IFT

Instituto de Física Teórica
Universidade Estadual Paulista

TESE DE DOUTORAMENTO IFT–T.02/15

A GEOMETRIC APPROACH TO COSMOLOGICAL BOUNDARY CONDITIONS

Lucas Lolli Savi

Orientador

José Geraldo Pereira

Abril de 2015

Acknowledgments

I would like to acknowledge the advice of Prof. J. G. Pereira, the useful conversations with my colleague H. Jennen and the financial support by CAPES.

Abstract

The observed T-asymmetry of macroscopic physics is traced back to the extremely low entropy configuration of the early universe. Although matter was very hot and with a uniform temperature at that stage, the gravitational degrees of freedom were largely suppressed, which fact contributes to the lowness of the entropy and is encoded in the high level of spatial symmetry (nearly Friedman-Lemaître-Robertson-Walker character) of the last scattering surface. I analyze different attempts to explain the origin of such special configuration. The inflation paradigm is probed with respect to this problem, and it is concluded that the initial low entropy cannot be accounted for within it. Similar conclusions are reached with respect to statistical (i.e. anthropic) reasonings. On the other hand, the paradigm known as conformal cyclic cosmology (CCC) presents itself as a new alternative which surpasses many of the difficulties faced by its rivals, although raising its own open questions. I introduce the model together with the mathematical structure of Cartan geometries as a possible means of achieving a better understanding of cosmological boundary conditions. One element which is crucial in this analysis is the modeling of the Cartan geometric structure over a de Sitter space $SO_e(4,1)/SO_e(3,1)$ with varying length parameter. The introduction of a length parameter in the kinematics is favored by the observation of a positive cosmological constant and also desirable for quantum gravity reasons, due to the natural scale set by Planck's constant.

Keywords: cosmological boundary conditions, conformal cyclic cosmology, cone space-time, de Sitter relativity, Cartan geometries.

Subjects: gravitation, cosmology, kinematics, temporal asymmetry.

Resumo

A assimetria temporal observada na física macroscópica se deve à configuração de entropia extremamente baixa do universo primordial. Apesar de a matéria estar muito quente e com uma temperatura uniforme naquele estágio, os graus de liberdade gravitacionais estavam em grande medida suprimidos, fato este que contribui para o baixo valor da entropia e está codificado no alto grau de simetria espacial (caráter aproximadamente Friedman-Lemaître-Robertson-Walker) da superfície de último espalhamento. Analisamos diferentes tentativas de explicar a origem de tal configuração especial. O paradigma inflacionário é testado com respeito a esse problema, e é concluído que a baixa entropia inicial não pode ser explicada dentro dele. Conclusões similares são obtidas com respeito a formulações estatísticas (i.e. antrópicas). Por outro lado, o paradigma conhecido como cosmologia cíclica conforme (CCC) se apresenta como uma nova alternativa que ultrapassa muitas das dificuldades enfrentadas pelos seus rivais, apesar de levantar suas próprias questões em aberto. Introduzimos o modelo juntamente com a estrutura matemática das geometrias de Cartan como um meio possível de atingir um melhor entendimento das condições de contorno cosmológicas. Um elemento que é crucial nessa análise é a modelagem de uma estrutura geométrica de Cartan sobre o espaço de de Sitter $SO_e(4,1)/SO_e(3,1)$ com um parâmetro de comprimento variável. A introdução de um parâmetro de comprimento na cinemática é favorecida pela observação de uma constante cosmológica positiva e também desejável por motivos oriundos da gravitação quântica, devido à escala natural determinada pelo comprimento de Planck.

Contents

Chapter 1

Introduction

1.1 Quantum gravity and kinematics with invariant length

It is expected that quantum effects on the fabric of spacetime become significant at the Planck scale, which is the natural scale of nature. This means that an invariant scale (which can be thought of as a scale of length as well as one of time, mass, energy, momentum, curvature etc) should be present in the fundamental structure of spacetime [1]. Ordinary special and general relativity do not take this into account, and treat every scale in the same fashion, and spacetime as a smooth continuum. Doubly special relativity (DSR) is an attempt to solve this by modifying the kinematics of special relativity through the introduction of an invariant length parameter, corresponding to the Planck length [2–7]. In the limit when that parameter goes to zero, the smoothness of spacetime is recovered, and with it ordinary special relativity. Such hypothetical scale would then mean the scale of "lumpiness" or "indivisibility" of spacetime, the size of its fundamental "blocks". However, DSR has been facing several difficulties, and as yet has no unique and rigorous mathematical formulation [7]. Nevertheless, the idea of including an invariant interval scale in the kinematics, in terms of which the classical description of spacetime should be amended to account for quantum effects insists to make sense.

Another approach to the same problem and in the same spirit is de Sitter special relativity. This amounts to simply replacing Minkowski space and its Poincaré group of isometries by their more general counterparts, the de Sitter space and group. An invariant spacetime length parameter l is then naturally introduced in the kinematics, corresponding to the curvature radius of the de Sitter spacetime. The crucial difference between DSR and de Sitter relativity is that, in the former, ordinary special relativity is recovered in the limit when the length parameter goes to *zero*, while in the latter it is recovered in the limit when the length parameter goes to *infinity.* This might sound as a difficulty at first, since the Planck length is a very small rather than a very big length. A beautiful solution to this arises from imagining that such curvature radius is not constant, but varying in spacetime. In other words, spacetime in the large would not be a perfect de Sitter space, but an irregularly curved space whose local geometry at each point should reduce to that of a de Sitter space with a certain value of the parameter l depending on the point. This has the additional advantage that now gravitational effects can be included in the theory. This scenario is indeed analogous to ordinary *general* relativity, with the crucial difference that now spacetime looks locally like de Sitter space instead of Minkowski space. To allow for such a generalization, however, one has to make use of a kind of geometry more general than Riemannian geometry, called Cartan geometry [8, 9, 122, 123]. Basically, while Riemannian geometry consists of a manifold endowed with a structure which makes it locally "look like" an affine space (Minkowski space in the 4D Lorentzian case), Cartan geometry deals with a manifold endowed with a more general structure, such that it can look locally like any homogeneous manifold (physically, this means either Minkowski, de Sitter or anti-de Sitter space, or the cone space described in section 5.4).

But how does allowing for the curvature radius of the local de Sitter model geometry to vary in spacetime solve the problem of the Planck length being very small instead of very large? The answer reveals itself naturally if we now look at the Planck scale not as a length scale, but as an energy, or better yet, an energy density one. Recall that general relativistically ($c = G = 1$) energy density has the same dimension as inverse area, and hence corresponds naturally to a cosmological constant, and to a de Sitter length parameter l. We can then naturally associate the curvature radius l of the local model de Sitter space to the local energy density at each point. Introducing then the quantum scale $\hbar = 1$, we get that in the limit when the energy density approaches the Planck (energy density) scale, the parameter l should approach the Planck length. In the converse limit, when energy density vanishes, l should go to infinity and the local structure should be just that of Minkowski space. So, instead of expecting spacetime to exhibit an *a priori* structure with a universal length scale, as is done in DSR, we expect that the matter content at each region dictates the spacetime structure there, and how much it should deviate from the geometry of empty space, that is, Minkowski space.

So in an energy density scale close to the Planck energy density, l would be very small, and considerable deviations from ordinary relativistic kinematics should be expected. However, in almost every region of spacetime accessible to our observation, and within our current resolution, the *averaged* energy density lies far below such order, and hence we usually observe phenomena develop in accordance with ordinary relativity. According to de Sitter relativity, though, this should fail to be true if we could probe much greater densities. Although observable effects on particle experiments are not to be expected with current technology (just like ordinary gravitational effects are negligible in the same context), highly energetic astronomical phenomena might present such evidence. Also, cosmological data might bear an imprint of such effect due to their cumulative character, that is, taking into account the approximate homogeneity observed of the universe at cosmological scales, there should be a uniform averaged residual de Sitter curvature on those scales, corresponding to the mean cosmological energy density. This should work exactly like a cosmological constant (positive or negative depending on the model geometry being a de Sitter or anti-de Sitter space). Since an accelerated expansion compatible with a positive Lambda is favored by present data, we suppose de Sitter space to be the model geometry of our universe, and a way to test the theory is to calculate the value it predicts for such residual de Sitter term and compare it with the observed one. If agreement is achieved, de Sitter relativity provides a natural solution to the dark energy problem.

So, unlike DSR and similar models, de Sitter relativity does not a priori advocate a "minimum length", let alone an actual rigid division of spacetime into blocks of a certain size, although it also does not in principle forbid a further modification which might implement such features. However, it still does accomplish the proposed goal of such models, which is to include testable quantum effects (coming from the intrinsically quantum invariant scale $\hbar = 1$) into the fabric of spacetime. Another advantage of the de Sitter approach against competing proposals of deforming the Poincaré group is that Lorentz symmetry, which is responsible for causality, is explicitly preserved throughout. All that is changed in the isometry group are the translations, which now correspond to motions on the de Sitter pseudo-sphere, and hence are not commutative anymore.

Apart from the quantum gravity motivation presented above, a Cartan geometry based on de Sitter space (which I will follow [119] in calling *de Sitter-Cartan geometry* from now on) is also motivated by cosmology. First and most simply, the mere existence of a positive cosmological term renders it much more natural to describe spacetime in terms of a de Sitter-Cartan structure, in which the cosmological term is naturally included, than in terms of a Riemannian structure, where the cosmological term is an added feature. But beyond that, a de Sitter-Cartan geometry with varying length parameter may naturally produce the geometric features of conformal cyclic cosmology (CCC), a kind of cosmological model which will be discussed later, and that if correct might solve the question of the the macroscopic temporal asymmetry of our universe. Let us pass to this question now.

1.2 Temporal asymmetry and statistics

Although fundamental physics is strictly CPT-invariant, macroscopic phenomena exhibit a clear temporal asymmetry. The human feeling of the passing of time in a specific temporal direction might be the most obvious of these, giving rise to the impression that time has a preferred orientation. But all around the macroscopic world, phenomena in general tend to have a consistent directionality in time, or "irreversibility", which marks a clear distinction between both temporal directions — that direction along which, for example, wood turns into ashes but never the opposite, and that in which ashes turn into wood but never the opposite.

The work of such people as Carnot, Joule, Clausius and Kelvin on the mutual conversion between heat and work first brought light to this mystery by noticing the general tendency of thermodynamic systems towards energy dissipation, that is, consumption of free energy. The concept of entropy, which Clausius defined through the formula

$$dS = \frac{dQ}{T},$$

came to be understood as a measure of the amount of energy which is not available to perform work. All the so-called irreversible processes were thus seen to be manifestations of a general law: that entropy never decreases in isolated systems.

The statement above is of course just the popular enunciation of the second law of thermodynamics. However, for such an enunciation to make sense, one temporal direction must have already been assigned as the correct one. In fact, it is implicit in such statements that the psychological direction of time is being chosen instead of the opposite one. The problem with stating a physical law with respect to the time direction which is perceived to be preferred by the human mind is that there is nothing in fundamental physics to back up such perception. In fact, this very perception is most likely itself the result of thermodynamic processes in some way, given the very nature of living systems, so that the general entropy gradient of nature and the human feeling of time are hardly independent [10]. Because of this, it makes more sense to state the second law as saying that entropy in all isolated systems has a monotonic temporal gradient, which is consistent in sign across the observable universe. Whether this sign is positive or negative depends upon an arbitrary choice of temporal direction[1].

The acknowledgment of the consistent entropy gradient represents a major scientific advance, since it exposes all previously known irreversible processes as particular instances of a general tendency. That might righteously be considered one of the most important unifications in the history of physics. Still, the theory of thermodynamics alone must assume such tendency as a postulate. Its explanation, however, came soon after, or at least part of it, when the problem was handed from thermodynamics to the newly emerging field of statistical mechanics.

Following work by Maxwell on the kinetic theory of gases, which applied statistics to physical systems composed of a great number of particles to derive distributions of quantities such as molecular velocities, Boltzmann arrived at an independent definition

[1]For this reason, the traditional expression "arrow of time", or "time's arrow", coined and popularized by Eddington [11], is avoided in this work, and replaced by "consistent monotonic entropy gradient", or "entropic temporal asymmetry". Since such gradient by itself does not endow time with a preferred direction, but only shows a "lop-sidedness" in the behavior of entropy with respect to it, it is clear that the only actual "arrow of time" is the psychological one, which remains a problem for the sciences of the brain.

of entropy as a property of a macroscopic state of a system, with respect to some coarse-graining of its phase space, which reflects the degree of ignorance with respect to the actual microscopic state. In other words, entropy is a measure of the number W of microstates compatible with the observed macrostate. However, since this number is multiplicative and Clausius's entropy is additive with respect to the size of the system, entropy must be the logarithm of that number. Thus Boltzmann arrived at the formula for the entropy of a macrostate:

$$S = -k \log W. \tag{1.1}$$

The constant k is necessary to achieve accordance with Clausius's units, namely energy per temperature, while the negative sign makes the quantity positive. This concept was taken up and mathematically refined by Gibbs, who noticed that Boltzmann's formula only applies in the case of equiprobable microstates. That is the case of an ideal gas of identical particles, but not the general one. Gibbs then devised the general formula[2]

$$S = -k \sum_i p_i \log p_i, \tag{1.2}$$

with p_i being the probability of the microstate i, and the sum is over all microstates composing the macrostate considered. When all p's are equal, this reduces to (1.1). In the continuum limit, (1.2) becomes:

$$S = -k \int dX p(X) \log p(X), \tag{1.3}$$

where the integral is taken over the whole space of microstates X compatible with the referred macrostate. For Hamiltonian systems with equiprobable microstates, this can also be expressed as

$$S = -k \log V,$$

where V is the phase space volume corresponding to the macrostate considered. All the above definitions agree with the Clausius formula when the latter can be applied, that is, for equilibrium states.

At this point, entropy started to be understood as a partly subjective concept, dealing with the information which is not accessible to an examiner. That view was ultimately vindicated by Shannon's development of the concept of information entropy [15], of which the statistical mechanics concept is a particular example. In Shannon's theory,

[2] The actual historical development of entropy formulas by Boltzmann and Gibbs might be more complicated than the simplified version presented here for didactic reasons [12]. Part of the difficulty in probing these matters is due to the relevant works of Boltzmann [13, 14] being still untranslated.

entropy is related to the amount of information which is expected to be gained from the disclosure of some message. In fact, entropy is a property of any probability distribution. The *self-information*, or *surprisal*, is a well defined notion in probability and information theory, which measures how unexpected an event is, or equivalently, the amount of new information gained by knowing that such event happened. If the probability of the event X is p, its surprisal is

$$I(X) = -\log p(X).$$

The entropy of a given probability distribution is just the expectation value of the surprisal,

$$S = E(I) = \int dX p(X) I(X) = -\int dX p(X) \log p(X),$$

which represents how much information is ignored. This gives exactly the formula (1.3), up to the dimensional constant. In the case of a thermodynamic system, each event X is a microstate, and the probability space is the space of of all microstates compatible with the observed facts, that is, the macrostate.

Since a greater entropy means a greater multiplicity of possibilities for the exact state of a system, the entropy of a macrostate is related to the probability that the system will be found occupying that macrostate instead of others. This understanding leads naturally to the result that if a system is let to evolve in time in a sufficiently random fashion, it has a statistical tendency, which is stronger in proportion to the size of the system, to occupy states of increasing entropy. For typical macroscopic systems such tendency approaches a law-like character. That is basically the content of Boltzmann's famous H-theorem, which was published as a proof of the second law, though with one caveat. Let us get to it.

At first, the considerations above seem to elucidate the second law of thermodynamics. This is what Boltzmann himself believed after he proved his theorem. However, this is not so, and the reason is simple, yet subtle: the tendency to increase the entropy works equally well in *both* time directions. This means that, even if one direction of time is deemed preferable for some external reason or convention, a system which is found in a state of not maximum entropy will not only most likely evolve towards states of increasing entropy into the *future*, but it also most likely reached the present state from states which had ever larger entropy towards the *past*. This is clear since the higher entropy states are the most probable ones. The very fact that the system is found in a state of low entropy is itself an unlikely one, and the a priori most likely reason for it is that it fluctuated away from the maximum entropy state, this one the only really probable one for large systems. In other words, if all which is known about a system is that it is in a state of entropy below maximum, the most likely extrapolation is that its entropy was maximum at the far past and will be maximum again in the far future, which means that the present state is the result of a fluctuation.

The behavior described above is perfectly symmetrical in time. That is of course what should be expected since the fundamental laws of physics are time-symmetric. This is the famous Lochschmidt objection towards Boltzmann: it is impossible to derive a time-asymmetric behavior only from time-symmetric laws. This means that there must be some component of time asymmetry already present in the derivation of the second law as it stands. The answer comes easily when one realizes that the solution of a dynamical problem depends not only on the laws, but also on the boundary conditions. Even with time-symmetric laws, there can be time-asymmetric boundary conditions, leading to a time-asymmetric behavior. That is actually the case of the H-theorem: Boltzmann had to assume that molecular velocities are uncorrelated in the (conventional) past [16,22,25]. This is a boundary condition, and a time-asymmetric one at that. With this assumption, then, the second law is obtained. This assumption adds new information, which was missing in the symmetric case. There, the only boundary condition was the present state of the system, and the entropy was unbounded at distant times in either temporal direction. In Boltzmann's case, a condition is set which fixes the entropy in one direction, since the lack of velocity correlations implies a low entropy state [16, 22, 25], but leaves it unconstrained in the other direction. The conclusion which is thus achieved is that, since the dynamics are time symmetric, any asymmetric behavior must follow from asymmetric boundary conditions.

In real systems, what we observe is that they did not reach their low entropies by random fluctuations from previous higher entropy states. Instead, we see entropy always increasing in the psychological direction of time, which means that even a state of relatively low entropy normally came from states of even lower entropy in the past. Since this observation is consistent all across the observed universe, the conclusion is that it has always been gaining in entropy, following the psychological direction. That can only mean that at some distant point in conventional past the entropy of the universe as a whole had a very low value. This might a priori be either the result of a boundary condition imposed by some fundamental cosmological principle, or else the result of a very big fluctuation, of literally cosmic proportions. But even in the second case, the result would be an *effective* cosmological boundary condition. So we known that the observed temporal asymmetry of all macroscopic physics is traced back to such a cosmological boundary condition. Thus the problem of macroscopic time asymmetry, which had already been handed from thermodynamics to statistical mechanics, is handed once more, now from the realm of statistical mechanics to that of cosmology. The point arrived thus far is already very advanced, in the sense that the temporal asymmetry of all macroscopic phenomena, from stellar dynamics to cooking, is reduced to one single cosmological boundary problem. But some serious questions still remain. For instance, which point is this where cosmic entropy is so low? How far is it from us? Is it a proper boundary condition or just an effective one, due to a statistical fluctuation? To advance any further and try to answer these question we must probe the fields of cosmology.

1.3 Cosmological boundary conditions

We have already seen briefly how macroscopic time-asymmetry was reduced to thermodynamic energy dissipation, and from that to the statistical tendency of systems to evolve towards more probable states. It stands to find out how the universe as a whole came to find itself in a state of very low entropy to begin with. Now to discuss the low entropy cosmological condition we need to understand some crucial properties of the current cosmological model. The first of these is that matter distribution (and hence spacetime structure) at cosmological scales is homogeneous and isotropic over a set of spatial hypersurfaces, with a definite curvature characteristic (either positive, negative or null). In other words, our universe conforms to a Friedman-Lemaitre-Robertson-Walker (FLRW) model. Empirically, the first and strongest evidence for maximal spatial symmetry comes from the observed near isotropy of the cosmic microwave background (CMB) [26, 27], which implies the universe was very homogeneous and isotropic at the time of last scattering. Formally, the existence of an isotropic radiation field with respect to an expanding geodesic congruence implies a Robertson-Walker metric [28–31], a result which is stable with respect to small deviations from the field isotropy [32], as is the empirical case. More recently, the CMB evidence is supplemented by large scale structure surveys [33, 34]. Theoretically, an FLRW model gives rise to a preferred time parameter (without an a priori preferred direction, however), whose flow is orthogonal to those hypersurfaces: that is cosmic time. Moreover, since those constraints prevent space from changing shape, there is only one spatial degree of freedom freedom left in the spacetime metric: the cosmic scale factor, which may vary with respect to cosmic time.

Although the above constraints do not prevent a static universe theoretically, it was shown by Robertson [35] that the only two static models allowed by them are the early models of Einstein and de Sitter (in a particular foliation). The first [17] is the so-called cylindrical world, which exhibits a perfect fine tuning of the cosmological term in order to match the matter density and achieve equilibrium, which is furthermore unstable. The second one [18, 19] is devoid of matter. Not only are these models very far from the actual situation, they are idealizations which are also far from any possible realistic situation. Thus any realistic FLRW model with matter will be one where the scale factor changes with cosmic time. According to Friedman's equations, this variation is monotonic, except in universes where the cosmological term is sufficiently small and the matter density sufficiently great. In such cases, there is a point of return, where the scale factor assumes a maximal value, decreasing in both time directions from that point, until vanishing. In all other cases, the scale factor is unbounded from above in one direction. In all cases, then, the scale factor shrinks to zero in at least one time direction. This means that cosmic time always has at least one boundary in realistic FLRW models. The present concordance model of our actual universe includes a cosmological term which leads to an a priori eternal spatial expansion in one direction of cosmic time. Hence a priori only in the opposite direction there is a boundary to cosmic time. That cosmological boundary is what will be referred to in this work as the Big Bang.

To the above stated cosmological facts it must be added the extra empirical fact that the temporal direction of cosmic expansion coincides with the psychological direction of time, and thus also with the direction of the general entropy increase. This means that the low entropy boundary condition which gives rise to the consistent macroscopic time asymmetry lies at the boundary of cosmic time, that is, at the Big Bang. All observations to date agree with the fact that entropy in every process, from primordial nucleosynthesis and stellar dynamics to planetary climate and biological evolution, has been increasing in the temporal direction of cosmic expansion ever since the Big Bang. Moreover, entropy and cosmic scale factor are the *only* two a priori independent quantities that vary in the universe as a whole, all other varying quantities, and hence any instances of temporal asymmetry, being linked to these two.

We have seen that the cosmological temporal asymmetry — the monotonic variation of the scale factor — is an expected result from the homogeneity and isotropy of space. However, space is not *exactly* homogeneous and isotropic. And that is precisely what allows the other temporal asymmetry, the so-called thermodynamic one, which is that given by the monotonic variation of entropy. To understand this, one must consider a crucial fact of thermodynamic systems. It is known that, for systems where pressure dominates, like the typical gas in a box, homogeneity and isotropy mean maximum entropy. However, this picture changes in the presence of long-distance forces. The attractive character of gravity prevents a homogeneous state of equilibrium. In fact, for gravitating systems, such as the universe, spatial homogeneity and isotropy mean a state of very *low* entropy [20–25, 36, 86, 105–107]. The tendency in such systems is that any irregularities will grow, increasing the clumping together of matter and destroying homogeneity and isotropy. This means that a universe described by a FLRW model has a relatively low entropy, which is the lower the closer it is to perfect spatial symmetry. At the last scattering surface, for instance, although matter was very hot, the temperature differences were very small, not greater than the order of 10^{-5}, which means a great spatial uniformity. Although there is entropy corresponding to the hot matter, the *gravitational* entropy is very low, way below the value it could have if that same matter would find itself confined into black holes, for instance. So we already knew that the entropic temporal asymmetry is due to the fact that the universe had a very low entropy around the time of the Big Bang. To this we must now add the fact that this low entropy state was a very smooth (in the sense of homogeneous and isotropic) one, which means that the kind of entropy which was constrained to a very low value was the gravitational one. This in the end the existence of a smooth Big Bang entails both the scale factor gradient (by enforcing an FLRW cosmology) *and* the entropy one (by giving a very low entropy boundary condition). This means that the smooth Big Bang is the key to all the temporal asymmetry in the universe. Our task is then to search for a theory which explains in a natural way why the universe has a smooth Big Bang. That I will call the *smooth Big Bang problem*. It will be the main focus of this thesis.

Chapter 2

Survey of approaches

The present chapter is devoted to the analysis of some of the most popular attempts to solve the smooth Big Bang problem and related issues: observer selection bias (also known as the weak anthropic principle) and cosmic inflation, plus the combination of both. A new alternative is also presented: Roger Penrose's conformal cyclic cosmology, which will be pursued further in the remainder of this thesis.

2.1 Observer selection bias

Generally in physics one would like to be able to get rid of free parameters, that is, finding a way to *deduce* from first principles the numbers which are currently taken as given, be they fundamental constants (like the Standard Model masses) or boundary conditions (such as most parameters in cosmology). Some of these parameters, or combinations thereof, are considered by some specialists to be fine-tuned for the existence of physicists [84], in the sense that anything resembling complexity would be precluded from existing were these parameters only slightly different. Now in statistics it is known that the selection of data for any scientific study is filtered by the precondition that there is some agent able and willing to collect that data. This effect is known as *observer selection bias*[1]. Forgetting to take it into account may lead to wrong probability estimates and some apparently unexplainable coincidences [37]. For this reason, many physicists have been trying to explain the values of arguably fine-tuned parameters by invoking it. The idea is that such parameters would not have fixed values, but would

[1]This kind of statistical bias is more commonly known among physicists as the "weak anthropic principle", a term coined by Brandon Carter [96], who defined it thus:

> "[W]hat we can expect to observe must be restricted by the conditions necessary for our presence as observers."

This is precisely the definition of observer selection bias. However, the name chosen by Carter is problematic. First, because the statement does not relate specifically to humans, but to any observers, here understood as any beings capable of inquiring about the world. Second, because it is not a principle, since a principle, strictly speaking, is something that should be assumed without proof; instead it is a tautology, which means it is always valid and should be taken into account in every scientific inquiry. So here I will stick to the nomenclature of the statistics literature: "observer selection bias".

vary in a specific manner across widely separated spacetime regions, or even more general settings, so that we would observe the values we do because in the regions where they differ the chance of there being any being asking about them is much reduced. This is a satisfactory answer, for example, for the question "why is the Earth hospitable to life?". Since there are vastly many planets in the observable universe, each with different properties, the existence of a few bio-friendly ones is bound to happen statistically, and we would only find ourselves living is those few. However, for any explanation of this type to be scientific valid, the existence of the ensemble of spacetime regions or other scenarios across which the parameters vary in the specified manner must be established, or at least falsifiable, as is required of any scientific theory. This requirement is satisfied by the above example of the Earth bio-friendliness, because our current understanding of the universe makes us believe there exist billions and billions of planets our there, which vary in size, temperature window and composition in such a way that a small but non-null chance of Earth-like planets forming exists. However, such requirement is frequently missed by most attempts to solve cosmological problems invoking observer selection bias, which prevents these attempts from being scientific [97]. A quick survey of some notable attempts is illuminating and teaches some important lessons.

Probably the first invocation of observer selection bias in physics is found in a paper by Boltzmann, where the idea is attributed to his assistant, Dr. Schuetz [95]. The goal in that case was exactly to explain why the entropy of the observed universe is so low, and was even lower in the past (though at that time it was not known that this was connected to homogeneity and isotropy, nor that the universe expanded). The solution proposed was to assume the universe arbitrarily extended in both space and time. This way, even if it is generally in equilibrium, one can make the probability to find a region of a certain size in any given state at any time as high as desired. Of course physicists would only find themselves in regions which are far from equilibrium. The discovery of cosmic expansion and the role of gravity made this world view impossible, but the theory could not live up even despite those advances. For if the low entropy of our world is the result of a fluctuation, a natural prediction is that such fluctuation should cover the minimum region necessary for our existence[2], or at least not extend beyond observations of the time. The actual case, on the contrary, is that it covers a vastly greater extension, and each new observation keeps stretching this extension. The theory was thus falsified [85, 87]. This simple example teaches the fundamental lesson that any attempt to solve the low entropy problem by means of observer selection bias must face the difficulty of explaining why there is such a huge reservoir of low entropy as we see in the standard Big Bang cosmology instead of just a "Boltzmann brain", that is, a much more probable small fluctuation. Other invocations of observer selection bias have been

[2]This is sometimes referred to as the "Boltzmann brain" paradox [86], since a typical asking being which emerged as a fluctuation should expect to be no more than a short-lived disembodied brain amidst a thermodynamic "soup".

notable in the history of cosmology. Not all of them dealt with the low entropy of the universe, but they also teach or reinforce important lessons which are very useful for our case.

In a letter to Nature in 1937 [98], and further in a 1938 paper [99], Dirac called attention to the closeness in the order of magnitude of some natural quantities, among them the age of the universe in atomic units and the inverse square of atomic masses in natural units[3]. To solve the coincidence, Dirac proposed that the parameters of atomic theory continually changed with time in order to keep the agreement. This became known as the “large numbers hypothesis”. Later, in 1961, Dicke [100] argued that physicists could normally only exist while there were stars in the universe, and showed that the location of such interval of cosmic history is a function of the atomic masses in the exact way observed by Dirac. He concluded that the coincidence was explained by already existing theories, taking observer selection bias into account, and thus Dirac’s theory was uncalled for. This satisfactory explanation started the era of anthropics in cosmology [101]. What is important to notice in this case is that the ensemble invoked by Dicke was nothing more than the sequence of epochs in the standard Big Bang universe, which was already established as a solid theory due to other observations, and not some hypothetic multiverse with properties tailored at will. For this reason the argument is sound and constitutes a valid scientific explanation.

The opposite case happened in 1973, when Collins and Hawking [38] studied the isotropy of the universe in light of a hypothesis by Misner [39, 40], called “chaotic cosmology”, according to which the universe would have come from a chaotic (homogeneous but not isotropic) state, and that the anisotropies would have been damped by dissipation processes. Collins and Hawking conclude that no such process could have produced an isotropic universe (which is not surprising in light of the considerations of the last chapter regarding the entropic behavior of gravitating systems). In fact, isotropic universes were found by them to be extremely rare. At the face of the impossibility to explain cosmic isotropy, the authors appealed to observer selection bias:

> “One possible way out of this difficulty is to suppose there is an infinite number of universes with all possible different initial conditions. Only those universes which are expanding just fast enough to avoid recollapsing would contain galaxies, and hence intelligent life. However, it seems that this subclass of universes which have just the escape velocity would in general approach isotropy. On this view, the fact that we observe the universe to be isotropic would be simply a reflection of our own existence.”

[3] Dirac expressly stated that the exact atomic units used (electron or proton mass, or mass-to-charge ratio, for instance) did not matter, since the results differ only by a few orders of magnitude, which in his view did not diminish the coincidence. The examples explicitly given by him are the age of the universe (estimated at 2×10^9 ages at the time) in units of e^2/mc^3 (7×10^{38}) and the ratio between the Coulomb and Newton forces between and electron and a proton (2.3×10^{39}).

This is a blatantly unscientific attitude: to *suppose* the existence of an infinite number of universes with all possible different initial conditions, without any independent empirical support, amounts to the same as no explanation at all. Simply *any* feature could be "explained" by such a move, that is, no glimpse of falsifiability exists for such hypothesis. Furthermore, the presence of observers (in the sense of beings who ask about the universe) is confounded by Collins and Hawking with the existence of galaxies. Although this might not be completely unreasonable, this move draws attention to another crucial fact: we do not know exactly which form observers might take and which conditions might allow their existence, and with what probabilities. This is a serious issue which haunts all arguments which use observer selection bias, so extreme caution must be taken when announcing results based on it.

A case similar to that of Collins and Hawking happened in 1989, where exactly the same mistake was made, though the technical details made its recognition much more subtle, leading to a wide acceptance of the fallacious argument even to the present date [97]. This case was the famous paper on the cosmological constant by Steven Weinberg [70]. Weinberg's paper dealt with three entities: Einstein's cosmological constant Λ, the cosmological constant equivalent of the energy density of the quantum field vacuum Λ_V, and the effective cosmological constant resulting from the combination of the two, $\Lambda_{\text{eff}} = \Lambda + \Lambda_V$. At the time, $|\Lambda_{\text{eff}}|$ already had an upper bound which was 118 orders of magnitude below the positive Λ_V calculated by particle physicists. If that calculation was correct, this meant that Einstein's Λ had to be negative and cancel Λ_V to at least 118 decimal places. Apart from the possibility that these two a priori independent quantities coincide so greatly in magnitude by sheer luck, two alternatives arise:

1. The quantities are actually not independent, but tied by some fundamental law as yet unknown.

2. There is actually no contribution to gravity from quantum vacuum, so that $\Lambda_{\text{eff}} = \Lambda$ and there is no spectacular fine tuning.

Weinberg analyses theories which might fulfill either 1 or 2, but also the possibility that both are wrong. In this case, one is forced to embrace the coincidence. And the reasonable consequence of this is that there is no reason for the coincidence to go even further than what we already see, exactly because it is a coincidence. So we should expect to find a nonzero Λ_{eff} somewhere near its upper bound. Weinberg then tries to show that the coincidence might be the effect of selection bias. For that sake, he invokes some multiverse models where Λ and/or Λ_V vary, stating that galaxies (and hence presumably life) could only form with low values of Λ_{eff}. However, this is only true if all other physical parameters remain unchanged throughout the multiverse [97]. Since there is no way to probe the existence of such a multiverse, let alone to know which and how parameters vary along it in case it exists, Weinberg's explanation cannot be held.

The lesson reinforced here is that an argument of selection bias can only survive if the ensemble on which it relies has independent empirical support. As already noted, this is the case with respect to Earth's bio-friendliness, since we know of the existence of other planets and their properties. However, the application of anthropics to cosmology requires the existence of a multiverse, which typically cannot be accessed empirically. If some multiverse theory makes predictions which can be falsified, then it might be pursued as a good scientific hypothesis. Otherwise there is nothing to be expected from it.

Many physicists hold that observer selection bias is to be used as a last resort: when everything else fails to explain a certain feature of our world, let us assume that there is much more world out there where such feature does not realize and then it is explained by observer selection bias. This attitude is completely unscientific. Observer selection bias is not a last resort, it is a factor which must *always* be taken into account *within* existing theories, *before* declaring that some feature is a coincidence and start searching for new physics. That was exactly the case of Dicke's argument, which dismissed a coincidence by taking into account the observer selection bias which follows directly from the already existing physical theories, thus rejecting the need for new theories such as the one by Dirac. On the other hand, when the established physical theories do not by themselves give rise to some observer selection bias which explains a given feature, it makes no sense to postulate a new theory which produces a multiverse whose existence is completely untestable just to explain such feature. Observer selection bias is a kind of restriction on what things can be observed which results from physical theories which are empirically established by specific, *independent* observations, not a card which can be pulled from the sleeve by postulating some new kind of multiverse tailored to agree with data every time some feature needs explaining. This does not mean that any multiverse theory is invalid. As an example, Lee Smolin's cosmic natural selection [97] is an instance of a falsifiable multiverse theory. However, most current such theories do not present any testable consequences.

Above I have just outlined the main difficulties faced by attempts to solve cosmological problems by means of observer selection bias, drawing from notable examples in the literature. In particular, I exposed one attempt to solve the problem focused on in this thesis, which is the the smooth Big Bang (low entropy) problem. As regards this problem in particular, the failed naïve Boltzmann-Schuetz hypothesis leaves an additional challenge to any other attempt in the form of the Boltzmann brain paradox. One way which has been proposed to evade this paradox is in a certain view regarding cosmic inflation. So let us talk about this paradigm now.

2.2 Cosmic inflation

In the previous chapter I exposed the view that the peculiar properties of the observed universe, in particular the high degree of spatial symmetry which allows for a monotonic entropy gradient, are due to the very special conditions realized at the boundary of cosmic time, i.e., the Big Bang. The inflationary paradigm [41–46, 56, 58] sustains the opposite view: that the properties of our universe are not surprising, but should be expected to arise dynamically, independently of the boundary conditions. This would be achieved through an exponential expansion phase driven by the potential of a scalar field. In particular, inflationists all agree that inflation explains why our universe is so homogeneous and isotropic, although the explanation itself varies. In spite of the confusion in the literature, after careful study it is seen that all arguments can be reduced to three logically independent views.

The first view is the one expressed in the original paper by Guth [41]. According to this view, spatial uniformity was not present from the beginning, but should have been *achieved* by causal processes, that is, thermalization. The smoothness problem thus reduces to the so-called "horizon problem", which is to explain how regions which were never in causal contact in the standard Big Bang model could have thermalized. As stated by Guth [41]:

> "The initial universe is assumed to be homogeneous, yet it consists of at least 10^{83} separate regions which are causally disconnected [...] Thus, one must assume that the forces which created these initial conditions were capable of violating causality".

Inflation allows those regions to have been in causal contact in the past, thus purportedly solving the horizon problem and hence the smoothness problem. Again citing Guth, now in a 2006 paper [58]:

> "The uniformity is created initially on microscopic scales, by normal thermal-equilibrium processes, and then inflation takes over and stretches the regions of uniformity to become large enough to encompass the observed universe."

However, a serious flaw makes this view utterly impossible: a gravitating system would never thermalize to a uniform state. As discussed in the previous chapter, it is the growing of inhomogeneities instead of their smoothing out that is entropically favored in the presence of gravity, so the universe could never have smoothed out from a previous inhomogeneous state. This means that however homogeneous and isotropic we observe the universe to be at any time, it could only have been even *more* homogeneous and isotropic in the past. The Big Bang itself must have been an extremely homogeneous and isotropic event. Hence *there is no horizon problem*, simply because no causal contact was needed to achieve smoothness. Smoothness was not *achieved*, but existing since the beginning. Still, most cosmology textbooks insist on the horizon problem [64–74], and many, even recent ones, still explicitly hold that inflation solves it by allowing previous causal contact between presently distant regions [75–83]. However, some inflationists admit the impossibility of such process, and recur to other means of explaining cosmic smoothness through inflation.

The second view admits that the universe could not have thermalized before inflation, and holds that the inflation period itself smoothed out the irregularities which existed before it. This would be possible because the inflaton potential is modeled to have negative pressure, acting as an effective cosmological constant and making gravity dominantly repulsive instead of attractive on large scales [22, 86]. However, there is no guarantee that irregularities shrink instead of growing during the exponential expansion phase [23]. Furthermore, the very onset of inflation already requires a high degree of spatial symmetry [47, 49–51, 53]. In fact, inflation requires even more special conditions than the normal Big Bang scenario [24, 36, 52, 86, 88, 89]. This is easy to see if we remember that time does not have an objective direction, since a contracting universe is not expected to "suck all the energy of the universe into the coherent (and extremely low entropy form) of a homogeneous rolling scalar field" [54, 88]. So again nothing is explained. The crucial fact which is missed in this view is that inflation, being based on time-symmetric fundamental physics, cannot reduce the entropy from whatever conditions existed before it. So it is simply impossible to explain the special properties of the observed universe as arising dynamically from natural boundary conditions (either by means of inflation, thermalization, Misner's chaotic cosmology mentioned in the previous section, or any other mechanism). The second law of thermodynamics, understood as the tendency of any system to evolve towards ever more probable states, enforces the boundary conditions to be very non-natural [25, 36]. This fact is accepted by some inflationists [43, 86], which leads to the third view on how inflation could explain the spatial uniformity of the universe.

The third view admits the above difficulty and accepts that the observed universe must have started off smooth. Instead of trying to show that this configuration is typical, it recognizes it as atypical, and turns to observer selection bias in a postulated big chaotic meta-universe to explain it. I have shown in the previous section that any attempt to explain the low entropy of the universe using observer selection bias must face the Boltzmann brain paradox: there is way more low entropy than needed for observers. The idea of the third view is to use inflation as an added feature to overcome this problem. Basically, inflation would turn a smooth region the size of ours into a matter of all or nothing: once a smooth region appeared as a fluctuation in the big chaotic meta-universe, it would readily inflate and become Hubble-sized like our observable universe [45–47, 86] (this is known as *chaotic inflation* [45]). So even if conditions to start inflation are rare, the exponential expansion generated by inflation would make post-inflation regions dominant among hospitable ones.

However, the truth is that regions like ours would be much more likely reached without inflation than with inflation [52, 53, 55, 88, 89]. This can again be easily seen if one rejects preconceptions about time having a preferred direction. Then it is clear that in such meta-universe, no matter which time convention is chosen, there should be as many inflating bubbles as deflating ones. Since a contracting smooth bubble is much more likely to shrink back into chaos without deflating, an expanding one must also be more likely to arise from it without inflating. Actually both situations are the same, the temporal direction being completely arbitrary. Furthermore, our conditions are still much more special than they needed to be even with inflation, i.e. even if the inflationary multiverse exists, we are not a typical habitable island. According to Paul Steinhardt, "the universe is flatter, smoother and more precisely scale-invariant than it had to be to support life. More typical islands, such as those younger than ours, are almost equally habitable yet much more numerous" [89]. Moreover, due to quantum fluctuations which randomly raise the value of the inflaton field (in the case of chaotic inflation), or due to the rate of exponential expansion being much greater than the rate of exponential decay (in the case of new inflation [42, 43]), if the inflaton field really exists, inflationists state that inflation keeps starting again (or going on) forever [48,49,58,61,62] (this is known as *eternal inflation* [58, 90]). As said by Guth, because of this, "anything that can happen will happen; in fact, it will happen an infinite number of times" [58]. This renders the theory completely unfalsifiable [60]. To try to remedy this, inflationists search for some kind of measure in order to calculate probabilities in the eternal inflation scenario, but the simplest and most natural ones disfavor the observed universe [55–58]. For example, weighting by volume gives young inflating regions the most weight, since at any time there are exponentially more inflating regions than before, and thus exponentially more pocket universes are being created (this is known as the "youngness problem") [55–58]. Instead of ruling out inflation, inflationists rule out the measures, and keep searching for a measure which yields the wanted data [55–58, 103]. Without some guiding principle other than just agreeing with observations, this reinforces the unfalsifiable character of the inflationary research project. It also should be noted that eternal inflation is time-asymmetric from the start, since inflation is supposed to happen only on one time direction. Then, apart from other problems which plague the paradigm[4], what becomes

[4]Inflation (even without a multiverse) is completely unconstrained, and can be adjusted to fit virtually any data [55, 59, 89] (even non-flat cosmologies [91–94]), by assuming an arbitrary number of scalar fields with arbitrarily complicated potentials. And it seems that the more natural settings (simpler potentials) are disfavored by present data [55, 56]. According to Steinhardt, inflationary potentials need orders of magnitude of fine tuning to reproduce the observed spectrum of density fluctuations [89]. Also "bad inflation" (inflation which leads to a universe very different from ours) is way more likely than "good inflation" (inflation which yields the properties we see), and also produces more structure, so there is not

clear is that it cannot explain the low entropy of the early universe.

2.3 Conformal cyclic cosmology

In section 1.3 it was argued that gravity contributes to entropy by clumping matter together. In particular, this creates event horizons which hide information from observers. Formulas for the entropy of black holes and cosmological horizons have already been obtained in the literature [108, 109]. It was also said that the Big Bang constitutes an extremely low entropy singularity, which is the opposite of typical black hole singularities. How is such fact to be expressed in terms of the geometrical properties of spacetime in the vicinity of the singularity? Penrose [21, 36, 110, 111] has proposed that the crucial element in such expression in the Weyl curvature tensor. Recall that the Riemann tensor in n dimensions can be decomposed as [112–115]

$$\begin{aligned} R_{abcd} = C_{abcd} + \frac{1}{n-2}\left(R_{ac}g_{bd} - R_{ad}g_{bc} - R_{bc}g_{ad} + R_{bd}g_{ac}\right) \\ -\frac{1}{(n-1)(n-2)}R\left(g_{ac}g_{bd} - g_{ad}g_{bc}\right) \end{aligned}$$

where $R_{bc} = R^a{}_{bac}$ is the Ricci tensor, $R = g^{ab}R_{ab}$ is the Ricci scalar and C_{abcd} is the Weyl tensor, which is basically what is left from the Riemann tensor after all possible contractions have been subtracted, that is, it is the traceless part of the Riemann tensor. The Ricci tensor gives the absolute volume scaling of spacetime, while the Weyl tensor gives the deformation, that is, the change in angles with respect to a spacetime without curvature. If C_{abcd} is zero (in four dimensions or more) then the metric is conformally flat and spacetime is in a certain sense uniform. That is exactly the description of the smooth Big Bang which is needed for the second law of thermodynamics to be valid. Thus Penrose has hypothesized that the Weyl tensor must vanish, or at least be bounded, at the Big Bang singularity. This is the original formulation of the well known *Weyl curvature hypothesis* (WCH) [21, 36, 110].

More recently, Paul Tod has found a new way to formulate the WCH [117], based on the theory of conformal gauge singularities. Each equivalence class of all the metrics which are conformally equivalent to a given metric on a given manifold is called a *conformal class*. The choice of a specific metric within a given conformal class is called a *conformal gauge*. It may happen that some conformal gauges in a given conformal class are singular, while others are not. Such singularities are called *conformal gauge singularities*, or *isotropic singularities*. Another way to put this is that conformal gauge singularities are singularities which can be removed by conformally rescaling the metric. Such procedure produces a smooth spacelike hypersurface in the new rescaled spacetime

---

even a selection bias argument to favor good inflation [89].

corresponding to the singularity in the original one. This also allows geodesics to be extended beyond the singularity, which is now just a smooth spacelike hypersurface, and a Cauchy problem can thus be posed with initial data given on this hypersurface. Tod has shown that a singularity with a finite Weyl tensor is a conformal gauge singularity, which means that spacetime can be extended beyond it in a smooth way after a conformal rescaling [117]. This means that the low gravitational entropy of the Big Bang can be expressed as the condition that it constitutes a conformal boundary to spacetime. This would mean that spacetime could be continued beyond the Big Bang in a smooth fashion, at least mathematically [111].

Penrose uses the above fact as the first step to his theory of *conformal cyclic cosmology* (CCC) [90, 111], according to which the final stage of the accelerated cosmic expansion is connected smoothly to a new Big Bang, and so on in a cyclic fashion, so that our own Big Bang is connected to a previous expanding phase to the conventional past. Each such cosmic cycle is called an *aeon* by Penrose. This connection between aeons would be possible mathematically due to Tod's theorem, and physically because the final stages of the cosmic expansion would resemble the Big Bang singularity in the sense that the clumping character of gravity has ceased to be dominant and matter has become homogeneously spread across the universe again after the final black holes have evaporated. There is, however, a huge difference between the final expanding universe scenario and the initial hot Big Bang epoch in terms of metric scale. Nevertheless, Penrose has presented arguments for believing that metric scale in not a physically significant concept in either situation.

Penrose [90, 111] has stressed the important fact that the metric character of spacetime is only perceptible in the presence of massive particles. Massless particles such as the photon can only sense the conformal spacetime structure, without any notion of distance. He notes that such a situation is physically realized at some stages of cosmic evolution. First, the early universe was extremely hot, which means the kinetic energy of fundamental particles greatly overwhelmed their mass term. This means the particles behaved in an ultrarelativistic fashion, their masses being essentially irrelevant. Thus the metric character of spacetime is the less present the closer one gets to the Big Bang. In effect, before symmetry-breaking, all fields should be exactly massless, thus only the conformal structure of spacetime should be present in the early stages of cosmic evolution. Second, the late universe is likely to be composed mainly by massless or effectively massless particles, since black holes will evaporate leaving only photons and gravitational waves. Thus in the far future the metric structure of spacetime is also expected to become irrelevant. Future infinity becomes a a region like any other for such massless particles, which can be crossed to some hypothetical extension without problems.

Summing all these information up, one has that both the past and future cosmic spacetime boundaries might allow for conformal extensions, which furthermore should be populated by massless matter, whose activity might be extended across such boundaries. Thus one might hypothesize that the universe works in a cyclic fashion, the far end of the expansion in one cycle connecting to the Big Bang of the next one. That is the proposal of conformal cyclic cosmology.

CCC is thus based on two theoretical observations. First, that the low entropy of the Big Bang might be expressed as the condition that it can be described as a spacelike conformal boundary beyond which spacetime could be extended. Second, that the late stages of our universe might approach such a boundary if massive particles are absent, and that immediately after the Big Bang the same is true for the kind of matter present.

Appealing as it might seem, CCC raises its own issues. Probably the most urgent of these is the doubt about whether in the far future there will be any massive particles left. Electric charge must be conserved, so electrons and positrons could not decay directly into photons. Of course they could anihilate, but one would not expect all electrons and positrons to disapeear this way. Another possibility is that de Sitter expansion at some point reaches, in some specific region, a complete horizon empty of any massive particles, and that such region might then connect with a Big Bang conformal boundary. However, as the cosmological constant gradually becomes dominant and spacetime asymptotically approaches de Sitter space, its symmetry group deviates ever more from the Poincaré group. Since rest mass is an exact Casimir operator of the Poincaré group, but *not* of the de Sitter group, it does not need to be conserved. In fact, what is asymptotically approached as the metric loses significance is the conformal cone spacetime, which represents in a satisfactory way both the final and initial stages of the universe, satisfying the necessary properties to consist of the boundary connecting two aeons.

In the next chapters the mathematics necessary to combine the ideas of conformal cyclic cosmology with that of an invariant kinematic length scale as presented in section 1.1 will be developed.

Chapter 3

Differential geometry and physics

3.1 Geometries and connections

Apart from spacetime intervals, physical quantities in general do not "belong" to spacetime. Observables which are described by such objects as scalars, vectors, tensors or spinors are elements of some space which is somehow "attached" to each point of spacetime. Such construction in mathematically formalized through the notion of fiber bundles. These quantities may have both "internal" degrees of freedom and others which are related to the spacetime structure itself. Even in classical mechanics, for example, simple observables such as the velocity of a particle do not "live" in the Euclidean space of positions, but in its tangent space. Since velocities are rates of displacement with respect to time, they can be understood as differences between points, which in the case of Euclidean space can be defined and correspond to Euclidean vectors. Hence in this particular case the tangent space is isomorphic to the space of positions. Such isomorphism however is limited in a sense: while the space of velocities is a *vector* Euclidean space, i.e., there exists a null velocity vector, the space of positions is a more general *affine* Euclidean space, where no preferred origin is specified. So in Euclidean geometry, the tangent space is isomorphic to the space of positions, also called base space, as an affine space, but the tangent space also has an additional vector structure (a specified origin). One particular feature of this construction is that vectors located at tangent spaces to different points can be compared directly. Since vectors like the velocity are given by spacetime displacements, a coordinate system in the base space assigns naturally a basis in the tangent space at each point. In a Cartesian coordinate system, the components of tangent vectors are enough to compare vectors at different points.

Riemannian geometry generalizes Euclidean geometry in that the space of positions is now allowed to have different shapes, that is, Euclid's fifth postulate is dropped. In particular, it does not need to be an affine space anymore. In fact it can have any shape which is smooth. The tangent space on the other hand must still be a vector space.

However, because the base space does not have an affine structure in general, its points cannot be subtracted to yield vectors. Vectors can still be defined locally as equivalence classes of directions of curves. This can be done rigorously by associating to each curve at a point a natural derivative operator acting on scalar fields. However we can not compare vectors at tangent spaces to different points anymore. To be able to do so one must specify a rule, which is called a *connection*. This tells how one transports vectors in a way so that they remain parallel to themselves. Unlike in Euclidean space, however, the result of such parallel transport depends in general on the path chosen. It may even happen that a vector transported in a parallel way along a loop comes back as a different vector from the original one. Such difference is measured by a quantity called *curvature*.

However, Riemannian geometry, in Riemann's original formulation (which was the one studied by Ricci-Curbastro and Levi Civita, later adopted by Einstein, and still the most popular among physicist in the study of general relativity), is in a certain sense limited, especially because of two of its features. First, all its statements refer to a metric. It is really an essentially metric geometry. This is clearly seen when we observe the Riemannian definition of curvature. In abstract differential geometry, curvature is a property of a connection, and it tells at each point the difference between the results one obtains when parallel transporting a vector (or any more general geometric object for that matter) around an infinitesimal loop. For Riemann, on the other hand, curvature at a point is the amount to which the volume of an infinitesimal figure at that point differs from what it should be in Euclidean geometry. This is a metric property — without a metric, it does not make sense, and hence neither does Riemannian geometry itself. The second limitation of traditional Riemannian geometry is that it is formulated entirely in terms of coordinates. This does not mean that the results of Riemannian geometry are coordinate-dependent. It just means that never did Riemann (nor Einstein, for that matter) formulate his equations in an abstract, coordinate-free fashion. Of course coordinates are indispensable as calculation tools. But today we know that they are not indispensable for building geometries. Furthermore, we know that we can learn a lot about essential geometric features if we dispense them. And in order to effectively do so it is really necessary to understand Riemannian geometry inside the more general and abstract framework of fiber bundles.

As said above, Riemannian geometry is built upon Euclidean geometry, and locally always reduces to it. This means that the *model* for Riemannian geometry is the same as its tangent space: just old plain Euclidean space. Thus the most natural way in which one could think of generalizing Riemannian geometry is arguably by asking the question "Why not generalize the model space?". The answer is that this is indeed feasible. A first natural generalization would be to let the model space be an affine space instead of a vector space. Geometrically, this would mean that a tangent vector transported around a loop in a parallel manner might fail not only to keep its original direction, but also its original starting point in the affine tangent space. Such fail is called *torsion*, and is the companion concept to curvature. While curvature represents the rotation suffered by the vector after the parallel loop journey, torsion represents the translation of its origin in the affine model space after the same journey.

But we can go even further and ask "Why restrict ourselves to affine model spaces? Why not more general spaces?". And then we arrive at Cartan geometries. These are geometries where the model space is any homogeneous space, including vectors spaces, affine spaces, projective spaces, spherical spaces, hyperbolic spaces and so on. Cartan's idea of parallel transport along a curve corresponds to a geometric notion which is very simple and intuitively easy to visualize: that of the model space rolling without slipping or twisting along the curve on the base manifold. In this process, the point of contact of the two surfaces changes in both surfaces. If after rolling around a loop on the base space the point of contact in the tangent space differs from what it was in the beginning of the journey, then there is torsion at that point of the base.

Cartan geometries are still just particular examples of fiber bundles with connections, called Ehresmann connections. The specific kind of Ehresmann connection on a Cartan geometry is a Cartan connection, while the specific kind of Cartan connection in Riemannian geometry is the affine connection, and the one chosen in GR is a still more special one, called the Levi-Civita connection. This one is considered a "natural" connection in a Riemannian manifold, which can be explained in at least two ways. First, it is the connection for which self-parallel curves (i.e., geodesics of the connection) extremize length (i.e., are geodesics of the metric). Second, it is the torsion-free metric connection. Saying the connection is metric means that the metric is parallel transport by it along any path. In other words, the covariant derivative of the metric always vanishes. Since the metric is the inner product in tangent space, and since parallel transportation defines an isomorphism between tangent spaces to distant points, a metric connection preserves the inner products of vectors when they are parallel transported. The torsion-free condition, on the other hand, means that vectors which are parallel transported around an infinitesimal loop come back to the same origin in the tangent space of the starting point.

3.2 Fiber bundles

A fiber bundle generalizes the Cartesian product of topological spaces, that is, it is the general formal concept behind the idea of a space B to every point of which a copy of another space F has been "attached". A connection on a fiber bundle by its turn represents the specific geometric way in which the many copies of F have been attached to each point of B, that is, it gives a means to compare objects belonging to copies of F "glued" to different points on B.

If we want to attach a copy of F at every point of B, we will get a "bigger" space E. This is a "synthetic" view: we start by the parts and build the whole. The problem with that construction in this case is that in the end we do not have a nice way to tell one copy of F from the other. Mathematicians then found out that the best way to define a fiber bundle is the converse, "analytical" way, in which one starts with the whole and then "breaks it down" to the parts. So we start with the "big" space E and define a *projection* of it onto B, which means to tell, for each point in B, what collection of points in E is "attached" to it. We then consider such construction to be a fiber bundle if this collection is in fact a copy of F for every point in B.

So what we need is a "big" space E which locally "looks like" $B \times F$. For that we need a map from E to B which covers all B and such that the inverse image of each point is precisely F. In particular we want to focus on the smooth category. Let E, B and F be smooth manifolds and let π be a smooth surjection from E to B. The quadruple (E, B, π, F) is a **smooth fiber bundle** iff every point in B has an open neighborhood U for which a diffeomorphism ϕ exists satisfying

$$\mathrm{proj}_1 \circ \phi = \pi,$$

where proj_1 is the projection onto the first factor. This corresponds to the following commutative diagram.

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\ \phi\ } & U \times F \\ {\scriptstyle \pi}\downarrow & \swarrow {\scriptstyle \mathrm{proj}_1} & \\ U & & \end{array}$$

E is called **total space** of the bundle, B is the **base space**, π is the **projection** and F is the **typical fiber** or **abstract fiber**. When no confusion can arise, we will denote the bundle simply by E. Each pair (U, ϕ) is called a **chart**. It follows that $\mathrm{proj}_1^{-1}(b)$ is isomorphic to $\pi^{-1}(b)$ for every b in B. But $\mathrm{proj}_1^{-1}(b)$ is the set of all pairs (b, f) where $f \in F$, and hence is isomorphic to F. Thus $\pi^{-1}(b)$ is isomorphic to F for every b in B, and is called the **fiber** over b. Clearly, if (V, ψ) is another chart such that V intersects U,

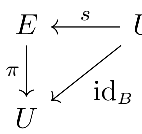
$$\mathrm{proj}_1 \circ \psi \circ \phi^{-1} = \mathrm{proj}_1.$$

This means that the first coordinate of any element of the total space is chart-independent. There is a priori no such restriction for the second entry. A G-bundle will be a special structure defined by introducing exactly that kind of restriction, in such a way that different charts are related by the action of a group on the second coordinate.

A **local section** of the bundle at a given open neighborhood U is a smooth right inverse of π at U, that is, a smooth map $s : U \to E$ such that

$$\pi \circ s = \mathrm{id}_B.$$

This corresponds to the following commutative diagram.



A section formalizes the notion of a field (like a vector field) over the base space B as an F-valued function on B.

Fiber bundles become more interesting when they posses groups of symmetries. These are encoded in group actions, which must fulfill some properties. A **left (or right) action** of a group G on a set X is a map from $G \times X$ (or $X \times G$) to X which is compatible with both the group composition and group identity, that is, $g(hx) = (gh)x$ and $ex = x$ (or $(xg)h = x(gh)$ and $xe = x$). The set X endowed with such action is called a **left/right G-space**. A given action is said to be **transitive** if given any two elements x and y of X there is always some g in G which brings x into y. It is said to be **effective**, if different transformations always have different results (for at least one element of X). A stronger requirement than effectiveness is that different transformations have different effects on *every* element of X. Such an action is called **free**.

The notion of group action gives rise to the concept of torsors. Just like an affine space over a vector space V is a set whose points can be "joined" by vectors in V, or more informally an instance of V without a specified origin, a torsor over a group G is a set whose points can be "joined" by elements of G, or more informally an instance of G without a specified identity. Since every vector space is an abelian group under vector addition and the origin is just the additive identity, every affine space is a torsor. Formally a **left/right G-torsor**, is a set X endowed with a left/right regular action of G on it. This amounts to saying that, for every x and y in X, there is always a g in G such that $y = gx$, and this g is unique. More generally, X can be any structure, provided the action of G always induces automorphisms on it. Hence affine spaces are torsors in the category of vector spaces. Klein geometries, as we will see shortly, are torsors in the category of smooth manifolds.

Now let G be a Lie group. Then the tuple (E, B, π, F, G) is a **smooth G-bundle** iff there is a smooth left action l of G on F by diffeomorphisms such that for every pair of charts (U, ϕ) and (V, ψ) of the bundle (E, B, π, F) a smooth map $h : U \cap V \to G$ exists satisfying

$$\mathrm{proj}_2 \circ \psi \circ \phi^{-1} = l \circ (h \times \mathrm{id}_F).$$

This corresponds to the following commutative diagram.

$$\begin{array}{ccccc} (U \cap V) \times F & \xrightarrow{\phi^{-1}} & \pi^{-1}(U \cap V) & \xrightarrow{\psi} & (U \cap V) \times F \\ & \searrow^{h \times \mathrm{id}_F} & & & \downarrow \mathrm{proj}_2 \\ & & G \times F & \xrightarrow{\quad l \quad} & F \end{array}$$

This means that a change of chart amounts to a group transformation on the fiber, that is,

$$\psi \circ \phi^{-1}(u, f) = (u, h(u)f).$$

G is called the bundle's **structure group**. Each map h is called a **transition function**. If we denote by h_{ji} the transition function corresponding to the change of chart $\phi_j \circ \phi_i^{-1}$, then for any other overlapping chart ϕ_k we have that

$$\begin{aligned} (u, h_{ki}(u)f) &= \phi_k \circ \phi_i^{-1}(u, f) \\ &= \phi_k \circ \phi_j^{-1} \circ \phi_j \circ \phi_i^{-1}(u, f) \\ &= (u, h_{kj}(u)h_{ji}(u)f), \end{aligned}$$

hence

$$h_{ki} = h_{kj}h_{ji}.$$

By the *fiber bundle construction theorem* [8, 116], if one is given a smooth manifold B with a covering $\{U_i\}$, another smooth manifold F with a left action of a Lie group G on it and a set of smooth maps satisfying (3.1), one can construct a smooth fiber bundle E with B as base, F as fiber, G as structure group and the h's as transition functions. In this construction F can be replaced by any left G-space F'. The resulting bundle, call it E', is then said to be **associated** to E.

Moreover, (E, B, π, F', G) is a **smooth principal G-bundle** iff (E, B, π, F') is a smooth fiber bundle and there is a right action r of G on E which preserves the fibers, that is,

$$\pi \circ r = \pi \circ \mathrm{proj}_1,$$

and also acts freely and transitively on each of them (this means that F' is a G-torsor).

The fiber preserving requirement corresponds to the following commutative diagram.

$$\begin{array}{ccc} E \times G & \xrightarrow{r} & E \\ {\scriptstyle \mathrm{proj}_1}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ E & \xrightarrow{\pi} & B \end{array}$$

Suppose we have a smooth G-bundle E and construct an associated bundle E' by replacing the fiber F by the group G itself. In this case the new fiber is endowed with not only a left G-action but also a right G-action. This right action of G on itself induces a right G-action on the bundle E', turning it into a principal bundle. This is called the **associated principal bundle** to the bundle E. For example, given a vector bundle P, i.e. a bundle whose fiber is a vector space V, the structure group is $GL(V)$, and the bundle of all ordered bases for all the fibers is the associated principal bundle to it, called its **frame bundle** $F(P)$, on which $GL(V)$ acts freely and transitively by change of basis. Unlike the original bundle, the associated principal bundle lacks a preferred origin at each fiber. In particular, the tangent bundle of a n-dimensional real smooth manifold is a smooth vector bundle with structure group $GL(n, \mathbb{R})$, and the principal bundle associated to it is the frame bundle $F(M)$.

Another important concept is that of **reduction of structure group**. Given a G-bundle (E, B, π, F, G), we say that the bundle admits a reduction of its structure group to another group H if it is isomorphic to a H-bundle with the same base and fiber. The reduction of the structure group in a fiber bundle is related to the introduction of some additional structure in the bundle. For example, a metric reduces the structure group to the orthogonal group, while an orientation reduces it to the special group, and so on.

3.3 Induced maps and Lie derivative

If M and N are two smooth manifolds, a smooth bijection ϕ from M to N induces the *pushforward* of objects from M to N and the *pullback* of objects from N to M. For any smooth scalar field g on N, its **pullback** by ϕ, denoted $\phi^* g$ is defined as simply the composition $g \circ \phi$. If γ' is a tangent vector field to a path γ on M, the tangent vector field to the corresponding path $\phi \circ \gamma$ on N is called its **pushforward** by ϕ, denoted $\phi_* \gamma'$. For a general vector field u acting on scalar fields M, then,

$$\phi_* u(g)|_q = u(\phi^* g)|_{\phi^{-1}(q)}$$

for any point q and smooth scalar field g on N, or, more abstractly,

$$\phi_* u(g) = u(\phi^* g) \circ \phi^{-1}. \tag{3.1}$$

If x and y are coordinates in M and N respectively, and $u^{\nu}_{[x]}$ and $(\phi_* u)^{\mu}_{[y]}$ are components of u and $\phi_* u$ on the respective coordinate basis, then we can define

$$\phi_{[xy]} = y \circ \phi \circ x^{-1}$$

and rewrite (3.1) as

$$\begin{aligned}
(\phi_* u)^{\mu}_{[y]}\, \partial_{\mu}(g \circ y^{-1}) \circ y &= u^{\nu}_{[x]} \partial_{\nu}(g \circ \phi \circ x^{-1}) \circ x \circ \phi^{-1} \\
&= u^{\nu}_{[x]} \partial_{\nu}(g \circ y^{-1} \circ \phi_{[xy]}) \circ x \circ \phi^{-1} \\
&= u^{\nu}_{[x]} \partial_{\nu} \phi^{\mu}_{[xy]} \partial_{\mu}(g \circ y^{-1}) \circ \phi_{[xy]} \circ x \circ \phi^{-1} \\
&= u^{\nu}_{[x]} \partial_{\nu} \phi^{\mu}_{[xy]} \partial_{\mu}(g \circ y^{-1}) \circ y,
\end{aligned}$$

so that

$$(\phi_* u)^{\mu}_{[y]} = u^{\nu}_{[x]} \partial_{\nu} \phi^{\mu}_{[xy]}.$$

Since $\partial_{\nu} \phi^{\mu}_{[xy]}$ is the Jacobian matrix of the transformation ϕ in the coordinates x and y, ϕ_* can be understood as the derivative of ϕ.

If b is a dual vector field on N, its pullback by ϕ is the dual vector field $\phi^* b$ on M satisfying

$$\langle \phi^* b, u \rangle = \langle b, \phi_* u \rangle \circ \phi$$

for any vector field u on M.

If the inverse of ϕ is also smooth (i.e. ϕ is a diffeomorphism) we can also define the pushforward of a scalar field f on M as

$$\phi_* f = \left(\phi^{-1}\right)^* f = f \circ \phi^{-1}.$$

Similarly, we can define the pullback of a vector field v on N as

$$\phi^* v = \left(\phi^{-1}\right)_* v$$

and the pushforward for a dual vector field a on M as

$$\phi_* a = (\phi^{-1})^* a,$$

which means that

$$\phi^* v(f) = v(\phi_* f) \circ \phi$$

and

$$\phi_* a(v) = a(\phi^* v) \circ \phi^{-1}$$

respectively. Then for any tensor field T on M we can define its pushforward by ϕ as

$$\phi_* T(v_1, v_2, ..., a_1, a_2, ...) = T(\phi^* v_1, \phi^* v_2, ..., \phi^* a_1, \phi^* a_2, ...) \circ \phi^{-1}$$

and the corresponding pullback as

$$\phi^* T(v_1, v_2, ..., a_1, a_2, ...) = T(\phi_* v_1, \phi_* v_2, ..., \phi_* a_1, \phi_* a_2, ...) \circ \phi.$$

The above general pullback operation allows us to compare the value of a tensor field T at a point p with its value at a different point q provided there is a diffeomorphism ϕ from M to itself such that $\phi(p) = q$. This is done by pulling T back through ϕ, that is, we can calculate

$$(\phi^* T - T)(p).$$

The pullback is necessary since $T(p)$ and $T(q)$ belong to different spaces and hence cannot be compared directly, whilst $\phi^* T$ effectively "pulls" the tensor $T(q)$ "back" to the point p so that it can be operated together with the tensor $T(p)$, that is, $\phi^* T(p)$ just "is" the tensor $T(q)$ "transported" to the point p.

Furthermore, any vector field u on M generates a one-parameter family of diffeomorphisms on M given by

$$\gamma_t^{[u]}(p) = \gamma\left(\gamma^{-1}(p) + t\right),$$

where γ is the flow of u. This allows us to compare the value of a tensor field at a point with its value at another point infinitesimally separated from the first in the direction given by the vector field in question:

$$\lim_{t \to 0} \left(\gamma_t^{[u]}\right)^* T - T.$$

Dividing by the parameter separation, this gives us a directional derivative operator for all smooth tensor fields on M:

$$\pounds_u T = \lim_{t \to 0} \frac{\left(\gamma_t^{[u]}\right)^* T - T}{t}. \tag{3.2}$$

This is called the **Lie derivative** of T in the direction of the tangent vector field u. It is a fundamental concept since it is a kind of derivative of tensor fields that can be defined on the manifold without any additional structure. It will be important in section 3.5.

3.4 Covariant derivatives and connections

Let (E, B, π, F) be a smooth fiber bundle and ψ a local section of it on an open neighborhood U. How can one compare the value of ψ at different points on U? In particular, how can one define a notion of differentiability for local sections of E? We have seen that the Lie derivative solves that problem for tensor fields on smooth manifolds, but what of general smooth fiber bundles? In that case, some additional structure is in general required. Such structure is called a **connection**. Formally, a connection on a smooth fiber bundle is a prescription to smoothly associate to each path on the base manifold a morphism between the fibers of the initial and final points of the path in such a way that the symmetry structure of the bundle is preserved (for instance, the G-structure in the case of a principal bundle, for which the connection is called a **principal connection**). Furthermore, it is required that, for a composition of paths, the connection the action of the connection results in the same transformation as the composition of its actions on the separate parts.

The above definition is better illustrated with an example. For simplicity, let us suppose that the fiber F is a (n-dimensional) vector space. Then let

$$\xi = (\xi_1 \ \xi_2 \ ... \ \xi_n)$$

be a row matrix composed of vectors who form a basis for F. Let furthermore

$$\psi_\xi = \begin{pmatrix} \psi_\xi^1 \\ \psi_\xi^2 \\ \vdots \\ \psi_\xi^n \end{pmatrix}$$

be a column matrix made of the components of ψ in the basis ξ, so that

$$\psi = \xi \psi_\xi.$$

If ζ is another basis, we can write

$$\xi = \zeta \xi_\zeta$$

and

$$\zeta = \xi \zeta_\xi,$$

where ξ_ζ is the square matrix whose ij entry is the i-th component of the vector ξ_j in the basis ζ, and similarly for ζ_ξ. Clearly, both matrices are inverse to each other.

For the section ψ we then have that

$$\begin{aligned}\psi &= \xi\psi_\xi \\ &= \zeta\xi_\zeta\psi_\xi \\ &= \zeta\psi_\zeta,\end{aligned}$$

so that the section components transform according to

$$\psi_\zeta = \xi_\zeta\psi_\xi,$$

leaving the geometric object ψ invariant.

Now we want to define an operator D which acts on local sections such as ψ and gives their derivative. First of all, which kind of object should such a derivative be? We know that the operator D must take as input not only a section but also a direction on the base manifold B. So what we will have is an operator D_u for each tangent vector field u on B. In the present case, where F is a vector space, D_u must clearly be linear. The tangent vector field u itself acts on scalar fields on B as a directional derivative. It can act on the components of sections of E, but this cannot be used as a good derivative operator for sections since its result is basis-dependent. That is, if we defined

$$D_u\psi = \xi\partial_u\psi_\xi$$

then we would have, after changing to another basis ζ,

$$\begin{aligned}D_u\psi &= \zeta\partial_u\psi_\zeta \\ &= \xi\zeta_\xi\partial_u(\xi_\zeta\psi_\xi) \\ &= \xi\zeta_\xi(\partial_u\xi_\zeta\psi_\xi + \xi_\zeta\partial_u\psi_\xi) \\ &= \xi\zeta_\xi\partial_u\xi_\zeta\psi_\xi + \xi\partial_u\psi_\xi \\ &= \xi\zeta_\xi\partial_u\xi_\zeta\psi_\xi + D_u\psi,\end{aligned} \tag{3.3}$$

which is absurd except if $\partial_u\xi_\zeta = 0$. The bottom line is that the variation of the section, if it is dissected into basis vectors and components, lies not only at the components but also at the basis vectors themselves. Both must be taken into account in order that the derivative operator acts in a basis-independent way. The correct way to obtain such an invariant derivative notion is to enforce it to satisfy the Leibniz rule

$$D_u(f\phi) = \partial_u f\phi + fD_u\phi,$$

where f is a scalar field on B and ϕ is a section of E. We thus get

$$\begin{aligned} D_u\psi &= D_u(\xi\psi_\xi) \\ &= \xi\partial_u\psi_\xi + D_u\xi\psi_\xi \\ &= \xi\left[\partial_u\psi_\xi + (D_u\xi)_\xi\psi_\xi\right] \\ &= \xi(\partial_u + A_u^\xi)\psi_\xi, \end{aligned}$$

where

$$A_u^\xi \equiv (D_u\xi)_\xi$$

is a square matrix whose entries are the components of the derivatives of each ξ_k. If we define the operator

$$\nabla_u^\xi = \partial_u + A_u^\xi,$$

then we have

$$D_u\psi = \xi\nabla_u^\xi\psi_\xi,$$

which is the expression for the derivative of a section in a given basis. This object is a true section itself. We can check that by finding the transformation law for A_u:

$$\begin{aligned} A_u^\zeta &= (D_u\zeta)_\zeta \qquad (3.4) \\ &= [D_u(\xi\zeta_\xi)]_\zeta \\ &= \xi_\zeta[D_u(\xi\zeta_\xi)]_\xi \\ &= \xi_\zeta[D_u\xi\zeta_\xi + \xi\partial_u\zeta_\xi]_\xi \\ &= \xi_\zeta[A_u^\xi\zeta_\xi + \partial_u\zeta_\xi]. \end{aligned}$$

A_u is clearly not an invariant object. However, this is exactly what is needed to compensate for the extra term in (3.3). We then have that

$$\begin{aligned} D_u\psi &= \zeta\nabla_u^\zeta\psi_\zeta \\ &= \xi\zeta_\xi(\partial_u + A_u^\zeta)\xi_\zeta\psi_\xi \\ &= \xi\zeta_\xi(\partial_u\xi_\zeta\psi_\xi + \xi_\zeta\partial_u\psi_\xi + \xi_\zeta A_u^\xi\zeta_\xi\xi_\zeta\psi_\xi + \xi_\zeta\partial_u\zeta_\xi\xi_\zeta\psi_\xi) \\ &= \xi\zeta_\xi\partial_u\xi_\zeta\psi_\xi + \xi\partial_u\psi_\xi + \xi A_u^\xi\psi_\xi + \xi\partial_u\zeta_\xi\xi_\zeta\psi_\xi \\ &= \xi\partial_u\psi_\xi + \xi A_u^\xi\psi_\xi \\ &= \xi\nabla_u^\xi\psi_\xi, \end{aligned}$$

where it is proven that $D_u\psi$ is invariant under changes of basis in F. The derivative operator D_u is sometimes called a *covariant derivative* or *connection* on E. Notice however that it is not unique. In fact, there are infinite ways to define such derivative. One way to see this is noticing that the object A^ξ_u is not constrained. Each different choice for this matrix gives a different connection on E. Each connection by its turn defines a notion of *parallel transport* on the bundle. A section ψ is said to be parallelly transported along a curve on B iff

$$D_u\psi = 0,$$

where u is the tangent vector field to the curve.

Now we know that A_u is not invariant, so that it may be non-zero in a given basis and still vanish in another basis. This means that the existence of this term is dependent not only on the intrinsic geometric properties of the bundle, but also on the choice of basis. Now we want to find an object which represents the non-triviality of the bundle but depends solely on intrinsic properties. In other words, we want to define and object such that if it vanishes *everywhere* in some basis it does so in *all* basis, and this means that A_u vanishes *everywhere* in *some* basis. For that, recall the transformation law for A_u (3.4):

$$A^\zeta_u = \xi_\zeta[A^\xi_u\zeta_\xi + \partial_u\zeta_\xi].$$

Let ζ be the basis in which A_u vanishes. Then we have

$$A^\xi_u\zeta_\xi = -\partial_u\zeta_\xi. \tag{3.5}$$

Now, using the identity

$$\partial_u\partial_v\zeta_\xi - \partial_v\partial_u\zeta_\xi = 0$$

and replacing (3.5), we get

$$\begin{aligned} 0 &= \partial_u(A^\xi_v\zeta_\xi) - \partial_v(A^\xi_u\zeta_\xi) \\ &= \partial_u A^\xi_v\zeta_\xi + A^\xi_v\partial_u\zeta_\xi - \partial_v A^\xi_u\zeta_\xi - A^\xi_u\partial_v\zeta_\xi \\ &= \partial_u A^\xi_v\zeta_\xi - A^\xi_v A^\xi_u\zeta_\xi - \partial_v A^\xi_u\zeta_\xi + A^\xi_u A^\xi_v\zeta_\xi \end{aligned}$$

Thus the object

$$F^\xi_{uv} \equiv \partial_u A^\xi_v - \partial_v A^\xi_u + A^\xi_u A^\xi_v - A^\xi_v A^\xi_u = \nabla^\xi_{[u}A^\xi_{v]}$$

vanishes everywhere in any basis when A_u vanishes everywhere in some basis, which means that the bundle is trivial.

The object A is a 1-form on the exterior algebra of the base manifold with values in the Lie algebra of the structure group of the bundle (in this simple case of a vector bundle, it is $\mathfrak{gl}(n)$). It is called the **connection 1-form** related to the covariant derivative operator D. The 2-form F is the **curvature** of the connection, whose general expression is

$$F = dA + \frac{1}{2}[A, A].$$

3.5 Symmetries in general and special Relativity

Einstein's *general covariance principle* states that coordinates do not exist a priori in spacetime, hence physical laws cannot depend on them. Though simple to understand and apply, this principle is really a cumbersome and incomplete way to formally understand the symmetries of the general and special theories of relativity, or any other theory. Firstly because tensor fields, the objects which appear in physical laws, need not be expressed in bases related to the spacetime coordinates (or even in any basis at all, for that matter). Secondly, because the spacetime points themselves need not be represented by coordinates either. So instead of coordinate transformations, the real fundamental physical symmetries deal with *active* transformations of the spacetime points (i.e. diffeomorphisms) on the one hand, and of the tensor fields themselves (local Lorentz transformations) on the other hand. These two classes of transformations are completely independent, a fact which is totally obscured by the coordinate point of view.

Let us follow a coordinate-independent treatment and look for all the possible symmetries of the action functional of general relativity in metric formalism, viz.

$$S\left[g, \boldsymbol{\psi}\right] = \int_{\mathcal{M}} \boldsymbol{\epsilon}[g] \left\{ \alpha(R[g] - 2\Lambda) + \mathcal{L}_M\left[g, \boldsymbol{\psi}\right] \right\}.$$

Here g is the spacetime metric and $\boldsymbol{\psi}$ represents all the matter fields. These are the dynamical variables of the theory. Functional derivation of the action with respect to them yields the desired Euler-Lagrange equations via Hamilton's principle. Furthermore, $\boldsymbol{\epsilon}$ is the natural volume element given by the metric, R is the Ricci scalar of its Levi-Civita connection, $\mathcal{L}_M$ is the matter Lagrangian density and α and Λ are empirical constants. The square brackets denote functional dependence, which possibly includes derivatives of the fields. For the matter fields, usually derivatives up to first order only are considered, though in the case of the metric second order derivatives are necessary to build the curvature scalar, and hence the action. The integration is to be taken over the whole spacetime manifold $\mathcal{M}$. An alternative approach (Palatini's) would be to consider the metric and the connection as independent dynamical variables. As long as the connection is supposed torsion free, this still gives Einstein's equations. The advantage here is that the metric compatibility condition for the connection is not introduced a priori but comes from the Euler-Lagrange equation for the connection. In both approaches, one can also replace the metric by the tetrad as independent variable.

The action above has in fact three independent classes of symmetries, which we explore below.

The first class comprises the internal (gauge) transformations of the matter fields, which by definition leave the matter Lagrangian invariant. These may be point-dependent, provided the corresponding connection forms (i.e. gauge fields with suitable transformation properties) are properly included as dynamical fields in $\mathcal{L}_M$. Of course these transformations have no effect on the remaining terms of S, since all the dependence on the matter fields is bore by $\mathcal{L}_M$, so they really are symmetries of the total action.

The second class of symmetries is made of all smoothly local (i.e. point-dependent) orientation-preserving (i.e. proper orthochronous) Lorentz rotations of the frame fields of the frame bundle $F(M)$. Not all linear transformations are allowed, but only the ones given above, because the action depends explicitly on the metric, by means of the volume element, the curvature scalar and also the inner products of matter fields which are taken with respect to it. This restricts us to transformations preserving the metric, that is, (pseudo-)orthogonal ones. Furthermore, the spacetime manifold has both an orientation and time-orientation. This is also reflected in the action in the volume element, which restricts our transformations even more, to the identity component of the Lorentz group.

The third and deepest class of symmetries corresponds to the diffeomorphisms of the spacetime manifold. These are very different in nature from the previous two. Gauge transformations alter only the *form* (that is, the value at each point) of fields defined in internal spaces. Similarly, local Lorentz transformations affect the form of tensor fields defined on the manifold's tangent structure. None of these actually transforms the spacetime points themselves. A diffeomorphism does exactly that, but in such a way that the manifold obtained is indistinguishable in every respect from the original one: a diffeomorphism is by definition a smooth bijection with smooth inverse from the manifold to itself, that is, an automorphism in the category of smooth manifolds.

However, should the spacetime points alone be modified, with the forms of the fields maintained, a diffeomorphism would not in general be a symmetry of the action. This is because one would change the variable of integration without changing the functional form of the integrand and the domain of integration accordingly. The actual symmetry thus consists of a diffeomorphism ϕ of the spacetime manifold accompanied by the corresponding induced transformation ϕ_* on the tensor fields, which is the *pushforward* operation defined in section 3.3, plus the change on the domain of integration D into $\phi(D)$. The pushforward is exactly that transformation necessary to make every new field $\phi_* T$ assume at each new point $\phi(p)$ the value that the old field T assumed at the corresponding old point p (which is the same *physical* point). With this induced map, all fields actually behave as scalars under diffeomorphisms: their change in form exactly compensates the change in the spacetime points. Any integral of fields on spacetime, including the action functional, is thus kept intact under the combined transformation

$$
\begin{aligned}
p &\to \phi(p) \\
T &\to \phi_* T \\
D &\to \phi(D),
\end{aligned}
$$

A diffeomorphism accompanied by these induced maps is then only a change on the "label" of each spacetime point, and what diffeomorphism invariance means is that such label is unphysical. Hereafter, when we refer to a diffeomorphism, we will mean this whole package.

We conclude then that S is gauge invariant, local Lorentz invariant and diffeomorphism invariant. In other words, it is a scalar with respect to these three kinds of transformation. It is worth to be stressed that all these three kinds of transformation are also symmetries of the matter Lagrangian alone, and hence *every* relativistic field theory, in either flat or curved spacetime, with or without gravity, necessarily presents diffeomorphism and local Lorentz symmetry (besides its internal gauge symmetries). Thus neither diffeomorphisms nor local Lorentz transformations can be considered as the characteristic "gauge freedom of gravity". They are "gauge" freedoms of spacetime alright, but not of gravity in particular. They would still be symmetries of the action if the gravitational effect were "turned off", e.g. by letting the constant α vanish, so that matter would not couple to spacetime anymore. In fact, since Special Relativity is a particular solution of General Relativity, it must and does have all the symmetries of GR, *plus* its own particular Poincaré group of isometries.

So what does one actually mean when one says that fixed parameter (i.e. global) translations and Lorentz rotations are "symmetries" of Minkowski spacetime, for example? To begin with, we are in this case talking about diffeomorphisms, that is, transformations of the spacetime points, and not of the tangent vectors directly (the latter are just "pushed forward"). In particular, a global Lorentz rotation of this kind is *not* a rotation of the tangent vectors by everywhere the same angle. It is rather a rotation of the *spacetime points* themselves around some given origin, which is itself a fixed point in spacetime. This is a diffeomorphism, and what we are saying is that, no matter which point is chosen to be the origin, no matter the plane or the angle of the rotation, such a transformation is always a "symmetry" of Minkowski spacetime. But wait, have we not just said that *all* diffeomorphisms are physical symmetries? Yes! So what kind of special symmetry is it that these transformations represent which deserves them any particular status? It happens that these diffeomorphisms are *isometries*, which means that the *metric tensor* is preserved by the pushforward operation induced by them. In other words, its *Lie derivative* (4.3) vanishes in the direction of the vector fields which generate them. This implies that in any equation which is written in component form the components of the metric can be replaced by their values in a given basis, and yet such equations remain the same in that basis after an isometry. This fact is what Einstein called *special covariance.* It means that physical laws do not change their form under a certain special class of transformations (the isometries) even if the metric tensor components are replaced by their values in a given basis. Another importance of isometries lies in that each linearly independently continuous family of isometries gives rise to a linearly independent conserved current according to Nöther's theorem.

At this point, one thing which should be very clear (and which we are going to stress in order to secure that it is clear) is the radical difference of character between diffeomorphisms and local Lorentz transformations. Diffeomorphisms are maps of the spacetime manifold to itself, and hence take spacetime points into other spacetime points. Tensors in this case are just carried along unaltered to new points by the corresponding pushforward operation. Local Lorentz transformations, on the other hand, are local transformations of the tangent bundle itself, and hence take vectors from the tangent space of each point into new vectors in the tangent space of the *same* point. They effectively change the *form* of vector fields, while every spacetime point remains "where it is".

Some people have mistakenly argued that diffeomorphisms and local Lorentz transformations are generalizations, respectively, of the translations and Lorentz isometries of Minkowski space. However, from what was said in the above paragraph, this is clearly not true. Every isometry, including these last two kinds, are particular examples of diffeomorphisms (remember isometries are defined as diffeomorphisms which preserve the metric). This means that they are transformations of the spacetime points. As we said, a Lorentz isometry in Minkowski space, for example, is a rotation of the whole spacetime around some given point. Of course, if one uses coordinate bases, the components of tensors will be transformed accordingly, and the expression for this transformation will also be a Lorentz component transformation. This is however different from a local Lorentz transformation, which can also be performed in Minkowski space. This one keeps spacetime points fixed and is independent of coordinates. So the conclusion is once again that Special Relativity possesses both diffeomorphism and local Lorentz symmetry, *plus* its particular isometries, which do not change the form of equations even when the metric is not explicit.

Chapter 4

Klein geometries and Cartan geometries

4.1 Klein geometries

Felix Klein's Erlangen Programme was a successful attempt to classify all geometries known at that time (now called homogeneous spaces, or Klein geometries) in terms of Lie groups. Informally, a homogeneous space is a space in which all the points are equivalent. Intuitively, this means that if one moves from any point to any other on the space one still sees everything exactly the same as before. To be able to check the homogeneity of a space we should then know a set of transformations which bring any point into any other point in the space, preserving whatever structure that space might posses. We know that sets of transformations of an object, in order to behave as actual transformations, must form *groups* (they must be associative, there must be an identity transformation and every transformation must have an inverse). Given an abstract group, one can turn it into the group of transformations of an object by defining an action of the group on that object. An action which allows for any element of a set to be brought to any other element by some some transformation is called *transitive*. Furthermore, if we want our homogeneous space to be smooth, the group in question must be a Lie group, and its action must be smooth. Rigorously, then, a homogeneous space is a set endowed with a smooth transitive action of a Lie group.

Klein's ingenious insight was to reverse the above logic: instead of first giving the space and then searching for its group of symmetries, one could start with the group and obtain a homogeneous space from it by a quotient of Lie groups. To see this, we must observe that every homogeneous space is contained in its automorphism group. It corresponds exactly to that set of transformations which move every point. In other words, it is the full automorphism group, "except for" those transformations which fix some point. These last form a subgroup of the full group, called its *stabilizer*, or *little group*. Formally, a Klein geometry is a quotient G/H, where G is the full automorphism group of the space, H is its stabilizer subgroup, and the manifold G/H corresponds to the homogeneous space itself. Inside this construction Klein could include every homogeneous space then known, including vector and affine spaces with and without metric, projective spaces and curved geometries such as spheric and hyperbolic spaces. Let us then pass to the formal construction of Klein geometries.

Let R be an equivalence relation on a set X, that is, a subset of $X \times X$ such that for all x, y and z in X:

$$(x, x) \in R$$
$$(x, y) \in R \Rightarrow (y, x) \in R$$
$$(x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R.$$

Denote by $[x]_R$ the equivalence class of an element x of X under such relation, that is, the set of all elements equivalent to x:

$$[x]_R = \{y \in X \mid xRy\}.$$

The **quotient set of X modulo R**, denoted X/R, is the set of all equivalence classes in X with respect to R:

$$X/R = \{\, [x]_R \mid x \in X\}.$$

The **quotient map** $q_R : X \to X/R$ associated to R is the map which takes each element of X into its equivalence class:

$$q_R(x) = [x]_R.$$

If X is endowed with a topology T, the **quotient topology of T modulo R**, denoted T/R, is the topology on X/R whose open sets are the subsets of X/R whose preimages under the quotient map are in T:

$$T/R = \left\{U \in X/R \;\middle|\; q_R^{-1}(U) \in T\right\}.$$

The quotient set X/R with the quotient topology X/T is a topological space, called the **quotient space of $S = (X, T)$ modulo R**, and denoted S/R.

Given a group G, a subgroup H of G and an element a of G, the **left** (resp. **right**) **coset** of a with respect to H, denoted aH (resp. Ha) is defined as the set of all elements of G of the form ah (resp. ha) for some h in H,

$$aH = \{ah \mid h \in H\}$$
$$Ha = \{ha \mid h \in H\}.$$

The left/right cosets of G with respect to a given subgroup H form equivalence classes in G. That is, a binary relation $a \sim b$ in G given by either $b \in aH$ or $b \in Ha$ is an equivalence relation. The three defining properties of an equivalence relation are easily proved to be satisfied as follows:

$$Ha \ni ea = a = ae \in aH,$$
$$\text{with } e \text{ the identity in } G,$$
$$\text{hence } \; a \sim a \; \text{ (reflexivity).}$$

$$b{\in}aH \;\Rightarrow\; b{=}ah \;\Rightarrow\; bh^{-1}{=}a \;\Rightarrow\; a{\in}bH \quad \text{and}$$
$$b{\in}Ha \;\Rightarrow\; b{=}ha \;\Rightarrow\; h^{-1}b{=}a \;\Rightarrow\; a{\in}Hb,$$
$$\text{with } h \text{ in } H,$$
$$\text{hence } \; a \sim b \Rightarrow b \sim a \; \text{ (symmetry).}$$

$$b{\in}aH \wedge c{\in}bH \Rightarrow b{=}ah \wedge c{=}bh' \Rightarrow c{=}ahh' \Rightarrow c{\in}aH \quad \text{and}$$
$$b{\in}Ha \wedge c{\in}Hb \Rightarrow b{=}ha \wedge c{=}h'b \Rightarrow c{=}h'ha \Rightarrow c{\in}Ha$$
$$\text{with } h \text{ and } h' \text{ in } H,$$
$$\text{hence } \; a \sim b \wedge b \sim c \Rightarrow a \sim c \; \text{ (transitivity).}$$

The **left** (resp. **right**) **coset space of** G **modulo** H, denoted G/H (resp. $H\backslash G$), is just the quotient set under the corresponding equivalence relation, that is, the set of all cosets in G with respect to H:

$$G/H = \{aH \,|\, a{\in}G\}$$
$$H\backslash G = \{Ha \,|\, a{\in}G\}.$$

There is a natural left (resp. right) action of G on G/H (resp. $H\backslash G$) given by

$$a(bH) = (ab)H$$
$$(Ha)b = H(ab).$$

We can see that this action is transitive because given any bH in G/H, we can get to any other cH in G/H since $c = ab$ for some a in G, namely cb^{-1}, and the analogous argument is valid for right cosets and actions. If G is a topological group, each coset space with the quotient topology is a quotient space, and the action above is a homeomorphism of such space for each element of G. If G is a Lie group and H a closed subgroup of G, there is a unique smooth structure on the coset space such that the action above is a diffeomorphism of the resulting manifold for each element of G. This manifold is called a **homogeneous manifold**, and since its smooth structure is unique it is denoted also by G/H without ambiguity. In fact, any smooth manifold endowed with a transitive action of a Lie group G by diffeomorphisms is diffeomorphic to a homogeneous manifold G/H for some H. Moreover, this H is always the stabilizer of a point in the manifold, which is defined below.

Given an action A of a group G on an object X, the stabilizer of a point x in X under A, which we will denote by $St_A(x)$, is the subset of transformations of G which keep x fixed. This is clearly a subgroup of G, since $ex = x$ (e is the identity in G) and $gx = x \Rightarrow g^{-1}x = x$. Now if for some x and y in X and g in G we have $gx = y$, then it is easy to see that $St_A(y) = gSt_A(x)g^{-1}$. Since conjugation is a group isomorphism, $St_A(y) \cong St_A(x)$. Hence, if the action is transitive, the stabilizers of all points in X are isomorphic. In a homogeneous manifold G/H, the subgroup H is always the stabilizer of some point on the manifold. Since the manifold has infinite points, there are infinite isomorphic copies of H in G. The resulting homogeneous manifolds G/H are clearly all isomorphic. This equivalence class of isomorphic manifolds is called a **Klein geometry**. A Klein geometry G/H is specified by choosing a Lie group G and an *abstract* closed subgroup H, to which the stabilizers of any point in the manifold will be isomorphic. Additionally, a Klein geometry is also a smooth H-principal bundle, with G as total space, G/H as base space and the natural action described above as projection.

Minkowski space is the Klein geometry $ISO_e(3,1)/SO_e(3,1)$, that is, affine Lorentzian oriented and time-oriented $\mathbb{R}^4$. De Sitter space is the Klein geometry $SO_e(4,1)/SO_e(3,1)$, that is, the oriented and time-oriented Lorentzian four-sphere. However, for physical purposes additional structure is necessary, namely a metric.

According to Sylvester's Law, given a nondegenerate symmetric bilinear form B on a vector space V, there is always a basis in V in terms of which B is pseudo-orthonormalized (i.e. represented by a diagonal matrix where each diagonal elements is equal to either 1 or -1). The general linear group $GL(n,\mathbb{R})$ is the automorphism group of the vector space $\mathbb{R}^n$. The orthogonal group $O(n)$ is the automorphism group of $\mathbb{R}^n$ with Euclidean metric. The generalized orthogonal group $O(p,q)$ is the automorphism group of R^n with a pseudometric having p positive and q negative eigenvalues. In the natural representation, the generators of the Lie algebra $\mathfrak{so}(p,q)$ are the $(p+q)\times(p+q)$ square matrices

$$(J_{ab})^c{}_d = \delta^c_a \eta_{bd} - \delta^c_b \eta_{ad},$$

where η is the diagonal matrix with p entries equal to 1 and q entries equal to -1. The algebra is given by the brackets

$$[J_{ab}, J_{cd}] = J_{ad}\eta_{bc} - J_{ac}\eta_{bd} - J_{bd}\eta_{ac} + J_{bc}\eta_{ad}.$$

In particular, for the rotation group $SO(3)$ we have

$$[L_i, L_j] = \epsilon_{ij}{}^k L_k$$

where

$$L_i \equiv \frac{1}{2}\epsilon_i{}^{jk} J_{jk}, \tag{4.1}$$

and for the Lorentz group $SO_e(3,1)$ we have

$$\begin{aligned}
[L_i, L_j] &= \epsilon_{ij}{}^k L_k \\
[B_i, B_j] &= \epsilon_{ij}{}^k L_k \\
[L_i, B_j] &= \epsilon_{ij}{}^k B_k,
\end{aligned}$$

where we have broken the symmetry of the Lorentz generators into the rotation generators L_i (4.1) and the boost generators $B_i \equiv J_{0i}$, which physically corresponds to the choice of a preferred time direction, i.e. an observer. This symmetry breaking defines the Klein geometry $SO_e(3,1)/SO(3)$. Given a point in a Lorentzian manifold, each point in this geometry represents the choice of a time direction from that point, that is, an instantaneous observer on it.

Expanding still, we can consider the Lie algebra of the full Poincaré group $ISO_e(3,1) \equiv SO_e(3,1) \oslash \mathbb{R}^4$:

$$\begin{aligned}
[P_a, P_b] &= 0 \\
[P_a, M_{bc}] &= -P_b\eta_{ca} + P_c\eta_{ba} \\
[M_{ab}, M_{cd}] &= M_{ad}\eta_{bc} - M_{ac}\eta_{bd} - M_{bd}\eta_{ac} + M_{bc}\eta_{ad}
\end{aligned} \tag{4.2}$$

where the $M_{ab} = J_{ab}$ ($a = 0, ..., 3$) are the Lorentz generators, and the P_a are the translation generators. The symmetry here is already broken into an $SO_e(3,1)$ and an $\mathbb{R}^4$ sector, which defines the Klein geometry $ISO_e(3,1)/SO_e(3,1)$, that is just Minkowski space. If we also consider the symmetry breaking between timelike and spacelike generators, as done above for the homogeneous Lorentz group, this becomes

$$[H, H] = 0$$
$$[P_i, P_j] = 0$$
$$[H, P_i] = 0$$
$$[H, L_i] = 0$$
$$[P_i, L_j] = -\epsilon_{ij}{}^k P_k$$
$$[H, B_i] = -P_i$$
$$[P_i, B_j] = -H\eta_{ij}$$
$$[L_i, L_j] = \epsilon_{ij}{}^k L_k$$
$$[B_i, B_j] = \epsilon_{ij}{}^k L_k$$
$$[L_i, B_j] = \epsilon_{ij}{}^k B_k,$$

where $H \equiv P_0$. This symmetry breaking defines the Klein geometry $ISO_e(3, 1)/SO(3)$. Each point in this manifold amounts to the specification of a point in Minkowski space plus a preferred time direction from that point, that is, an instantaneous observer. This geometry is then the space of all instantaneous observers on Minkowski space, also called its *observer space* [104]. This is a 7-dimensional smooth manifold.

The Lie algebra of the de Sitter group $SO_e(4, 1)$ can also have the symmetry of its generators broken into Lorentz generators and de Sitter translation generators, respectively given by

$$M_{ab} = J_{ab}$$
$$\Pi_a = -\frac{1}{l} J_{4a}$$

where $a, b \neq 4$ and l is the length parameter of the de Sitter hyperboloid considered, which determines the scale of the symmetry breaking (see section 5.1). This division of the algebra corresponds physically to the choice of a point in the hyperboloid. The brackets are then

$$[M_{ab}, M_{cd}] = M_{ad}\eta_{bc} - M_{ac}\eta_{bd} - M_{bd}\eta_{ac} + M_{bc}\eta_{ad}$$
$$[\Pi_a, \Pi_b] = -\frac{1}{l^2} M_{ab}$$
$$[M_{ab}, \Pi_c] = \frac{1}{l}(\Pi_a\eta_{bc} - \Pi_b\eta_{ac}).$$

This symmetry breaking defines the Klein geometry $SO_e(4, 1)/SO_e(3, 1)$, which is just de Sitter space. One notable fact stemming from the second bracket is that in de Sitter space translations in different directions do not commute. Now dividing the Lorentz generators into boosts and rotations again, we finally have the brackets

$$
\begin{aligned}
[\Pi_0, \Pi_0] &= 0 \\
[\Pi_0, \Pi_i] &= -\frac{1}{l^2} B_i \\
[\Pi_i, \Pi_j] &= -\frac{1}{l^2} \epsilon_{ij}{}^k L_k \\
[L_i, \Pi_0] &= 0 \\
[L_i, \Pi_j] &= -\frac{1}{l} \epsilon_{ij}{}^k \Pi_k \\
[B_i, \Pi_0] &= \frac{1}{l} \Pi_i \\
[B_i, \Pi_j] &= \frac{1}{l} \Pi_0 \eta_{ij} \\
[L_i, L_j] &= \epsilon_{ij}{}^k L_k \\
[B_i, B_j] &= \epsilon_{ij}{}^k L_k \\
[L_i, B_j] &= \epsilon_{ij}{}^k B_k .
\end{aligned}
$$

This amounts to choosing both a point and a frame in de Sitter space, that is, the Klein geometry $SO_e(4,1)/SO(3)$. This geometry is the space of all Lorentz frames on de Sitter space, that is, its observer space. We see that in general, for a given homogeneous spacetime (i.e. Minkowski, de Sitter or anti-de Sitter) with automorphism group G, its observer space is defined as the Klein geometry $G/SO(3)$.

4.2 Cartan geometries

Cartan geometries are manifolds endowed with Cartan connections. A Cartan connection is, roughly speaking, an principal connection with a rule for a symmetry breaking. So for example a $SO(3)$ connection might have its symmetry broken into a $SO(2)$ part and a $\mathbb{R}^2$ part. That is possible because the Lie algebra $\mathfrak{so}(3)$ can be formally split into a $\mathfrak{so}(2)$ part and a $\mathbb{R}^2$ part. However, this process is better understood geometrically. Let a sphere S^2 roll over some two-dimensional surface. Then, if the sphere does not slip nor twist, the point of the sphere which gets in contact with the surface moves according to definite rules. Defining a point of contact amounts to breaking the symmetry of the sphere. After some finite motion of the sphere on the surface, the new configuration of the sphere can be compared to the original one. The difference between the two is an element of the sphere's symmetry group $SO(3)$. However, since now the symmetry of the sphere is broken by the introduction of a tangency point, the comparison between these two configurations can be divided in two parts: one which encodes the information about the "translation" of the point of tangency, and a another one which tells about the rotation of the sphere around the axis defined by the original point of tangency.

The Cartan connection then encodes all this information, properly separated into a spin connection and a coframe field.

The crucial role of a Cartan connection, as with any other connection, is to provide a rule for parallel transport of objects along the manifold on which it is defined. However, a Cartan connection is more general than, for example, the affine connections of Riemannian geometry. This is because the parallel transport defined by the Cartan connection can be based on more general Lie groups. In the case of four dimensional Lorentzian spacetime, the rotation of frames is always governed by the Lorentz group $SO_e(3,1)$. However, the displacement of frames along spacetime depends on the choice of a more larger group containing the Lorentz group, which might be for example the Poincaré group $ISO_e(3,1)$, the de Sitter group $SO_e(3,1)$ or the anti-de Sitter group $SO_e(3,2)$. This choice of a symmetry amounts to the choice of a model Klein geometry to approximate the local structure of spacetime. That is why Cartan geometries can be understood as the geometries of generalized waywisers [122], that is, homogeneous spaces rolling over surfaces. Such rolling of the model Klein geometry over the base manifold determines the rule for parallel transport which is encoded in the Cartan connection. When the model Klein geometry is rolled along an infinitesimal loop starting from a given point in the base manifold, at the end of the journey it might have changed both orientation and point of contact. The part of the Cartan connection encoding information about the change in orientation is the spin connection ω, while the part encoding information about the change in contact point is the coframe field e. This is reflected in the split of the Lie algebra as a graded vector space, a direct sum of two subspaces via the isomorphism [9]

$$\mathfrak{g} \cong \mathfrak{h} \oplus \mathfrak{g}/\mathfrak{h}$$

corresponding to the symmetry breaking defined by the choice of a point on the model Klein geometry. When the model Klein geometry performs a finite rolling along the base manifold, it traces a curve on the manifold, and also a curve on itself, marking the points where both surfaces were in contact. This motion provides the desire rule for parallel transport of objects belonging to the model Klein geometry along any path on the base manifold in terms of elements of the group of symmetry of the model Klein geometry.

Formally speaking, a Cartan connection based on the Klein geometry G/H is a G-principal connection 1-form A on a smooth manifold M equipped with a reduction of its structure group to H and such that the connection 1-form θ linearly identifies each tangent space of M with the tangent space $\mathfrak{g}/\mathfrak{h}$ of the Klein geometry G/H, that is, for each point p in M the composition of the action of the connection 1-form A with the canonical projection from $\mathfrak{g}$ to $\mathfrak{g}/\mathfrak{h}$,

$$T_p \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{h},$$

is an isomorphism of vector spaces. The base manifold M together with such a connection one-form is called a Cartan geometry. The **spin connection** ω and **coframe field** e are obtained from the Cartan connection by the splitting of the Lie algebra (4.3), which gives

$$A = \omega + e$$

where ω is a $\mathfrak{h}$-valued 1-form and e is a $\mathfrak{g}/\mathfrak{h}$-valued 1-form. The 1-form ω is itself a connection on the reduced H-bundle. If one applies this same splitting to the curvature 2-form F of the Cartan connection A, one gets one part [123]

$$R = d\omega + \frac{1}{2}[\omega, \omega]$$

related to the spin connection ω and one part

$$T = de + \frac{1}{2}[\omega, e]$$

related to the coframe field e. The first one is just the curvature of the principal connection ω. The other one is called the **torsion** of the Cartan geometry. Note that the curvature R is such that it vanishes if M is isomorphic to the Klein geometry G/H, so that curvature here means deviation from the model Klein geometry, which is considered flat in this context. By its turn, the coframe field (also called **vielbein** of **soldering form**) finds no analogy in terms of principal connections (and thus neither does it have any analog in Yang-Mills types of theories). It expresses the infinitesimal dislocation of the point of contact between the model Klein geometry and the base manifold, while the spin connection represents the infinitesimal rotation of the model Klein geometry.

Chapter 5

De Sitter-Cartan geometry and cosmology

5.1 De Sitter space and stereographic coordinates

De Sitter space is defined as the Klein geometry

$$dS = SO_e(4,1)/SO_e(3,1)$$

endowed with a metric g induced from its natural embedding in 5-dimensional Minkowski vector space, that is, vector $\mathbb{R}^5$ with a Lorentzian metric. This embedding is defined as the set of points of such vector $\mathbb{R}^5$ with constant norm l under that metric:

$$||\chi||^2 = -\left(\chi^0\right)^2 + \left(\chi^1\right)^2 + \left(\chi^2\right)^2 + \left(\chi^3\right)^2 + \left(\chi^4\right)^2 = l^2. \tag{5.1}$$

This manifold is simply connected and geodesically complete, but not compact. The natural topology inherited from 5D Minkowski vector space by means of the induced metric is $\mathbb{R} \times S^3$.

With a Levi-Civita connection, the curvature scalar is constant throughout the manifold, which is equivalent to saying that it lodges the maximal number of 10 independent Killing fields for a Riemannian manifold of 4 dimensions, i.e., it is maximally symmetric. These fields correspond exactly to the 10 elements of the de Sitter algebra which generate the de Sitter group.

For our purposes it will be useful to study the limits of the de Sitter space and group for vanishing and infinite radius. The must suitable coordinate patch for such purpose is the stereographic one. It is defined in analogy with the stereographic projection of the 2-sphere:

$$\frac{\chi^\alpha}{x^\alpha} = \frac{\Sigma}{\sigma} = \frac{1}{2}\left(1 - \frac{\chi^4}{l}\right) \equiv \Omega_l,$$

where

$$\Sigma \equiv \left(\eta_{\alpha\beta}\chi^\alpha\chi^\beta\right)^{1/2} \quad \text{and} \qquad \sigma \equiv \left(\eta_{\alpha\beta}x^\alpha x^\beta\right)^{1/2}.$$

Here and henceforth, Greek indices run from 0 to 3 and $\eta = \mathrm{diag}(-1,1,1,1)$. Thus the coordinate change is given simply by

$$x^\alpha = {\Omega_l}^{-1}\,\chi^\alpha.$$

This projection is defined for the whole de Sitter hyperboloid except for the points which have $\chi^4 = l$ (these are analogous to the "north pole" in the stereographic projection of the 2-sphere).

In order to get the inverse transformations, which depend on Ω_l, we must start by finding χ^4 in terms of x. For such, let us get back to the definition of the de Sitter space, rewritten in a more convenient fashion (5.1):

$$\frac{1}{l^2}\Sigma^2 + \left(\frac{\chi^4}{l}\right)^2 = 1\,.$$

Here we use

$$\begin{aligned}\Sigma^2 &= \sigma^2\,{\Omega_l}^2 \\ &= \left(\frac{\sigma}{2}\right)^2\left(1-\frac{\chi^4}{l}\right)^2.\end{aligned}$$

Defining

$$\zeta \equiv \left(\frac{\sigma}{2l}\right)^2$$

and

$$\hat{\chi}^4 \equiv \frac{\chi^4}{l},$$

he have

$$\begin{aligned}\zeta\left(1-\hat{\chi}^4\right)^2 &= 1-\left(\hat{\chi}^4\right)^2 \\ &= \left(1+\hat{\chi}^4\right)\left(1-\hat{\chi}^4\right).\end{aligned}$$

Since $\hat{\chi}^4 \neq 1$, then

$$\zeta\left(1-\hat{\chi}^4\right) = 1+\hat{\chi}^4\,,$$

so that

$$\zeta = \frac{1+\hat{\chi}^4}{1-\hat{\chi}^4} \qquad \text{and} \qquad \hat{\chi}^4 = \frac{\zeta-1}{\zeta+1} \ ,$$

which gives

$$\begin{aligned} \Omega_l &= \frac{1}{2}\left(1-\frac{\zeta-1}{\zeta+1}\right) \\ &= \frac{1}{\zeta+1} \ . \end{aligned}$$

Now, with the expression for Ω_l in terms of x in hands, we can write the inverse transformations

$$\chi^\alpha = \Omega_l \, x^\alpha \qquad \text{and} \qquad \chi^4 = l \, \frac{\zeta-1}{\zeta+1} = l \ \Omega_l \, (\zeta-1) \, .$$

The conformally flat character of the de Sitter metric is explicit in the stereographic coordinates, since the metric components in these coordinates are [125]

$$g_{\alpha\beta} = {\Omega_l}^2 \, \eta_{\alpha\beta} \, . \tag{5.2}$$

5.2 De Sitter group and contractions

The group of all the symmetries of de Sitter space is the de Sitter group $O(4,1)$. In terms of the 5D coordinates its generators are written as

$$J_{AB} = H_{AC}\chi^C \frac{\partial}{\partial \chi^B} - H_{BC}\chi^C \frac{\partial}{\partial \chi^A},$$

where $H = \text{diag}(-1,1,1,1,1)$. Transforming to stereographic coordinates we have, for $\alpha, \leq 4$,

$$\begin{aligned} J_{\alpha\beta} &= H_{\alpha C}\chi^C \frac{\partial}{\partial \chi^\beta} - H_{\beta C}\chi^C \frac{\partial}{\partial \chi^\alpha} \\ &= \eta_{\alpha\gamma}\chi^\gamma \frac{\partial}{\partial \chi^\beta} - \eta_{\beta\gamma}\chi^\gamma \frac{\partial}{\partial \chi^\alpha} \\ &= \eta_{\alpha\gamma} \ \Omega_l x^\gamma \ \Omega_l^{-1} \frac{\partial}{\partial x^\beta} - \eta_{\alpha\gamma} \ \Omega_l \, x^\gamma \ \Omega_l^{-1} \frac{\partial}{\partial x^\alpha} \\ &= \eta_{\alpha\gamma} x^c \frac{\partial}{\partial x^\beta} - \eta_{\alpha\gamma} x^\gamma \frac{\partial}{\partial x^\alpha} \\ &= M_{\alpha\beta}, \end{aligned}$$

where we identify the Lorentz generators from (4.2). Here we note explicitly that the Lorentz group is a subgroup of the de Sitter group.

For the remaining four generators we have

$$J_{4\beta} = H_{4C}\chi^C \frac{\partial}{\partial \chi^\beta} - H_{\beta C}\chi^C \frac{\partial}{\partial \chi^4} \; . \tag{5.3}$$

Now we use

$$H_{4C}\chi^C = H_{44}\chi^4 = \chi^4 = l\,\Omega_l\,(\zeta - 1) \; ,$$

so that the first term in (5.3) becomes

$$\begin{aligned} H_{4C}\chi^C \frac{\partial}{\partial \chi^\beta} &= l\,\Omega_l\,(\zeta-1)\;\Omega_l^{-1}\frac{\partial}{\partial x^\beta} \\ &= l\;(\zeta-1)\,\frac{\partial}{\partial x^\beta}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \frac{\partial}{\partial \chi^4} &= \frac{\partial x^\alpha}{\partial \chi^4}\frac{\partial}{\partial x^\alpha} \\ &= \frac{\partial}{\partial \chi^4}\left(\Omega_l^{-1}\,\chi^\alpha\right)\frac{\partial}{\partial x^\alpha} \\ &= \chi^\alpha\left(-\Omega_l^{-2}\right)\frac{\partial}{\partial \chi^4}\left(\frac{1-\bar{\chi}^4}{2}\right)\frac{\partial}{\partial x^\alpha} \\ &= x^\alpha\,\Omega_l^{-1}\,\frac{1}{2l}\,\frac{\partial}{\partial x^\alpha}, \end{aligned}$$

so that the second term in (5.3) becomes

$$\begin{aligned} H_{\beta C}\chi^C \frac{\partial}{\partial \chi^4} &= \left(\eta_{\beta\gamma}\,\Omega_l\,x^\gamma\right)\left(\frac{1}{2l}\,\Omega_l^{-1}\,x^\alpha\frac{\partial}{\partial x^\alpha}\right) \\ &= \frac{1}{2l}\,\eta_{bc}\,x^c x^a\,\frac{\partial}{\partial x^a}. \end{aligned}$$

Thus we have finally

$$\begin{aligned} J_{4\beta} &= l\;(\zeta-1)\,\frac{\partial}{\partial x^\beta} - \frac{1}{2l}\,\eta_{\beta\gamma}\,x^\gamma x^\alpha\,\frac{\partial}{\partial x^\alpha} \\ &= \frac{\sigma^2}{4l}\frac{\partial}{\partial x^\beta} - l\frac{\partial}{\partial x^\beta} - \frac{1}{2l}\,\eta_{\beta\gamma}\,x^\gamma x^\alpha\,\frac{\partial}{\partial x^\alpha} \\ &= -l\frac{\partial}{\partial x^\beta} + \frac{1}{4l}\left(\sigma^2\delta^\alpha_\beta - 2\eta_{\beta\gamma}\,x^\gamma x^\alpha\right)\frac{\partial}{\partial x^\alpha} \\ &= -lP_\beta - \frac{1}{4l}K_\beta\,, \end{aligned}$$

where we idenfity the generators of Poincaré translations, P_β, and special conformal transformations, K_β [125]. The generators $J_{4\beta}$ produce the so-called de Sitter transvections, which take any point to any other in de Sitter space [118].

To study the limit of the de Sitter group for big l, it is useful to define the normalized generator

$$\Pi_\beta \equiv \frac{J_{4\beta}}{-l} = P_\beta + \frac{1}{4l^2} K_\beta. \tag{5.4}$$

Then, when $l \to \infty$,

$$\Pi_\beta \to P_\beta$$

and the Poincaré translations are obtained. Thus, in this limiting process, the de Sitter group contracts to the Poincaré group, and de Sitter space approaches Minkowski space.

On the other hand, for the limit of vanishing l, we define the generator

$$\bar{\Pi}_\beta \equiv -4lJ_{4\beta} = 4l^2 P_\beta + K_\beta.$$

This way, for $l \to 0$ we have that

$$\bar{\Pi}_\beta \to K_\beta.$$

Thus when l vanishes the de Sitter transvections reduce to pure special conformal transformations. Thus, in this limiting process, a new group is obtained which analogous to the Poincaré group, except that translations are replaced by special conformal transformations [121], and de Sitter space approaches the cone space [120], which is transitive under those transformations.

5.3 De Sitter algebra and conformal transformations

As seen in section 4.1, the generators $M_{\alpha\beta}$ and Π_β of the de Sitter group satisfy the algebra

$$[M_{\alpha\beta}, M_{\gamma\delta}] = M_{\alpha\delta}\eta_{\beta\gamma} - M_{\alpha\gamma}\eta_{\beta\delta} - M_{\beta\delta}\eta_{\alpha\gamma} + M_{\beta\gamma}\eta_{\alpha\delta}$$
$$[\Pi_\alpha, \Pi_\beta] = -\frac{1}{l^2} M_{\alpha\beta}$$
$$[M_{\alpha\beta}, \Pi_\gamma] = \frac{1}{l}(\Pi_\alpha \eta_{\beta\gamma} - \Pi_\beta \eta_{\alpha\gamma}).$$

From (5.4), we have that the components of a Killing vector field $\xi_{(\alpha)}$ related to a de Sitter transvection are expressed in stereographic coordinates as

$$\xi^\mu_{(\alpha)} = \delta^\mu_\alpha + \frac{1}{4l^2}(\eta_{\alpha\rho} x^\mu x^\rho - \delta^\mu_\alpha \sigma^2).$$

The Lie derivative of an arbitrary vector field v with respect to $\xi_{(\alpha)}$ is

$$\pounds_{\xi_{(\alpha)}} v = \xi^\mu_{(\alpha)} \partial_\mu v - v^\mu \partial_\mu \xi^\nu_{(\alpha)}.$$

In stereographic coordinates, we have

$$\left(\pounds_{\xi_{(\alpha)}} v\right)^{\nu} = \Pi_{\alpha} v^{\nu} - v^{\mu} \left(\Sigma_{\alpha}\right)_{\mu}{}^{\nu},$$

where Π_{α} is given by (5.4) and

$$\left(\Sigma_{\alpha}\right)_{\mu}{}^{\nu} = \frac{1}{2l^2} x^{\rho} \left(\eta_{\alpha\rho} \delta^{\nu}_{\mu} + \eta_{\mu\alpha} \delta^{\nu}_{\rho} - \eta_{\rho\mu} \delta^{\nu}_{\alpha}\right).$$

The matrix operators Σ_{α} correspond to the special conformal part of the de Sitter transvections. They do not satisfy the de Sitter algebra [118]. In fact, an explicit calculation gives

$$\begin{aligned} \left[\Sigma_{\alpha}, \Sigma_{\beta}\right]_{\mu}{}^{\nu} = \frac{1}{4l^2} x^{\rho} x^{\sigma} & \left(\eta_{\alpha\mu} \eta_{\beta\rho} \delta^{\nu}_{\sigma} - \eta_{\mu\alpha} \eta_{\rho\sigma} \delta^{\nu}_{\beta} - \eta_{\rho\mu} \eta_{\beta\sigma} \delta^{\nu}_{\alpha} \right. \\ & \left. - \eta_{\beta\mu} \eta_{\alpha\rho} \delta^{\nu}_{\sigma} + \eta_{\mu\beta} \eta_{\rho\sigma} \delta^{\nu}_{\alpha} + \eta_{\rho\mu} \eta_{\alpha\sigma} \delta^{\nu}_{\beta}\right). \end{aligned}$$

This means that they cannot be viewed as de Sitter generators [118]. Only the derivative operators Π_{α} and the whole Lie derivative correspond to true representations of de Sitter transformations. Indeed, the Lie derivative also satisfies the de Sitter algebra, as should be expected [118]:

$$\left[\pounds_{\xi_{(\alpha)}}, \pounds_{\xi_{(\beta)}}\right] = -\frac{1}{l^2} \pounds_{\zeta_{(\alpha\beta)}},$$

where $\zeta_{(\alpha\beta)}$ is a Killing vector corresponding to a Lorentz transformation.

When applied to the metric tensor g, the Lie derivative with respect to ξ is written

$$\left(\pounds_{\xi_{(\alpha)}} g\right)_{\mu\nu} = \Pi_{\alpha} g_{\mu\nu} + g_{\mu\lambda} \left(\Sigma_{\alpha}\right)_{\nu}{}^{\lambda} + g_{\nu\lambda} \left(\Sigma_{\alpha}\right)_{\mu}{}^{\lambda}.$$

In de Sitter space, where the ξ are isometries, the above expression vanishes identically, and we have, using (5.2),

$$\begin{aligned} \Pi_{\alpha} g_{\mu\nu} &= -g_{\mu\lambda} \left(\Sigma_{\alpha}\right)_{\nu}{}^{\lambda} - g_{\nu\lambda} \left(\Sigma_{\alpha}\right)_{\mu}{}^{\lambda} \\ &= -\frac{x_{\alpha}}{l^2 \Omega_l} g_{\mu\nu}. \end{aligned}$$

For a transformation parameter ϵ, this corresponds to a conformal factor of [118]

$$\beta^2 = 1 - \frac{\epsilon \cdot x}{l^2 \Omega_l}.$$

Thus the application of Π_{α} on the de Sitter metric results in a conformal rescaling.

5.4 The cone space and cosmology

In the formalism of Cartan geometries one may have a spacetime modeled over de Sitter space, but with varying l [119]. In this case, the cosmological term would be dynamic and could vary with respect to local properties of the matter fields, possibly ranging from Planck to cosmological scales. While the spacetime obtained from de Sitter space in the limit $l \to \infty$ is the well known Minkowski spacetime, the other limit, $l \to 0$, produces the so-called cone spacetime [120]. This spacetime is transitive under special conformal transformations alone. Its group of symmetries is then analogous to the Poincaré group, except that the translation subgroup is replaced by the special conformal subgroup.

To study the cone space in detail as well as its approach from de Sitter space it is convenient to define the 'inverse' host space coordinates [124]

$$\bar{\chi}^A = \chi^A / 4l^2, \tag{5.5}$$

in terms of which relation (5.1) assumes the form

$$\eta_{AB}\, \bar{\chi}^A \bar{\chi}^B + \left(\bar{\chi}^4\right)^2 = \frac{1}{16 l^2}. \tag{5.6}$$

The inverse stereographic projection is now given by

$$\bar{\chi}^\alpha = \bar{\Omega}\, x^\alpha \tag{5.7}$$

and

$$\bar{\chi}^4 = -\, l\, \bar{\Omega} \left(1 + \sigma^2 / 4l^2\right) \tag{5.8}$$

where

$$\bar{\Omega} \equiv \frac{\Omega}{4l^2} = \frac{1}{4l^2 - \sigma^2}. \tag{5.9}$$

In these coordinates, the de Sitter metric assumes the form

$$\bar{g}_{\alpha\beta} = \bar{\Omega}^2\, \eta_{\alpha\beta}. \tag{5.10}$$

If we now take the limit $l \to 0$ leading to the cone space, the metric (5.10) becomes

$$\bar{g}_{\alpha\beta} \to \bar{\eta}_{\alpha\beta} = \sigma^{-4}\, \eta_{\alpha\beta}.$$

The cone spacetime has vanishing Riemann, Ricci and scalar curvature tensors [120]. The cosmological term, on the other hand, diverges. In fact, its metric tensor $\bar{\eta}_{\alpha\beta}$, given by Eq. (5.11), is singular for points in which $\sigma^2 = 0$. However, if we perform a conformal re-scaling of the metric [124]

$$\bar{\eta}_{\alpha\beta} \;\to\; \bar{\eta}'_{\alpha\beta} = \omega^2\,\bar{\eta}_{\alpha\beta},$$

with the conformal factor given by

$$\omega^2 = \sigma^4\,\alpha^2, \tag{5.11}$$

the resulting metric tensor

$$\bar{\eta}'_{\alpha\beta} = \alpha^2\,\eta_{\alpha\beta} \tag{5.12}$$

is no longer singular. It is then clear that we are dealing with a conformal gauge singularity, as discussed in section 2.3. This is a hint that the cone spacetime might represent the very early universe. Further theoretical evidence comes from its thermodynamic properties. The de Sitter horizon has entropy [108]

$$S = \frac{\pi c^3 k_B l^2}{G\hbar}, \tag{5.13}$$

where k_B is the Boltzmann constant, and temperature

$$T = \frac{\hbar c}{2\pi k_B l}. \tag{5.14}$$

The flux of energy across the horizon must then satisfy the relation

$$dE = T\,dS. \tag{5.15}$$

For a constant cosmological term, the energy, entropy and volume will be constant, and the above equation will be trivially satisfied. However, for a varying cosmological term, the de Sitter pseudo radius l will change with time, and so will do E, S, and T. In this case, (5.15) can be integrated to give [124]

$$E = \frac{c^4\,l}{2G}. \tag{5.16}$$

The corresponding energy density is given by

$$\varepsilon \equiv \frac{E}{V} = \frac{3c^4}{8\pi G l^2}. \tag{5.17}$$

In the limit $l \to 0$ the temperature of the de Sitter horizon diverges. The entropy, on the other hand, vanishes, as does the energy. However, the energy density becomes infinite. We can thus say that the cone spacetime is not only geometrically but also thermodynamically consistent with what one should expected for an initial condition for the universe. Furthermore, if conformal cyclic cosmology is the answer to the smooth Big Bang problem, then the cone spacetime might play the role of a satisfying model for the interface connecting two aeons. It certainly satisfies the Weyl curvature hypothesis both in the original sense of having vanishing Weyl tensor and in Tod's sense of possessing a conformal gauge singularity, which can be removed by a conformal rescaling of the metric so that geodesics can be continued past it. Besides that, as we said before, the cone spacetime is transitive under pure special conformal transformations. This means that ordinary translational degrees of freedom are turned off, and thus ordinary gravitational degrees of freedom are also turned off [124]. This corresponds to the idea of a very low gravitational entropy around the Big Bang, as discussed in section 1.3. This fact is matched by the already mentioned vanishing of the curvature tensors and scalar in this spacetime, again confirming its conformity with a scenario of vanishing gravitational degrees of freedom.

Although it is at this point clear that the cone spacetime suits well the role of an initial cosmic condition, it might still seem puzzling that it might also represent the latest stages of cosmic evolution. After all, why should Λ become infinite at late cosmic times? To answer this question one must understand the physical meaning of the parameter l. As can be seen, for example, in equation (5.4), it represents the abundance of not conformally invariant matter over the conformally invariant type. In late cosmic times, the universe is supposed to be dominated by conformally invariant matter. What we are proposing here is a deeper connection between the matter content of the universe and its geometry than even ordinary general relativity leads one to believe exists. More precisely, the proposal is that, if the matter content per itself does not allow the distinction between certain physical properties, then it does not make sense to talk about such distinction in the geometric structure either. Particularly, if the matter content does not allow for proper distances and proper times to be meaningfully defined, no spacetime translations might exist, and no metric based on them might play any physical role. In this case, only conformal transformations have meaning, and only the part of the de Sitter metric associated with them is physically relevant. So the limit of l going to zero *physically* means that massive fields are completely overwhelmed by their massless counterparts. That is why, in the case os a positive cosmological term, the late universe, just like the early one, should be expected to approach the cone spacetime, which is the limit of de Sitter spacetime when massive fields tend to become irrelevant.

Chapter 6

Conclusions

The time asymmetry of the second law of thermodynamics, despite the time symmetry of the fundamental laws of physics, points to the presence of a special boundary condition on the beginning of the universe. Such specialness is encoded in the high level of spatial uniformity of the Big Bang universe. Observer selection bias and cosmic inflation seem unable to account for that uniformity. However, Paul Tod has noticed that the special character of the Big Bang can be mathematically encoded in the requirement that a suitable conformal rescaling of the metric allows it to become a smooth hypersurface across which geodesics could be extended. Roger Penrose notes that the later stages of cosmic evolution, after Hawking evaporation of final black holes has ended, should be dominated by uniformly distributed massless matter, just like the early stages, where particles are ultrarelativistic and rest mass becomes physically irrelevant. Together with Tod's theorem, this might point to a cyclic nature for cosmic history, where each cycle (called *aeon*) should end in a state which becomes the Big Bang to the next one. Such proposal is called conformal cyclic cosmology.

Another fundamental problem in theoretical physics is the inclusion of a length scale parameter into kinematics. This is motivated both by the introduction of a fundamental length scale brought in by quantum mechanics, and by the observation of the existence of a positive cosmological constant. However, these two scales are widely different, one being about the shortest length possibly to make physical sense, the other about the curvature radius of the universe. This fact suggests that the length scale parameter might not be a constant, but a variable parameter depending on local physical properties. This kind of picture is not allowed by standard general relativity and Riemannian geometry, but can possibly be devised under the formalism of Cartan geometries. Hendrik Jennen is currently developing a mathematical model of such kind of geometry [119].

The presence of a length parameter l in the kinematics requires that the fundamental local symmetry of the universe be governed by the de Sitter group $SO(4,1)$, and that the structure of spacetime be locally approximated to that of de Sitter space. The stereographic projection of the de Sitter hyperboloid shows that its transitivity transformations can be split into a pure translational part and a pure proper conformal. This split is physically significant since it gives information about the massive and massless sectors of the matter fields separately. In the limit of large l de Sitter space reduces to Minkowski space, which is transitive under pure translations, while in the limit of small l it contracts to another flat homogeneous spacetime called the cone spacetime. This spacetime seems an appropriate approximation for both early and late stages of cosmic history since it is transitive under pure special conformal transformations, which correspond well to a scenario where standard rest mass, together with proper distance and proper time, bear no physically meaningful part [118].

It is important to remark that de Sitter, anti-de Sitter, Minkowski and the cone spacetime are all maximally symmetric Lorentzian Klein geometries, which can be constructed simply as coset spaces of Lie groups, without mention to any gravitational theory. In these sense, all are equally fundamental and apt to be used as a basis for kinematics. The choice of one over the others must be made based on physical facts. In the absence of a cosmological constant, Minkowski space is the correct choice. However, current empirical evidence points to the existence of a positive cosmological constant, which signals towards de Sitter spacetime as a fundamental basis for kinematics. On the other hand, the relative dominance of massive fields over massless ones varies along cosmic history, effectively changing the relative physical importance of the pure translational sector of the de Sitter transvections over the pure special conformal sector. In the limit where massless fields dominate, the cone spacetime becomes the appropriate description of local physics. Also of relevance is the fact that, from the point of view of Cartan geometries, all the above four spacetimes are considered flat, since they are perfect Klein geometries, which furthermore stresses their equivalence in terms of fundamental geometric status.

Lastly, the cone spacetime satisfies Penrose's original Weyl curvature hypothesis, since it has vanishing Weyl tensor, and also the version of the hypothesis based on Tod's theorem, in that its singularity can be removed by a conformal rescaling to provide a smooth hypersurface across which geodesics may be extended. Other than that, its thermodynamic properties are also consistent with what should be expected about both the very early universes, since it has vanishing entropy, infinite temperature and infinite energy density. For all these reasons, insofar as conformal cyclic cosmology remains a viable framework, the cone spacetime might serve as a good model for the interface between two consecutive aeons in this theory [124].

Bibliography

[1] Martin Bojowald, *Initial Conditions for a Universe*, Gen. Rel. Grav. **35** 1877–1883 (2003)

[2] Giovanni Amelino-Camelia, *Testable scenario for Relativity with minimum-length*, Phys. Lett. B **510** 255–263 (2001), arXiv:hep-th/0012238

[3] Giovanni Amelino-Camelia, *Relativity in space-times with short-distance structure governed by an observer-independent (Planckian) length scale*, Int. J. Mod. Phys. D **11** 35 (2002), arXiv:gr-qc/0012051

[4] João Magueijo & Lee Smolin, *Lorentz invariance with an invariant energy scale*, Phys. Rev. Lett. **88** 190403 (2002), arXiv:hep-th/0112090

[5] Giovanni Amelino-Camelia, *Doubly-Special Relativity: First Results and Key Open Problems*, Int. J. Mod. Phys. D **11** 1643 (2002), arXiv:gr-qc/0210063

[6] Giovanni Amelino-Camelia, *Relativity: Special treatment*, Nature **418** 34–35 (2002) arXiv:gr-qc/0207049

[7] Giovanni Amelino-Camelia, *Doubly-Special Relativity: Facts, Myths and Some Key Open Issues*, Symmetry **2**, 230–271 (2010), arXiv:1003.3942

[8] R. W. Sharpe, *Differential geometry: Cartan's generalization of Klein's Erlangen program*, (Springer, New York, 1997)

[9] Derek K. Wise, *MacDowell-Mansouri gravity and Cartan geometry*, Class. Quantum Grav. **27**, 155010 (2010), arXiv:gr-qc/0611154

[10] J. B. Hartle, *The physics of 'Now'*, Am. J. Phys. **73**, 101–109 (2005), arXiv:gr-qc/0403001

[11] A. S. Eddington, *The Nature of the Physical World* (The Macmillian Company, New York, 1928), Chapter IV

[12] Jos Uffink, *Rereading Ludwig Boltzmann*, in Petr Hájek, Luis Valdés-Villanueva & Dag Westerstahl (eds.), *Logic, Methodology and Philosophy of Science. Proceedings of the Twelfth International Congress* (King's College Publications, London, 2004)

[13] Ludwig Boltzmann, *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, Wiener Berichte **66**, 275-370 (1872)

[14] Ludwig Boltzmann, *Über die beziehung dem zweiten Haubtsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht*, Wiener Berichte **76**, 373–435 (1877)

[15] Claude Shannon, *A mathematical theory of communication*, The Bell System Technical Journal **27**, 379–423 (1948)

[16] P. C. W. Davies, *The Physics of Time-Asymmetry*, (Surrey University Press, 1974)

[17] A. Einstein, "Cosmological Considerations on the General Theory of Relativity" in: H. A. Lorentz, A. Einstein, H. Minkowski & H. Weyl, *The Principle of Relativity* (Methuen, 1923), translated by W. Perrett & G. B. Jeffery from the German original: *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, Sitz. d. Preuß. Akad. d. Wiss. (Berlin) 142–152 (1917)

[18] W. de Sitter, *On the relativity of inertia*, Proceedings of the Royal Netherlands Academy of Arts and Sciences (KNAW) **19**, 1217–1225 (1917)

[19] W. de Sitter, *On the curvature of space*, Proceedings of the Royal Netherlands Academy of Arts and Sciences (KNAW) **20**, 229–243 (1918)

[20] Roger Penrose, *Singularities in Cosmology*, in M. S. Longair (ed.), *Confrontation of Cosmological Theories with Observational Data* (D. Heidel, Boston, 1974)

[21] R. Penrose, *Singularities and time-asymmetry*, in S. W. Hawking & W. Israel (eds.), *General Relativity and Cosmology: An Einstein Centenary Survey* (Cambridge University Press, Cambridge, 1979)

[22] P. C. W. Davies, *Inflation and time asymmetry in the universe*, Nature **301**, 398–400 (1983)

[23] Don N. Page, *Inflation does not explain time asymmetry*, Nature **304**, 39–41 (1983)

[24] Stefan Hollands & Robert M. Wald, *An alternative to inflation*, Gen. Rel. Grav. **34**, 2043–2055 (2002), arXiv:gr-qc/0205058

[25] Huw Price, *On the Origins of the Arrow of Time: Why There is Still a Puzzle about the Low Entropy Past*, in C. Hitchcock, (ed.), *Contemporary Debates in the Philosophy of Science* (Blackwell, 2004).

[26] R. H. Dicke, P. J .E. Peebles, P. G. Roll & D. T. Wilkinson, *Cosmic Black-Body Radiation*, ApJ **142**, 414–419 (1965)

[27] A. A. Penzias & R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s*, ApJ **142**, 419–421 (1965)

[28] G. E. Tauber & J. W. Weinberg, *Internal State of a Gravitating Gas*, Phys. Rev. **122**, 1342–1365 (1961)

[29] J. Ehlers, P. Geren & R. K. Sachs, *Isotropic solutions of the Einstein-Liouville equations*, J. Math. Phys. **9**, 1344–1349 (1966)

[30] S. W. Hawking & G. F. R. Ellis, *The large scale Structure of space-time*, (Cambridge University Press, Cambridge, 1973), pages 348–359

[31] Chris A. Clarkson & Richard K. Barrett, *Does the Isotropy of the CMB Imply a Homogeneous Universe? Some Generalised EGS Theorems*, Class. Quant. Grav. **16**, 3781–3794 (1999), arXiv:gr-qc/9906097

[32] W. R. Stoeger, R. Maartens & G. F. R. Ellis, *Proving almost-homogeneity of the universe: an almost Ehlers-Geren-Sachs theorem*, ApJ **443**, 1–5 (1995)

[33] G. Efstathiou et al., *Evidence for a non-zero Λ and a low matter density from a combined analysis of the 2dF Galaxy Redshift Survey and cosmic microwave background anisotropies*, Mon. Not. R. Astron. Soc. **330**, L29–L35 (2002), arXiv:astro-ph/0109152

[34] Beth A. Reid et al., *Cosmological Constraints from the Clustering of the Sloan Digital Sky Survey DR7 Luminous Red Galaxies*, Mont. Not. R. Astron. Soc. **404.1**, 60–85 (2010) arXiv:0907.1659

[35] H. P. Robertson, *On the foundations of relativistic cosmology*, PNAS **15**, 822–829 (1929)

[36] Roger Penrose, *Difficulties with Inflationary Cosmology*, Ann. N. Y. Acad. Sci. **571**, 249–264 (1989)

[37] Nick Bostrom, *Anthropic Bias: Observation Selection Effects in Science and Philosophy* (Routledge, 2002)

[38] C. B. Collins & S. W. Hawking, *Why is the universe isotropic?*, ApJ **180**, 317–334 (1973)

[39] Charles W. Misner, *The isotropy of the universe*, ApJ **151**, 431–457 (1968)

[40] Charles W. Misner, *Relativistic Fluids in Cosmology*, in Cecile M. DeWitt & John A. Wheeler (eds.), *Battelle Rencontres: 1967 Lectures in Mathematics and Physics* (W. A. Benjamin, 1968)

[41] Alan Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, Phys. Rev. D **23**, 347–356 (1981)

[42] A. D. Linde, *A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Phys. Lett. B, **108**, 389–393 (1982)

[43] Andreas Albrecht and Paul Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, Phys. Rev. Lett. **48**, 1220–1223 (1982)

[44] A. D. Linde, *The new inflationary universe scenario*, in G. W. Gibbons, S. W. Hawking & S. Siklos (eds.), *The Very Early Universe* (Cambridge University Press, 1983)

[45] A. D. Linde, *Chaotic Inflation*, Phys. Lett. B **129**, 177–181 (1983)

[46] A. D. Linde, *The inflationary universe*, Rep. Prog. Phys. **47**, 925–986 (1984)

[47] A. D. Linde, *Initial conditions for inflation*, Phys. Lett. B **162**, 281–286 (1985)

[48] A. D. Linde, *Eternally existing self-reproducing chaotic inflationary universe*, Phys. Lett. B **175**, 395–400 (1986)

[49] A. D. Linde, *Eternally existing self-reproducing inflationary universe*, Phys. Scr. **T15**, 169–175 (1987)

[50] Dalia S. Goldwirth, *Inhomogeneous initial conditions for inflation*, Phys. Rev. D **43**, 3204–3213 (1991)

[51] Dalia S. Goldwirth & Tsvi Piran, *Initial conditions for inflation*, Physics Reports **214**, 223–292 (1992)

[52] Lisa Dyson, Matthew Kleban & Leonard Susskind, *Disturbing implications of a cosmological constant*, JHEP **0210**, 011 (2002), arXiv:hep-th/0208013

[53] G. W. Gibbons & Neil Turok, *Measure Problem in Cosmology*, Phys. Rev. D **77**, 063516 (2006), arXiv:hep-th/0609095

[54] Andreas Albrecht, *de Sitter equilibrium as a fundamental framework for cosmology*, in Thomas Elze (ed.), *4th International Workshop DICE 2008: From Quantum Mechanics through Complexity to Spacetime: The Role of Emergent Dynamical Structures 2008* (IOP, Bristol, 2008)

[55] Anna Ijjas, Paul J. Steinhardt & Abraham Loeb, *Inflationary schism*, Phys. Lett. B **736**, 142-146 (2014), arXiv:1402.6980

[56] Alan H. Guth, David I. Kaiser & Yasunori Nomura, *Inflationary paradigm after Planck 2013*, Phys. Lett. B **733**, 112-119 (2014), arXiv:1312.7619

[57] Matthew C. Johnson & Jean-Luc Lehners, *Cycles in the multiverse* Phys. Rev. D **85**, 103509 (2012), arXiv:1112.3360

[58] Alan H. Guth, *Eternal inflation and its implications*, J. Phys. A: Math. Theor. **40**, 6811–6826 (2007), arXiv:hep-th/0702178

[59] V. Mukhanov, *CMB, Quantum fluctuations and the Predictive Power of Inflation*, arXiv:astro-ph/0303077

[60] Paul Steinhardt, *Theories of Anything* (2014) http://edge.org/response-detail/25405

[61] Alexander Vilenkin, *Birth of inflationary universes*, Phys. Rev. D **27**, 2848 (1983)

[62] A. S. Goncharov, A. D. Linde & V. F. Mukhanov, *The global structure of the inflationary universe*, Int. J. Mod. Phys. A **2**, 561 (1987)

[63] A. D. Linde, *Eternal chaotic inflation* Mod. Phys. Lett. A **01**, 81 (1986)

[64] Andrew Liddle, *An Introduction to Modern Cosmology* (Wiley, New York, 2003), pp. 101–102

[65] Edward W. Kolb & Michael S. Turner, *The Early Universe* (Addison-Wesley, New York, 1990), pp. 261–263

[66] J. A. Peacock, *Cosmological Physics* (Cambridge University Press, Cambridgee, 1998), pp. 323

[67] Jayant V. Narlikar, *Introduction to Cosmology*, Second edition (Cambridge University Press, Cambridge, 1993), pp. 179–181

[68] Patrick Peter & Jean-Philippe Uzan, *Primordial Cosmology* (Oxford University Press, Oxford, 2009), pp. 453

[69] Peter Coles & Francesco Lucchin, *Cosmology — The Origin and Evolution of Cosmic Structure*, Second Edition (Wiley, New York, 2002), pp. 147–149

[70] Steven Weinberg, *Cosmology* (Oxford University Press, Oxford, 2008), pp. 205–206

[71] Viatcheslav Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, Cambridge, 2005), pp. 226–227

[72] Barbara Ryden, *Introduction to Cosmology* (Addison-Wesley, 2002) pp. 193–195

[73] Malcolm S. Longair, *Galaxy Formation*, Second Edition (Springer, New York, 2008), pp. 621

[74] Scott Dodelson, *Modern Cosmology* (Academic Press, 2003) pp. 142-144

[75] Adrew Liddle, op. cit., pp. 105–106

[76] Edward W. Kolb & Michael S. Turner, op. cit., p. 273

[77] J. A. Peacock, op. cit., p. 325

[78] Patrick Peter & Jean-Philippe Uzan, op. cit., pp. 537–538

[79] Peter Coles & Francesco Lucchin, op. cit., pp. 149–151

[80] Steven Weinberg, op. cit., pp. 201–202

[81] Barbara Ryden, op. cit., pp. 200–202

[82] Malcolm S. Longair, op. cit., p. 626

[83] Scott Dodelson, op. cit., pp. 144–150

[84] B. J. Carr & M. J. Rees, *The anthropic principle and the structure of the physical world*, Nature **278**, 605–612 (1979)

[85] Huw Price, *Time's Arrow & Archimedes Point: New Directions for the Physics of Time*, pp. 85–86 (Oxford University Press, Oxford, 1996)

[86] Andreas Albrecht, *Cosmic Inflation and the Arrow of Time*, in J. D. Barrow, P. C. W. Davies & C. L. Harper eds., *Science and Ultimate Reality: From Quantum to Cosmos* (Cambridge University Press, Cambridge, 2003)

[87] P. C. W. Davies, *Inflation in the universe and time asymmetry*, Nature **312**, 534–527 (1984)

[88] Huw Price, op cit., pp. 32–37

[89] Paul Steihardt, *The Inflation Debate*, Sci. Am. **304**, 36–43 (2011)

[90] Anthony Aguirre, *Eternal inflation, past and future*, in Rudy Vaas ed. *Beyond the Big Bang: Competing Scenarios for an Eternal Universe* (Springer, Berlin, 2008)

[91] Alexander Vilenkin, *Open Universes, Eternal Inflation, and the Anthropic Principle*, Int. J. Theor. Phys. **38**, 3135–3145 (1999)

[92] J. Richard Gott III, *Creation of open universes from de Sitter space*, Nature **295**, 304–307 (1982)

[93] Martin Bucher & Neil Turok, *Open inflation with an arbitrary false vacuum mass*, Phys. Rev. D **52**, 5538 (1995), arXiv:hep-ph/9503393

[94] Martin Bucher, Alfred S. Goldhaber & and Neil Turok, *Open universe form inflation*, Phys. Rev. D **52**, 3314 (1995), arXiv:hep-ph/9411206

[95] Ludwig Boltzmann, *On Certain Questions of the Theory of Gases*, Nature **51**, 413–415 (1895)

[96] Brandon Carter, *Large Number Coincidences and the Anthropic Principle in Cosmology*, in M. S. Longair (ed.), *Confrontation of Cosmological Theories with Observational Data* (D. Reidel, 1974)

[97] Lee Smolin, *Scientific alternatives to the anthropic principle*, in Bernard Carr (ed.), *Universe or multiverse?* (Cambridge University Press, Cambridge, 2009), arXiv:hep-th/0407213

[98] P. A. M. Dirac, *The Cosmological Constants*, Nature **139**, 323 (1937)

[99] P. A. M. Dirac, *A New Basis for Cosmology*, Proc. R. Soc. Lond. A **165**, 199–208 (1938)

[100] R. H. Dicke, *Dirac's Cosmology and Mach's Principle*, Nature **192** ,440–441 (1961)

[101] George Gale, *The Anthropic Principle*, Sci. Am. **275**, 154–171 (1981)

[102] Steven Weinberg, *The Cosmological Constant Problem*, Rev. Mod. Phys. **61**, 1–23 (1989)

[103] Vitaly Vanchurin & Alexander Vilenkin, *Predictability crisis in inflationary cosmology and its resolution*, Phys. Rev. D **61**, 083507 (1999), arXiv:gr-qc/9905097

[104] Steffen Gielen & Derek K. Wise, *Lifting General Relativity to Observer Space*, J. Math. Phys. **54**, 052501 (2013), arXiv:1210.0019

[105] Robert M. Wald, *The Arrow of Time and The Initial Conditions of the Universe*, arXiv:gr-qc/0507094

[106] Sean M. Carroll & Jennifer Chen, *Spontaneous Inflation and the Origin of the Arrow of Time* (2002), arXiv:hep-th/0410270

[107] Sean M. Carroll & Jennifer Chen, *Does Inflation Provide Natural Initial Conditions for the Universe?*, Int. J. Mod. Phys. D **14**, 2335 (2005), arXiv:gr-qc/0505037

[108] G. W. Gibbons & S. W. Hawking, *Cosmological event horizons, thermodynamics, and particle creation*, Phys. Rev. D **15**, 2738–2751 (1977)

[109] S. W. Hawking, *Black holes and thermodynamics*, Phys. Rev. D **13**, 191–197 (1976)

[110] Roger Penrose, *The Emperor's New Mind — Concerning Computers, Minds and the Laws of Physics*, Chapter 7 (Penguin Books, New York, 1989)

[111] Roger Penrose, *Cycles of Time — An Extraordinary New View of the Universe* (The Bodley Head, London, 2010)

[112] S. W. Hawking & G. F. R. Ellis, op. cit., pp. 41

[113] Charles W. Misner, Kip S. Thorne & John Archibald Wheeler, *Gravitation*, (W. H. Freeman and Company, New York, 1970), pp. 325

[114] Steven Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, (Wiley, New York, 1972), pp. 145

[115] Robert M. Wald *General Relativity*, (University of Chicago Press, 1984), pp. 40

[116] Norman Steenrod, *The Topology of Fiber Bundles* (Princeton University Press, Princeton, 1951)

[117] K. Paul Tod, *Isotropic Cosmological Singularities*, in Jörg Frauendiener, Helmut Friedrich (eds.), *The Conformal Structure of Space-Times* (Springer, Berlin, 2002)

[118] J. G. Pereira, A. C. Sampson & L. L. Savi, *De Sitter transitivity, conformal transformations and conservation laws*, Int. J. Mod. Phys. D **23**, 1450035 (2014), arXiv:1312.3128

[119] Hendrik Jennen, *Cartan geometry of spacetimes with a nonconstant cosmological function* Λ, Phys. Rev. D **90**, 084046 (2014), arXiv:1406.2621

[120] R. Aldrovandi, J.P. Beltrán Almeida & J. G. Pereira, *A singular conformal spacetime*, J. Geom. Phys. **56**, 1042–1056 (2006), arXiv:gr-qc/0403099

[121] R. Aldrovandi & J. G. Pereira *A Second Poincaré Group*, in H. Aratyn et. al eds., *Topics in Theoretical Physics II: Festschrift for Abraham H. Zimmerman* (Ed. Fundação Instituto de Física Teórica, São Paulo, 1998), arXiv:gr-qc/9809061

[122] H. F. Westman & T. G. Zlosnik, *Gravity, Cartan geometry, and idealized waywisers*, arXiv:1203.5709

[123] Gabriel Catren, *Geometrical Foundations of Cartan Gauge Gravity*, arXiv:1407.7814

[124] A. Araujo, H. Jennen, J. G. Pereira, A. C. Sampson & L. L. Savi, *On the spacetime connecting two aeons in conformal cyclic cosmology*, arXiv:1503.05005

[125] Lucas Lolli Savi, *Relatividade Restrita de de Sitter: Uma Abordagem Cinemática*, Master's Dissertation, Section 2.5, IFT-UNESP (2010).