[94] A.C.D.van Enter, R.Fernández, A.D.Sokal, Phys. Rev. Lett. **66** (1991) 3253; also J. Stat. Phys. **72** Nos 5/6 (1993).

[95] G.Esposito, Fortsch. der Physik **40** (1992) 1.

[96] C.R.Evans, in *General Relativity and Gravitation 1992*, Proceed. of GR13, eds R.J.Gleiser, C.N.Kozameh, O.M.Moreschi, IOP (1993).

[97] C.R.Evans, J.S.Coleman, Phys. Rev. Lett. **72** (1994) 1782 (gr-qc/9402041).

[98] G.Eyink, N.Goldenfeld, cond-mat/9407021.

[99] R.Fabbri and F.Melchiorri, Gen. Rel. Grav. **13** (1981) 201.

[100] H.V.Fagundes and S.F.Knok, Ap. J. **368** (1991) 337.

[101] M.E.Fisher, in *Critical Phenomena* Lecture Notes in Physics **186** ed. F.J.W.Hahne, Springer-Verlag Berlin (1983), p.82

[102] T.Frankel, *Gravitational Curvature*, W.H.Freeman and Company (1979).

[103] D.S.Freed, K.K.Uhlenbeck, *Instantons and Four-Manifolds*, 2nd ed. Springer-Verlag, N.Y. (1991).

[104] T.Futamase, Phys. Rev. Lett. **61** (1988) 2175.

[105] T.Futamase, M.N.R.A.S. **237** (1989) 187.

[106] T.Futamase, Prog. Theor. Phys. **86** No.2 (1991) 389.

[107] T.Futamase, Prog. Theor. Phys. **89** (1993) 581.

[108] T.Futamase and G.F.R.Ellis, An approximation to cosmological density fluctuations in the nonlinear stage, unpublished.

[109] S.Gallot, D.Hulin, J.Lafontaine, *Riemannian Geometry*, Springer-Verlag (1987).

[110] D.Garfinkle, gr-qc/9412008.

[111] M.J.Geller and J.P.Huchra, Science **246** (1989) 897.

[112] M.Gell-Mann and F.Low, Phys. Rev.**98** (1954) 1300.

[113] G.W.Gibbons, S.W.Hawking & J.M.Stewart, Nucl. Phys.B**281** (1987) 736.

[114] T.Goldman, J.Pérez-Mercader, F.Cooper, M.M.Nieto, Phys. Lett. B**281** (1992) 219.

[115] J.P.Gollub, Nature **367** (1994) 318.

[116] S.W.Goode, Phys. Rev. D **39** (1989) 2882.

[117] R.H.Gowdy, Ann. Phys. **83** (1974) 203.

[118] L.P.Grishchuk, Sov. Astr. A.J. **11** (1968) 881.

[119] M.Gromov, *Structures metriques pour les varietes riemanniennes*, Paris: Cedic/F Nathan (1981).

[120] S.R. de Groot and L.Suttorp, *Foundations of Electrodynamics*, Amsterdam North-Holland (1972).

[121] E.J.Groth and P.J.E.Peebles, Astron. Astrophys. **41** (1975) 143.

[122] K.Grove and P.V.Petersen, Annals of Math. **128** (1988) 195.

[123] R.Gupta, Scaling, the Renormalization Group and Improved Lattice Actions, Los Alamos preprint.

[124] R.S.Hamilton, J. Diff. Geom. **17** (1982) 255.

[125] R.S.Hamilton, J.Isenberg, Quasi-convergence of Ricci Flow for a Class of Metrics, preprint (1992).

[126] H.W.Hamber, preprint UCI-Th-92-51 (Oct.1992).

[127] Hanson, T.Regge, C.Teitelboim, *Constrained Hamiltonian Systems*, Acad. Naz. dei Lincei, Rome (1976).

[128] E.R.Harrison, Rev. Mod. Phys. **39** (1967) 862.

[129] E.R.Harrison, M. N. R.A.S. **140** (1968) 281.

[130] S.W.Hawking, Ap. J. **145** (1966) 544.

[131] S.W.Hawking and G.F.R.Ellis, *The Large Scale Structure of Space-Time*, C.U.P. (1973).

[132] C.Hellaby, Gen. Rel. Grav. **20** (1988) 1203.

[133] M.Heller, *Theoretical Foundations of Cosmology*, World Sc. (1992).

[134] A.Hemmerich, Astron. Astrophys. **185** (1987) 1.

[135] G.'t Hooft, Nucl. Phys.B **61** (1973) 455.

[136] J.Hubbard, P.Schofield, Phys. Lett. **40**A (1972) 245.

[137] R.H. Hulse, J.H. Taylor, Ap. J. Lett. **195** (1975) L51.

[138] J.Hwang and E.Vishniac, Ap. J. **353** (1990) 1.

[139] R.Isaacson, Phys. Rev. **166** (1968) 1263.

[140] R.Isaacson, Phys. Rev. **166** (1968) 1272.

[141] J.Isenberg, M.Carfora, M.Jackson, in *Proceed. of the Fifth* Marcel Grossmann Meeting on Gen.Rel., (eds) D.G.Blair, M.J.Buckingham, World Sc. (1989).

[142] J.Isenberg & M.Jackson, J. Diff. Geom. **35** (1992) 723.

[143] M.Jacobs, E.V.Linder and R.V.Wagoner, preprint SU-ITP-9233.

[144] M.W.Jacobs, E.V.Linder, R.V.Wagoner, Phys.Rev. D **45** (1992) R3292.

[145] D.A.Johnston, J.P.Kownacki, A.Krzywicki, hep-lat/9407018.

[146] B.J.T.Jones, in *Renaissance of GR and Cosmology*, eds G.Ellis, A.Lanza and J.Miller, C.U.P. (1993).

[147] L.P.Kadanoff, Physics **2** N.Y. (1966) 263.

[148] A.Karlhede, Gen. Rel Grav. **12** (1980) 693.

[149] M.Kasai, Phys. Rev. Lett. **69** (1992) 2330.

[150] M.Kasai, dissertation RITP Hiroshima University (Feb. 1988).

[151] K.I.Kellermann, Nature **361** (1993) 134.

[152] V.T.Khiet, Ukr. Fiz. Zh. **29** (1984) 969.

[153] Y.Kiem, preprint PUPT-1484 July 1994.

[154] C.W.Kim, preprint JAU-TIPAC 940020.

[155] C.W.Kim, A.Sinibaldi, J.Song, preprint DFF-215/12/94.

[156] H.Kodama and M.Sasaki, Prog. Theor. Phys. **78** (1984) 1.

[157] T.Koike, T.Mishima, preprint TIT/HEP-266/Cosmo-46.

[158] E.W.Kolb and M.S.Turner, *The Early Universe*, Addison - Wesley (1990).

[159] A.A.Kosiński, *Differentiable Manifolds*, Academic Press (1993).

[160] D.Kramer, H.Stephani, M.A.H.MacCallum and E.Herlt, *Exact Solutions of Einstein's Field Equations*, C.U.P. Cambridge (1980).

[161] A.Krasiński, *Physics in an Inhomogeneous Universe*, to be published by C.U.P.

[162] J.Kristian, Ap. J. **147** (1967) 864.

[163] J.Kristian, R.K.Sachs, Ap. J. **143** (1966) 379.

[164] Y.K.Lau, preprint AS-ITP-93-27.

[165] Y.W.Lee, Astron. J. **104** (1992) 1780.

[166] J.Lelong-Ferrand, in *Differential Geometry and Relativity*, eds M.Cahen and M.Flato, D.Reidel Pub. Co., Dordrecht (1976).

[167] G.Lemaître, Ann. Soc. Sci. Bruxelles A**53** (1933) 51.

[168] A.R.Liddle and D.H. Lyth, Phys. Rep. **231** (1993) 1.

[169] E.M.Lifshitz, J.Phys. USSR **10** (1964) 116.

[170] E.M.Lifshitz, ZhETF **16** (1946) 587.

[171] E.M.Lifshitz, I.M.Khalatnikov, Usp. Fiz. Nauk **80** (1963) 391.

[172] E.M.Lifshitz and I.M.Khalatnikov, Adv. Phys. **12** (1963) 185.

[173] A.Linde, preprint SU-ITP-94-36 (hep-th/9410082); A.Linde, A.Mezhlumian, D.Linde, preprint SU-ITP-94-39.

[174] R.W.Lindquist, J.A.Wheeler, Rev. Mod. Phys. **29** (1957) 432.

[175] J.Lott, Commun. Math. Phys. **107** (1986) 165.

[176] C.O.Lousto, Nucl. Phys. B **410** (1993) 155.

[177] D.H.Lyth and E.D.Stewart, Ap. J. **361** (1990) 343.

[178] R.Maartens, G.F.R.Ellis, W.R.Stoeger, Univ. of Cape Town preprint 1994/13.

[179] M.A.H.MacCallum, in *GR – an Einstein Centenary Survey*, eds. S.W.Hawking and W.Israel, CUP (1979), p.533.

[180] M.Maggiore, preprint IFUP-TH 22/94 (hep-th/9404172).

[181] L.S.Marochnik, N.V.Pelikhov, G.M.Vereshkov, Astrophys. Space Sci. **34** (1975) 249.

[182] L.S.Marochnik, N.V.Pelikhov, G.M.Vereshkov, Astrophys. Space Sci. **34** (1975) 281.

[183] L.S.Marochnik, P.D.Naselsky, N.V.Pelikhov, Astrophys. Space Sci. **67** (1980) 261.

[184] L.S.Marochnik, Astron. Zh. **57** (1980) 903; Sov. Astr. A.J. **24** (1981) 518.

[185] L.S.Marochnik, Astron. Zh. **57** (1980) 1129; Sov. Astr. A.J. **24** (1980) 651.

[186] L.S.Marochnik, Astrophys. Space Sci. **69** (1980) 3 and 31.

[187] D.R.Matravers, G.F.R.Ellis, W.R.Stoeger, Univ. of Cape Town preprint 1994/10.

[188] R.A.Matzner, J. Math. Phys. **9** (1968) 1063.

[189] R.A.Matzner, J. Math. Phys. **9** (1968) 1657.

[190] W.H.McCrea, Z.Astrophysik **18** (1939) 98.

[191] C.W.Misner, Phys. Rev. D**8** (1973) 3271.

[192] C.W.Misner, J.A. Wheeler, K.S.Thorne, *Gravitation* W.H.Freeman & Company, San Francisco (1973).

[193] J.W.Moffat and D.C.Tatarski, submitted to Ap. J. (astro-ph/9407036).

[194] V.Moncrief, Phys. Rev. D**16** (1977) 1702.

[195] M.Morikawa, in Abstracts of Contributed Papers GR13, eds P.W.Lamberti, O.E.Ortiz, Córdoba (1992).

[196] V.F.Mukhanov, H.A.Feldman and R.H.Brandenberger, Phys. Rep. **215** (1992) 203 - 333.

[197] A.H.Nelson, M.N.R.A.S. **158** (1972) 159.

[198] J.F.Nicoll, I.E.Segal, Proc. Nat. Acad. Sci. (U.S.A.) **72** (1975) 4691.

[199] T.W.Noonan, Gen. Rel. Grav. **16** (1984) 1103.

[200] T.W.Noonan, Gen. Rel. Grav. **17** (1985) 535.

[201] D.O'Connor and C.R.Stephens, in *Directions in General Relativity* Proc. 1993 Int. Symposium Maryland, vol.1, eds B.L.Hu, M.P.Ryan Jr. and C.V.Vishveshwara, C.U.P. (1993).

[202] Y.Oshiro, K.Nakamura, A.Tomimatsu, Nagoya Univ. preprint DPNU-94-04.

[203] T.Padmanabhan, *Structure Formation in the Universe*, Cambridge (1993).

[204] M.J.Pierce et al, Nature **371** (1994) 385.

[205] J.Pullin, gr-qc/9409044.

[206] Raychaudhuri, *Theoretical Cosmology*, O.U.P. Oxford (1979).

[207] A.K.Raychaudhuri, in *Highlights in Gravitation and Cosmology*, ed. B.R.Iyer et al. C.U.P. (1988).

[208] A.K.Raychaudhuri, G.Mukherjee, M.N.R.A.S. **209** (1984) 353.

[209] T.Regge, Il Nuovo Cimento **XIX** (1961) 558.

[210] G.F.R.N.Rhee, Nature **350** (1991) 211.

[211] G.Rosen, Nuovo Cimento **B57** (1980) 125.

[212] H.C.Rosu, astro-ph/9411112.

[213] D.Ruelle, *Chance and Chaos*, Princeton University Press (1991).

[214] S.E.Rugh, to appear in NATO ARW *Deterministic Chaos in GR*, ed. D.Hobile Plenum Press N.Y. (1994).

[215] H.Russ et al, Phys. Rev. D **48** (1993) 4552.

[216] J.Russo, L.Susskind, L.Thorlacius, Phys. Rev. D **46** (1992) 3444; ibid. **47** (1993) 533 and **48** (1993) 5769.

[217] E.Saar, Eesti NSV Akadeemia Toimetised (Izv. Akad. Nauk Eston. SSR) 20 (1971) 420; Tartu Aston. Obs. Publ. **39** (1971) 206, 234 and 249.

[218] R.K.Sachs and A.M.Wolfe, Ap. J. **147** (1967) 73.

[219] R.K.Sachs, H.Wu, Bull. Amer. Math. Soc. **83** no.6 (1977) 1101, sec. 4.2.

[220] R.Sancisi and T.S. van Albada, in *Dark Matter in the Universe*, (eds) J.Kormendy and G.Knapp, Reidel Dordrecht (1987).

[221] A.Sandage, Ap. J. **331** (1988) 583; *ibid* **350** (1990) 631.

[222] A.Sandage et al, Ap. J. **423** (1994) L13.

[223] R.H.Sanders, Astron. Astrophys. Rev **2** (1990) 1.

[224] R.Scaramella, Ap. J. **390** (1992) L57.

[225] S.Scheuer, Nature **361** (1993) 112.

[226] P.Schneider, J.Ehlers, E.E.Falco *Gravitational Lenses*, Springer-Verlag Berlin (1992).

[227] D.W.Sciama, *Modern Cosmology*, Cambridge C.U.P. (1971).

[228] R.Shankar, Rev. Mod. Phys. **66** (1994) 129.

[229] M.F.Shirokov and I.Z.Fisher, Astron. Zh. **39** (1962) 899 (in Russian); Sov. Astr. A.J. **6** (1963) 699.

[230] M.F.Shirokov, in *Contemporary Problems of Gravitation* Collection of proceed. of the 2-nd Soviet Conference on Gravitation, Tbilisi University Publishing House (1967), p.376.

[231] N.R.Sibgatullin, Doklady ANSSSR **200** (1971) 308; Sov. Phys. Doklady **16** (1972) 697.

[232] L.Smolin, Guanajuato Lectures Jan. 1995.

[233] G.F.Smoot et al, Ap.J. **396** (1993) L1; E.L.Wright et al, Ap.J. **396** (1993) L13.

[234] A.Spero and R.Baierlein, J. Math. Phys. **18** (1977) 1330.

[235] A.Spero and R.Baierlein, J. Math. Phys. **19** (1978) 1324.

[236] J.M.Stewart, Class. Qu. Grav. **7** (1990) 1169.

[237] J.M.Stewart, M.Walker, M.N. R.A.S. A**341** (1974) 49.

[238] W.R.Stoeger, G.F.R.Ellis, C.Hellaby, M.N.R.A.S. **226** (1987) 373.

[239] W.R.Stoeger, G.F.R.Ellis, B.G.Schmidt, Univ. of Cape Town preprint 1991/1.

[240] A.Strominger, L.Thorlacius, preprint NSF-ITP-93-144.

[241] E.C.G.Stuckelberg and A. Petermann, Helv. Phys. Acta **26** (1953) 499.

[242] Y.Suto, T.Suginohara, Y.Inagaki, submitted to Prog. Theor. Phys.

[243] P.Szekeres, Ann. Phys. **64** (1971) 599.

[244] P.Szekeres, Commun. Math. Phys. **41** (1975) 55.

[245] R.K.Tavakol and G.F.R.Ellis, Phys. Lett. A **130** (1988) 217.

[246] W.P.Thurston, *The Geometry and Topology of Three Manifolds* Lecture Notes, Princeton Univ. (1978).

[247] W.P.Thurston and J.R.Weeks, Sc.Amer. (July 1984) 94.

[248] R.C.Tolman, Proc. Nat. Acad. Sci. USA **20** (1934) 169.

[249] J.E.Traschen, Phys. Rev. D **29** (1984) 1563.

[250] J.E.Traschen, D.M.Eardley, Phys. Rev. D **34** (1986) 1665.

[251] G.de Vaucouleurs, Science **157** (1971) 1223.

[252] S.Weinberg, *Gravitation and Cosmology*, Wiley N.Y. (1972).

[253] S.Weinberg, Phys. Rev. D **8** (1973) 3497.

[254] B.Widom, J. Chem. Phys. **43** (1965) 3898.

[255] K.G.Wilson, Rev. Mod. Phys. **47** (1975) 773.

[256] K.G.Wilson, Sc.Amer. **241** (2), (1979) 140.

[257] K.G.Wilson, Phys. Rev. B4 (1971) 3174.

[258] K.G.Wilson, Phys. Rev. B4 (1971) 3174.

[259] K.Wilson and J.Kogut, Phys. Rep. C**12** (1974) 75.

[260] C.M.Will, Phys. Rep. **113** (1984) 345.

[261] A.Woszczyna, Phys. Rev. D **45** (1992) 1982.

[262] A.Woszczyna and A.Kulak, Class. Qu. Grav. **6** (1989) 1665.

[263] A. Woszczyna, A.Kulak and S.Zieba, Class. Qu. Grav. **8** (1991) L167.

[264] A.B.Zamolodchikov, JETP Lett. **43** (1986) 730.

[265] E.C.Zeeman, Nonlinearity **1** (1988) 115.

[266] N.V.Zotov and R.W.Stoeger SJ, Class. Qu. Grav. **9** (1992) 1023.