



ELSEVIER

Contents lists available at ScienceDirect

Comptes Rendus Physique

www.sciencedirect.com



Testing quantum gravity with cosmology / Tester les théories de la gravitation quantique à l'aide de la cosmologie

Do we *really* understand the cosmos?

Comprenons-nous vraiment le cosmos ?

Thanu Padmanabhan

IUCAA, Pune University Campus, Ganeshkhind, Pune 411007, India

ARTICLE INFO

Article history:

Available online 24 February 2017

Keywords:

Dynamics of universe

Emergent gravity

Quantum gravity

Spacetime information

Cosmological constant

ABSTRACT

Our *knowledge* about the universe has increased tremendously in the last three decades or so – thanks to the progress in observations – but our *understanding* has improved very little. There are several fundamental questions about our universe for which we have no answers within the current, operationally very successful, approach to cosmology. Worse still, we do not even know how to address some of these issues within the conventional approach to cosmology. This fact suggests that we are missing some important theoretical ingredients in the overall description of the cosmos. I will argue that these issues – some of which are not fully appreciated or emphasized in the literature – demand a paradigm shift: We should not think of the universe as described by a specific solution to the gravitational field equations; instead, it should be treated as a special physical system governed by a different mathematical description, rooted in the quantum description of spacetime. I will outline how this can possibly be done.

© 2017 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

R É S U M É

Notre *connaissance* de l'univers s'est drastiquement accrue au cours de ces trois dernières décennies – grâce au progrès accomplis en matière d'observation – mais notre *compréhension* de celui-ci ne s'est que très peu améliorée. Il est plusieurs questions fondamentales touchant notre univers pour lesquelles nous n'avons pas de réponses avec l'approche actuelle de la cosmologie, au demeurant couronnée d'un grand succès d'un point de vue opérationnel. Pire encore, nous ne savons même pas comment appréhender certains problèmes dans le cadre de l'approche conventionnelle de la cosmologie. Ceci suggère que des ingrédients théoriques importants d'une description complète du cosmos nous font encore défaut. J'affirmerai ici que ces questions – dont la littérature n'a pas encore pleinement pris la mesure ou suffisamment insisté dessus – exigent un changement de paradigme : nous ne devrions pas penser à l'univers comme étant décrit par une solution spécifique aux équations du champ gravitationnel, mais plutôt comme devant être considéré comme un système physique spécifique gouverné par une description mathématique différente, pre-

E-mail address: paddy@iucaa.in.

nant ses racines dans une description quantique de l'espace-temps. Je soulignerai comment ceci peut être fait.

© 2017 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Motivation

Spectacular progress in cosmological observations in the last four decades has helped us to develop a standard model of the universe which is very successful. In this model, the smooth universe is described by a *specific* solution to the field equations of gravity, say, Einstein's equations, and can be parameterized by a small set of numbers ($H_0, \Omega_B, \Omega_{DM}, \Omega_{DE}, \Omega_R$...with their usual meanings). In addition, the formation of structures in the universe is described quite adequately in terms of the growth of perturbations around this smooth universe. These perturbations, generated during an inflationary phase,¹ can be characterized by a power spectrum $P(k) = Ak^n$ with two more parameters, A and n . Both theory and observations are mature enough today to test not only the lowest order predictions (for, e.g., scale invariance of the perturbations, corresponding to $n = 1$), but also higher order effects (like, for, e.g., the deviation $(n - 1)$, in specific models). Thus, on the whole, the description of the universe can be based on a set of well defined parameters which are directly observable.

At the next level of probing, such a description encounters three kinds of difficulties, of which the first two are well-known in the literature and the third one will be the core topic of discussion in this article.

The first kind of difficulty is related to technical issues and details in the model. The following questions, for example, belong to this set: Can we correctly describe the properties and statistics of dwarf galaxies? Do we understand the detailed mechanism which caused the reionization in the universe? Most cosmologists (including me) believe that it is just a question of time before we have satisfactory and consistent answers to such issues within the standard description.

The second kind of difficulty which arises in cosmology is related to the description of the matter sector. Examples are questions like: What is the nature and abundance of the dark matter² particle? How can we explain the baryon-to-photon ratio in our universe? These issues are more fundamental than the first kind of problems but most of us believe that we do have an algorithmic procedure available to attack these problems, *within the framework of conventional cosmology*. For example, a successful extension of the standard model in high energy physics might allow us to compute such numbers from first principles. The current difficulty is then only due to our inadequate understanding of particle physics at high enough energies.

The third kind of problems – which, as I said, we will be concerned with – are those which we have no clue as to how to address. The most important example in this category is the extremely tiny – but non-zero – value of the cosmological constant.³ As regards this set, I am not so much concerned about the lack of a viable solution as with the fact that we do not even know how to properly attack these problems within the framework of conventional cosmology. In some cases, which I will discuss, it is not even clear how to precisely state these problems within the context of the standard model of cosmology.

After some clarifications on the notion of expansion of the universe (Sec. 2) I will describe, in Sections 3 to 6, these foundational conundrums in cosmology. Based on this discussion, I will argue (see Sec. 7) that it is fundamentally incorrect to describe the universe as a specific solution to the gravitational field equations. Instead we should think of the universe a special system and look for a different paradigm to describe its evolution. I will suggest, towards the end of the article, some possible ingredients of such a paradigm and explain (see Sec. 8) how it can solve the cosmological constant problem. I will use the mostly positive signature and set $c = 1$, $\hbar = 1$ and (occasionally) $G = 1$. The Greek indices range over 1, 2, 3 while the Latin indices range over 0–3.

2. Expansion of the universe is in the eye of the beholder

The standard cosmological model is based on a specific solution to gravitational field equations. It is generally believed that one key feature of this solution is the 'expansion' of the smooth background universe which is supposed to distinguish the Friedmann solution from, say, the spacetime describing the region around the Sun. What is not adequately emphasized

¹ I would love to have a viable alternative to the inflationary generation of perturbations, but there is none which can be considered a worthy challenger. So, in this article I will accept the inflationary paradigm as a working hypothesis.

² Verification of Einstein's equations at cosmological scales require testing the hypothesis $G_b^a - \kappa T_b^a = 0$ where $\kappa = 8\pi G$. When the directly observed values of these two tensors $G_b^a(\text{obs})$ and $T_b^a(\text{obs})$ lead to $G_b^a(\text{obs}) - \kappa T_b^a(\text{obs}) \equiv \kappa Q_b^a \neq 0$, as it happens in our universe, Einstein's theory appears to flunk the test. We can then either postulate a modified matter tensor $T_b^a = T_b^a(\text{obs}) + Q_b^a$, (as done in the case of dark matter) or a modification of theory by $G_b^a = G_b^a(\text{obs}) - \kappa Q_b^a$, (as done in the case of dark energy which I take to be the cosmological constant). I will accept both these modifications, viz, the postulates of dark matter and the cosmological constant, in this article. One can question these assumptions, but again I find that all alternatives are much worse theoretically.

³ In this article, I will assume that dark energy is cosmological constant. Other explanations for dark energy are more ad hoc, not demanded by observations, do not explain why cosmological constant is zero and leaves the fine tuning problem unanswered. I do not think these models are better alternatives to the postulate of a cosmological constant.

is the fact that the standard notion of expansion depends on the coordinates you choose to describe the Friedmann model. Geodesic observers in a spacetime will interpret not only the Friedmann metric but also the Schwarzschild metric as ‘expanding’, while non-geodesic observers can find both of them non-expanding – in a precise sense, described below. This tells you that one need not associate the theoretical difficulties we will be discussing later too strongly with the notion of expansion. Since this result, unfortunately, has not attracted necessary attention in the literature⁴ and comes as a surprise even to some experts, I will introduce the Friedmann model from a different perspective and highlight this fact.

It seems reasonable to assume that the smooth, large-scale, spatial 3-geometry of the universe should be homogeneous and isotropic and hence must have a constant 3-curvature which can be taken to be $k = 0, -1$ or $+1$. I will confine my attention to models in which $k = 0$, so that the spatial sections are flat.⁵ Such a maximally symmetric 3-space should also necessarily be spherically symmetric about any given spatial origin. It is, therefore, natural to foliate the 3-space by 2-dimensional spherical surfaces with the metric $dl^2 = r^2(d\theta^2 + \sin^2\theta d\phi^2) \equiv r^2 d\Omega^2$ where r is a radial coordinate with a clear physical meaning: $r = [A/(4\pi)]^{1/2}$ with A being the proper area of the foliating 2-surfaces. Consider now a spacetime metric given by

$$ds^2 = -\frac{1}{24\pi G} \frac{d\rho^2}{\rho(\rho+p)^2} + \left[dr + \frac{r}{3(\rho+p)} d\rho \right]^2 + r^2 d\Omega^2 \tag{1}$$

where the coordinates are chosen to be (ρ, r, θ, ϕ) and $p = p(\rho)$ is a specified function. This metric describes our universe with $p = p(\rho)$ representing an effective equation of state for the matter with p and ρ interpreted as total pressure and total density! If you compute the G_b^a for this metric, you will find that it satisfies Einstein’s equations $G_b^a = \kappa T_b^a$ with a source energy momentum tensor $T_b^a = [\rho + p(\rho)]u^a u_b + p(\rho)\delta_b^a$ where u^a is the four-velocity of geodesic observers in the spacetime. This, in turn implies that $\rho = T_{ab}u^a u^b$, $p = (1/3)T_{ab}h^{ab}$ where $h_{ab} = g_{ab} + u_a u_b$ is the projection tensor orthogonal to the four-velocity u^a of the geodesic observers and T_{ab} is essentially determined by ρ and the function $p(\rho)$.

The line element in Eq. (1) is remarkable in the sense that the spacetime geometry could be expressed directly in terms of the variables which occur in the matter sector of the theory through T_b^a . That is, we have now solved the Einstein’s equations $G_b^a = \kappa T_b^a$ for the metric⁶ and have expressed the components of the metric g_{ab} directly in terms of the components of the stress tensor T_b^a . This is, of course, impossible to do in general but works in the case of the universe only because of its high level of symmetry.

To reduce the metric in Eq. (1) into the conventional form is quite straightforward. Given a function $p = p(\rho)$, compute the indefinite integral

$$t = -\left(\frac{1}{24\pi G}\right)^{1/2} \int \frac{d\rho}{(\rho+p)\sqrt{\rho}} \tag{2}$$

to obtain the function $t = t(\rho)$. Invert this function, locally, to determine $\rho = \rho(t)$ and thus $p = p(\rho) = p(t)$, obtaining ρ and p as functions of t . Define, for convenience, the function $H(t)$ through

$$H(t) \equiv -\frac{1}{3} \frac{d\rho}{dt} \frac{1}{(\rho+p)} = \left(\frac{8\pi G\rho}{3}\right)^{1/2} \tag{3}$$

where the second equality follows from Eq. (2). Transform from the coordinates (ρ, r, θ, ϕ) to the coordinates (t, r, θ, ϕ) , and⁷ you will find that the line interval in Eq. (1) becomes

$$ds^2 = -dt^2 + (dr - H(t)r dt)^2 + r^2 d\Omega^2 \tag{4}$$

Some of you will recognize this line element as representing the Friedmann model in the Painlevé-type coordinates; if you do not, introduce a function $a(t)$ and a coordinate x through the relations

$$H(t) \equiv \frac{\dot{a}}{a}; \quad x \equiv \frac{r}{a} \tag{5}$$

and you will find that the line interval in the coordinates (t, x, θ, ϕ) is given in the familiar form:

$$ds^2 = -dt^2 + a^2(t)(dx^2 + x^2 d\Omega^2) \tag{6}$$

The line element in Eq. (4) contains a single unknown function of time, $H(t)$. The metric as well as the field equation can be expressed entirely in terms of the function $H(t)$. But the function $a(t)$, defined through the first equation in Eq. (5), is not unique and has a scaling degree of freedom, $a \rightarrow \lambda a$. This is obvious when the metric is written as in Eq. (4) because

⁴ I have discussed a few of these issues in Ref. [1] as well as in Ref. [2].

⁵ This is certainly preferred by both observations and theory. Most of the discussion can be generalized to $k \neq 0$, but not all.

⁶ The metric, as it is written, has a singularity if we choose the equation of state to be exactly $p = -\rho$; but this can be handled by a careful limiting procedure or in a different coordinate system. Our universe is never described by a strictly $p = -\rho$ equation of state. Both during the inflationary phase as well as the late time acceleration, this equation of state is approached only asymptotically, and hence the line element is well defined.

⁷ When ρ is a monotonic function of t you can switch from ρ to t trivially; if not, you can still do it locally and glue the definitions together appropriately.

H is invariant under constant rescaling of $a(t)$. This is not apparent if we start with the standard form of the Friedmann metric in Eq. (6) unless we also rescale x . This is usually considered to be a rather trivial matter but it is not. Equation (5) clearly shows that for a given $H(t)$ determined by the source, the corresponding $a(t)$ is *not* unique and is arbitrary with respect to a scaling by a constant. Such a scaling freedom does *not* exist if we use the coordinates in Eq. (4) or Eq. (1) to describe the Friedmann geometry. If we rescale a , then the second equation in Eq. (5) tells us that x is *automatically* rescaled, leaving r fixed. I stress that r has a direct geometrical meaning $(A/4\pi)^{1/2}$ in terms of the area of the $t = \text{constant}$, $r = \text{constant}$ surface. So, the real origin of the scaling freedom in $a(t)$ is from the fact that the geometrical description in Eq. (4) or Eq. (1) cares only for $H(t)$ and that the $a(t)$ arises through the definition in the first equation in Eq. (5). This, in turn, implies that we have the freedom to set $a(t) = 1$ for some value of t while describing the universe. I will say more about this choice later on.

The line elements in Eq. (1) or Eq. (2) give us a very different pictures about the ‘expansion of the universe’, compared to the line element in Eq. (6). Observers comoving with the coordinates (t, r, θ, ϕ) or (ρ, r, θ, ϕ) – i.e. observers with world lines having r, θ, ϕ fixed – will find that the spatial cross-sections [corresponding to $t = \text{constant}$ or $\rho = \text{constant}$] are described by *flat Euclidean 3-space*. In particular, the volume enclosed by the $r = \text{constant}$ surface in 3-space is just $(4\pi/3)r^3$ and the area of the $r = \text{constant}$ surface is $4\pi r^2$. Neither the volume nor the area “expands” as time evolves in this coordinate system! In contrast, observers using the coordinate system (t, x, θ, ϕ) will find that the spatial cross-sections are Euclidean 3-spaces scaled by an overall time dependent factor $a(t)$. The volume enclosed by the surface $x = \text{constant}$ is $(4\pi/3)x^3[a(t)]^3$ and the area of the $x = \text{constant}$ surface is $4\pi x^2[a(t)]^2$. Both this volume and the area change with time and the universe “expands” in this coordinate system if $a(t)$ is an increasing function of time.

The above result demonstrates the title of this subsection. The observers with $\mathbf{x} = \text{constant}$ are geodesic observers and the clocks carried by them measure the cosmic time t . These observers see the universe as expanding. The observers following the world line $\mathbf{r} = \text{constant}$ are *not* geodesic observers. When we use the metric in Eq. (4), the geodesics are described by the equation

$$r \exp \left[- \int H(t) dt \right] = \text{constant} \quad (7)$$

Since we like to think of galaxies to be in geodesic motion in our universe (though no real galaxy follows the geodesic of the smooth Friedmann model), it is useful to use coordinates in which each galaxy has a constant value of x, θ, ϕ rather than a constant value of r, θ, ϕ . This is merely a question of convenience and is not fundamental. The geometrical features of our universe (like, e.g., redshift) do not change under coordinate transformations, and we can indeed talk of all physical phenomena using the metric in Eq. (4) without ever introducing the notion of an ‘expanding’ universe.

In case you find this surprising, let me assure you that there is no swindle. The “expansion” of a spacetime defined in terms of the increase of, say, the proper areas of the surfaces with $t = \text{constant}$, $r = \text{constant}$, is *always* a coordinate dependent effect and can occur in *several* spacetimes. Consider, for example, the following metric:

$$ds^2 = -c^2 dt^2 + \frac{4}{9} \left[\frac{9GM}{2(x+ct)} \right]^{2/3} dx^2 + \left[\frac{9GM}{2} (x+ct)^2 \right]^{2/3} [d\theta^2 + \sin^2 \theta d\phi^2] \quad (8)$$

The proper area A of the 2-surfaces with $t = \text{constant}$, $x = \text{constant}$, increases with time as $A \propto (x+ct)^{4/3}$; similarly, the volume enclosed by the surface $t = \text{constant}$, $x = \text{constant}$, also increases with time. The observers using t, x, θ, ϕ in the spacetime described by the metric in Eq. (8) can claim – just like the observers using the coordinates in Eq. (6) – that their spacetime is expanding.⁸ But the metric in Eq. (8) describes the spacetime outside a spherical star like the Sun! Instead of the standard Schwarzschild coordinates, we are using the coordinates appropriate to freely falling observers [4]. Just as in the case of Eq. (6), the time coordinate t in Eq. (8) denotes the proper time shown by the geodesic observers with trajectories $\mathbf{x} = \text{constant}$. So, even a static spacetime can appear to be expanding when you use geodesic coordinates. The co-moving observers in the standard Schwarzschild coordinates, of course, are non-geodesic observers just as the observers who use Eq. (1) or Eq. (4). There is nothing sacred about geodesic observers and, in fact, we hardly use the coordinates adapted to geodesic observers anywhere in GR except in the case of cosmology.

The reason I brought in the coordinate dependence of the notion of “expanding” universe is the following: As we proceed to discuss several conundrums in Friedmann model, they may appear to be related to the fact that spatial sections of the universe are expanding. This, however, is incorrect. One can discuss all of cosmology including the big bang singularity, the inflationary phase, etc., without this notion, by choosing a different set of coordinates. The metric will still be time dependent – through $H(t)$ in Eq. (4) and through $p(\rho)$ in Eq. (1) – but the spatial sections will be non-evolving Euclidean

⁸ More formally, one can introduce in any spacetime the notion of a congruence of *geodesic* observers with a geodesic velocity field $u^i(x)$. You can define an expansion of this congruence by $\theta \equiv \nabla_i u^i$ which appears to give a geometric, coordinate-independent, definition of expansion in the Friedmann universe. This is true but θ will be non-zero for the geodesic congruence in most spacetimes, including the spacetime around the Sun. If you introduce synchronous coordinates in which the metric is $ds^2 = -dt^2 + h_{\alpha\beta} dx^\alpha dx^\beta$, then the geodesic velocity field is $u^i = \delta_0^i$ and $\theta = \partial_t [\ln \sqrt{h}]$ will be non-zero in general. In the Friedmann universe, the maximal symmetry of space itself gives you a preferred timelike vector, which coincides with the velocity vector of geodesic observers; so θ is independent of x^α , which will not happen in general. The definition of expansion still remains linked to a choice of observers in one way or another even when you use such – more geometric – ideas.

3-space. So it is conceptually inappropriate to attribute all the theoretical issues which we will come across to the fact that the universe is expanding.

3. Everything is allowed in cosmology

In general relativity, we think of the field equation $G_b^a = \kappa T_b^a$ as determining the metric tensor g_{ab} for a *given* source described by T_b^a . We never ask where T_b^a came from; nor do we impose any constraints on the nature of T_b^a except to demand that $\nabla_a T_b^a = 0$. This works quite well in all non-cosmological contexts like, e.g., when you want to determine the gravitational field around a binary pulsar or the gravitational field produced by a massive rotating black hole. You choose the appropriate T_b^a , solve Einstein's equation to get g_{ab} and then work out all the properties.

This approach runs into a curious problem in the context of cosmology. If you use the standard Friedmann coordinates in Eq. (6), then the energy momentum tensor should necessarily be of the form $T_b^a = \text{dia}(-\rho(t), p(t), p(t), p(t))$ having two undetermined functions of time. The metric has one undetermined function $H(t)$ and these *three* functions are related by *two* independent components of Einstein's equations:

$$\rho(t) = \frac{3H^2}{8\pi G}; \quad p(t) = -\frac{1}{8\pi G}[3H^2 + 2\dot{H}] \quad (9)$$

This, in turn, tells you two things: First, the Friedmann model of the universe is under-determined unless you externally specify a relation between p and ρ and put some conditions on what kind of p and ρ are acceptable. If you do not (and most modern cosmologists do not) you can have any evolutionary model of the universe described by any real function $a(t)$ [or $H(t)$] by a suitable choice of the equation of state $p = p(\rho)$! All you need to do is to choose your favorite model of evolution described by some $a(t)$ [which could give you a universe with any kind of features you like], compute $\rho(t)$ and $p(t)$ through Eq. (9) and eliminate t between these two variables to determine an equation of state $p = p(\rho)$ for the material medium populating the universe.⁹ The equation of state will be weird and fine-tuned but such models are routinely published in the literature. Somewhat gratifyingly, $\rho(t)$ given by Eq. (9) will be positive definite but p could have either sign. A source with $p < 0$ would have been unthinkable some five decades back but today, negative pressure sources are not only considered acceptable but also respectable by the current generation of cosmologists. So you can publish a paper with any evolutionary history $a(t)$ if you do not care for laboratory justification for the energy momentum tensor. *Cosmological evolution is, in principle, fundamentally unconstrained* which is an issue we need to recognize.

The situation is made worse by the following factors: If you start with an equation of state which is tested and justified in the laboratory, you will most probably have $\rho > 0, p > 0$. When you evolve the universe backward in time, you will eventually reach energy scales which are not tested in the laboratory. This is going to happen irrespective of how high an energy scale you can explore in the laboratory. So, unless you discover in the lab, forms of matter with an equation of state which violates $p + \rho > 0$ and/or $\rho + 3p > 0$, you are always going to hit an unknown domain. You cannot do cosmology by specifying a lab-tested T_b^a and solving for g_{ab} if (i) your lab-tested T_b^a has $\rho > 0$ and $p > 0$ and (ii) you go sufficiently far into the past. This makes cosmology rather special from the point of view of solving Einstein's equation. You *need* to postulate different forms of T_b^a , work out the observational consequences and iterate on the procedure. Unfortunately this never works out satisfactorily in practice. The near-infinite number of inflationary models available in the market is a simple proof that this procedure is unlikely to give you a predictive, unique description of the very early phases of the universe.

There is also a deeper conceptual issue. It was known right from the early days of GR that the equation $G_b^a = \kappa T_b^a$ equates something extraordinarily beautiful on the left hand side to some ugly structure on the right hand side. This distinction gets worse as we start probing the earlier and earlier phases of the universe. For example, most of the scalar field potentials used in inflationary model building have no particle physics justification and are theoretically unacceptable from the point of view of quantum field theory – even as an effective field theory. Add to it the fact that both matter and geometry has to “come out of nothing” at the beginning and you can see the knot we have tied ourselves in. This is the first of a series of questions I want to raise for which we not only do not have a solution but – even worse – we do not even know how to approach the problem. This difficulty arises because *the geometry does not constrain the matter sector sufficiently* in standard GR, but leaves us with far too many choices.

4. The strange (and stranger) numbers which characterize our universe

Observations are consistent with the idea that our universe can be described in terms of three distinct evolutionary phases: (i) An inflationary phase with an equation of state $p \approx -\rho$; (ii) a radiation-dominated phase with the equation of state $p \approx (1/3)\rho$ followed by a matter-dominated phase with $p \approx 0$. (iii) A late time accelerated phase with the equation

⁹ If the universe is populated by several species of energy density, then ρ and p will be the sum of densities and pressures of each species and, of course, each of them will have their own equation of state. But since total pressure $p(t)$ and total density $\rho(t)$ are just functions of time, you can eliminate t between the two and determine an effective equation of state $p = p(\rho)$ for an ‘effective fluid’ which will produce the same geometry. You can also play the same game – as is often done in various disguises – using a scalar field with a potential $V(\phi)$. I have given an explicit recipe for constructing a $V(\phi)$ for any $a(t)$ in Ref. [11].

of state $p \approx -\rho$ which I will take to be dominated by the cosmological constant. The standard way of describing the evolutionary history of such a model is through the equation

$$H^2(t) = \frac{\dot{a}^2}{a^2} = \begin{cases} H_{\text{inf}}^2 & (a < a_{\text{rh}}) \\ H_0^2 [(1 - \Omega_R - \Omega_m) + \Omega_R a_0^4/a^4 + \Omega_m a_0^3/a^3] & (a > a_{\text{rh}}) \end{cases} \tag{10}$$

where a_{rh} is the epoch of reheating at which the inflation ended. (I have assumed instant reheating and set $k = 0$ for simplicity.) Such a description uses the constant parameters, H_{inf} , H_0 , Ω_R , Ω_m and a_0 . (We usually set to $a_0 = 1$ but there is a subtlety about this choice which I will come back to.)

While these parameters are very convenient to compare observations with theory, they are completely unsuited for describing the universe as a physical system. For example, cosmologists living in a star system located in a galaxy at $z = 8$ will use corresponding parameters evaluated at $z = 8$ which, of course, will differ numerically from the ones we use. In other words, the parameters used in Eq. (10) have no epoch invariant significance and are tied to a very special epoch at which the CMB temperature is 2.73 K. This is unsatisfactory when we want to think of the universe as a physical system described by certain cosmic constants. It is necessary to describe the evolution of the universe using constant parameters which will have epoch-independent significance.

This is fairly easy to do and, in fact, one can do it in an infinite number of ways. One convenient set of parameters to use are the following: (i) We describe the inflationary phase by a constant density ρ_{inf} with the equation $H^2(t) \approx (8\pi G \rho_{\text{inf}})/3$ which will replace the first part of Eq. (10). (ii) Similarly, we introduce another constant density ρ_Λ corresponding to the cosmological constant so that at very late times the universe will be described by the deSitter expansion with $H^2(t) \approx (8\pi G \rho_\Lambda)/3$. (iii) To describe the radiation- and matter-dominated phase, it is convenient to introduce another constant density

$$\rho_{\text{eq}} \equiv \frac{\rho_m^4(a)}{\rho_R^3(a)} = \sigma T_{\text{eq}}^4 \tag{11}$$

where σ is the Stefan–Boltzmann constant and the second equality defines the temperature T_{eq} . We can also introduce the parameter a_{eq} by the epoch-independent definition

$$a_{\text{eq}} \equiv \frac{a \rho_R(a)}{\rho_m(a)} \tag{12}$$

and work with the variable $x \equiv (a/a_{\text{eq}})$. Equation (10) can now be rewritten in the form

$$\left(\frac{\dot{x}}{x}\right)^2 = \begin{cases} (8\pi G/3) \rho_{\text{inf}} \\ (8\pi G/3) [\rho_\Lambda + \rho_{\text{eq}} (x^{-3} + x^{-4})] \end{cases} \tag{13}$$

in terms of the three densities (ρ_{inf} , ρ_{eq} , ρ_Λ). This is a much more meaningful way of describing our universe than by using the parameterization in Eq. (10). In particular, our cosmologist friend who lived in the $z = 8$ galaxy would have written exactly the same equation with exactly the same numerical values¹⁰ for (ρ_{inf} , ρ_{eq} , ρ_Λ). We could also convey to the $z = 8$ cosmologist our normalization convention for $a(t)$: We tell her to set $a(t) = 1$ at the epoch when the CMB temperature was equal to T_{eq} ; this is again an invariant statement characterizing the description of our universe.¹¹ In other words, Eq. (13) describes the universe as a physical system (like, for, e.g., an elastic solid) determined by certain constants (like, for, e.g., the Young’s modulus, etc., for a solid).

Obviously it is meaningful to ask why our universe has certain numerical values for (ρ_{inf} , ρ_{eq} , ρ_Λ) in terms of, say, the Planck density. From observations we know that:

$$\rho_{\text{inf}} < (1.94 \times 10^{16} \text{ GeV})^4 \tag{14}$$

$$\rho_{\text{eq}} = \frac{\rho_m^4}{\rho_R^3} = [(0.86 \pm 0.09) \text{ eV}]^4 \tag{15}$$

$$\rho_\Lambda = [(2.26 \pm 0.05) \times 10^{-3} \text{ eV}]^4 \tag{16}$$

in natural units with $c = 1$, $\hbar = 1$. Several comments are in order vis-a-vis these values.

¹⁰ More precisely, if you divide each of these densities by the Planck density $\rho_{\text{pl}} = c^5/G^2\hbar$, you will get three dimensionless numbers which will be the same as those used by the $z = 8$ cosmologist. So it does not matter that we are using the CGS system which might not have existed at $z = 8$!

¹¹ This is a good time to point out the fact that there are certain constants in the universe the numerical value of which we cannot determine uniquely. For example, consider the combinations like $aT(a)$ or $a\rho_R(a)/\rho_m(a)$. These quantities remain constant, independent of the epoch a at which they are measured, as the universe evolves. But their numerical value depends on the numerical value you attribute to a_0 , which is not determined by theory. Recall that Einstein’s equation only fixes $H(t)$ and not $a(t)$. So while we know that $aT(a) = a_0 T_0 = a_{\text{eq}} T_{\text{eq}}$, we will never know its numerical value without making an additional assumption. This is why it is convenient to use the scaling freedom and set $a = 1$ at the epoch when the radiation temperature was equal to T_{eq} . This is a normalization which is independent of the current epoch and something with which our cosmologist at $z = 8$ will agree.

Today we have no firm theory which determines the numerical values of any of these three densities. However, the *nature* of our ignorance about them differs significantly. The bound in Eq. (14) could actually be replaced by a numerical value if future observations determine the energy scale of inflation. Further, high energy physics does provide a glimmer of hope in eventually coming up with some sensible¹² model of inflation which will determine the density in Eq. (14).

Consider next the numerical value of ρ_{eq} . From the definition, we can relate ρ_{eq} to the ratio between the number density of the photons and the number density of matter particles:

$$\rho_{\text{eq}} = \frac{\rho_m^4}{\rho_R^3} = C \frac{(n_{\text{DM}} m_{\text{DM}} + n_{\text{B}} m_{\text{B}})^4}{n_\gamma^4} = C \left[m_{\text{DM}} \left(\frac{n_{\text{DM}}}{n_\gamma} \right) + m_{\text{B}} \left(\frac{n_{\text{B}}}{n_\gamma} \right) \right]^4 \quad (17)$$

where $C = 15^3 (2\zeta(3))^4 c^3 / \pi^{14} \hbar^3$ is a numerical constant, $n_{\text{DM}}, n_{\text{B}}, n_\gamma$ are the current number densities of dark matter particles, baryons and photons respectively and $m_{\text{DM}}, m_{\text{B}}$ are the masses of the dark matter particle and baryon. We expect the physics at (possibly) GUTs scale to determine the ratios (n_{DM}/n_γ) and (n_{B}/n_γ) and specify m_{DM} and m_{B} . Indeed, we have a framework to calculate these numbers in different models of high energy physics (for a review, see, e.g., [5]) though none of these models can be considered as compelling at present.

Thus we do have a possible theoretical framework for determining ρ_{inf} and ρ_{eq} . But the situation is completely different as regards ρ_Λ . We have no clue what determines the astonishingly small but non-zero numerical value of the cosmological constant characterized by the number (in natural units with $c = 1, \hbar = 1$):

$$(\rho_\Lambda L_p^4) \approx 1.1 \times 10^{-123} \quad (18)$$

The late time evolution of the universe is characterized by the cosmological constant Λ , and the four constants (Λ, G, \hbar, c) describing nature thus lead to the dimensionless combination

$$\Lambda \left(\frac{G\hbar}{c^3} \right) = 8\pi \rho_\Lambda L_p^4 \approx 2.8 \times 10^{-122} \quad (19)$$

which is probably the smallest non-zero number relevant to physics! This issue is well-known and has often been thought of as *the* most fundamental problem in theoretical physics today.

Incidentally, there is another small number which has not acquired the notoriety it probably deserves. You can easily verify that:

$$\rho_{\text{eq}} L_p^4 \approx 2.5 \times 10^{-113} \quad (20)$$

is also extraordinarily small. (This fact comes as a surprise many cosmologists!) How come cosmology literature is totally silent on the fine-tuning problem of ρ_{eq} but gives such a bad press to ρ_Λ ? Do you really believe that 10^{-113} is *not* fine-tuning but 10^{-122} is? That would be a pretty ridiculous stand to take. When confronted with the smallness of $\rho_{\text{eq}} L_p^4$ most people distinguish it conceptually from the smallness of $\rho_\Lambda L_p^4$ by giving two arguments: (a) We *do* have a hope of determining $\rho_{\text{eq}} L_p^4$ from high energy physics through Eq. (17) but we have no clue as to how to go about determining $\rho_\Lambda L_p^4$. (b) It is more likely that Λ has “something to do” with L_p^2 than with ρ_{eq} . These arguments may not sound very satisfactory but they do highlight the importance of the cosmological constant problem as something unique.

So, our universe appears to be a hastily put together, make-shift job, using the three densities $\rho_{\text{inf}}, \rho_{\text{eq}}, \rho_\Lambda$ which *have no relation to each other either conceptually or numerically*. The fact that you can build such an ad-hoc universe is closely related to the theoretical feature I mentioned in Sec. 3. The current theoretical framework which uses *only* Einstein’s equations is insufficient to constrain the matter sector. You can build universes with any set of three numbers $(\rho_{\text{inf}}, \rho_{\text{eq}}, \rho_\Lambda)$ and there is no theoretical principle constraining their values. Clearly we need some *additional* theoretical principle to supplement the gravitational field equations if we have to make sense out of these numbers.

Having described the strangeness of the numerical values of $(\rho_{\text{inf}}, \rho_{\text{eq}}, \rho_\Lambda)$, let me show you something *still stranger*. I invite you to form a specific, dimensionless, number I out of these three densities by the definition:

$$I = \frac{1}{9\pi} \ln \left(\frac{4}{27} \frac{\rho_{\text{inf}}^{3/2}}{\rho_\Lambda \rho_{\text{eq}}^{1/2}} \right) \quad (21)$$

and evaluate its numerical value by plugging in the values in Eq. (15) and Eq. (16) for ρ_{eq} and ρ_Λ respectively and taking $\rho_{\text{inf}}^{1/4} = 10^{15}$ GeV. Surprisingly enough, you will get

$$I \approx 4\pi \left(1 \pm \mathcal{O}(10^{-3}) \right) \quad (22)$$

¹² By ‘sensible’, I mean a model of inflation in which the scalar field driving the inflation, for example, serves some useful purpose other than just driving the inflation.

That is, $I = 4\pi$ to an accuracy of one part in thousand for the values of parameters determined from observations and considered reasonable by cosmologists. This should make you wonder why the right hand side of Eq. (21) has such a nice value as 4π since it is not often that such strange things happen.

Later in this article (see Sec. 8), I will show that: (i) the right hand side of Eq. (21) can actually be interpreted, in a well-defined manner, as the amount of cosmic information accessible to an eternal observer and (ii) the reason it is 4π has to do with the quantum microstructure of spacetime.¹³ Obviously, turning this around and arguing that $I = 4\pi$ from theoretical considerations, one can determine the numerical value of cosmological constant in terms of the other cosmological parameters, ρ_{eq} and ρ_{inf} which – eventually – will be determined from the high energy physics. But, as I said before, this would require a shift in the theoretical paradigm and *cannot* be done within the conventional approach to cosmology. I will come back to this in Sec. 7.

5. Come back aether, all is forgiven?

Our universe selects a preferred Lorentz frame with respect to which you can measure the absolute velocity of your motion. In a few decades your car will be equipped with a gadget which will couple to the CMB, detect its dipole anisotropy and will tell you your velocity vector with respect to the absolute rest frame of the universe in which the CMB is homogeneous and isotropic. Operationally, to a limited extent, this is no different from the good old aether which was providing an absolute reference frame for defining the state of rest. In fact, the universe also provides you with an absolute *time* coordinate in terms of the CMB temperature. If you specify that you are using a coordinate system in which the CMB is homogeneous and isotropic and the CMB temperature is, say 30 K, you have uniquely specified your Lorentz frame (with the only residual symmetry allowed being that of spatial rotations and spatial translations).

The field equation of GR, of course, is generally covariant and does not select out any coordinate system – and, in fact, it is invariant under a much larger group than just the Lorentz group. A *specific solution* to this field equation need not possess the full symmetry of the equation, which is a rather trivial and well known fact. To obtain any specific solution, we need to specify T_b^a which could bring in a natural coordinate system. For example, the metric around Sun has the simplest description if you use a spherically symmetric coordinate system with its origin at the center of the Sun. Similarly, if T_b^a is spatially homogeneous and isotropic, it is probably simplest to describe the universe using the coordinates in Eq. (4) or Eq. (6). This fact, by itself, is not a cause for surprise or concern.¹⁴

Neither does the existence of cosmic ‘aether’ in the form of the CMB violate special relativity in any way. In fact, you can reinvent special relativity using the CMB observations along the following lines: Let A be a geodesic observer who sees the CMB as isotropic (in a local inertial frame in the Friedmann geometry) and let B be another inertial observer, moving with a boosted velocity \mathbf{v} with respect to A, who will see the CMB as anisotropic with a dipole anisotropy $T^{-1}(\theta, \nu) \equiv \beta(\theta, \nu) = \beta_0[a(\nu) + b(\nu) \cos(\theta - \theta_0)]$. By varying the direction and magnitude of her velocity, B can determine the functional forms of $a(\nu)$, $b(\nu)$ and the value of θ_0 purely from local observations. She will find that $a(\nu) = \gamma(\nu)$, $b(\nu) = \gamma(\nu)(\nu/c)$ where $\gamma^{-1} = \sqrt{1 - \nu^2/c^2}$, involving a parameter c which she will recognize is equal to the speed of light. By comparing the results in *three* inertial frames A, B and C and careful reverse engineering, one can motivate the standard velocity addition formula in SR involving the relative velocity of B and C. This, in turn, will tell you that the speed of light is the same in all inertial frames which are boosted with respect to the freely falling frame in Friedmann universe (in which the CMB is isotropic). The rest of SR will follow.¹⁵ All this is interesting but does not create any problem.

But the situation does have a deeper level of subtlety, which raises a conceptual issue. The real peculiarity is not the fact that our universe has a preferred Lorentz frame; *it is the fact that we see no trace of this preference at sub-cosmic scale physics*. To appreciate this issue, you have to recall that the smooth universe is an approximate entity and its description involves averaging the energy momentum tensor of actual clustered matter over sufficiently large scales. In arriving at the standard cosmological model, we *first* average the matter distribution over sufficiently large scales, determine an average $\langle T_b^a \rangle$ and *then* solve $G_b^a = \kappa \langle T_b^a \rangle$ to obtain the Friedmann metric g_{ab}^{FRW} . It is assumed that if we had solved the exact equations $G_b^a = \kappa T_b^a$, found the exact metric g_{ab} and then averaged it over large scales, we would have found that $\langle g_{ab} \rangle = g_{ab}^{\text{FRW}}$ to a sufficient level of accuracy.¹⁶

¹³ It is an *observational fact* that I defined via Eq. (21) has a numerical value 4π for our universe. You need to be a true believer in coincidences if you think such a result can be completely ignored as “just one of those things”!

¹⁴ More formally, in quantum field theoretic language, we assume that the local vacuum state is Lorentz invariant in the suitable limit. The universe with matter and CMB is interpreted as a highly excited state of the vacuum which does not have this symmetry.

¹⁵ Unfortunately, text books in SR often create the impression that the absence of an absolute frame of rest is important for the validity of SR. Our universe *does* have an absolute frame of rest but its existence does not violate either SR or GR.

¹⁶ Aside: This fact is used to raise a bogey in cosmology by someone once in a while but this averaging assumption is *always* implicit in the interpretation of GR field equations. In the Einstein’s equations $G_b^a = \kappa T_b^a$, you are expected to specify the matter energy momentum tensor at every event in the spacetime and then solve these differential equations. But you have no way of specifying the energy momentum tensor precisely at any given event where matter is present! For example, suppose you want to determine the metric inside the Sun. Usually you define T_b^a at an event inside the Sun by treating the matter as a fluid; this T_b^a is an approximate T_b^a obtained by averaging a more exact T_b^a over a region large compared to the mean free path. If you probe matter at still smaller scales you will discover averaging at various scales all the way to, say, quarks and gluons. So, strictly speaking, we are *always* solving the equations $G_b^a = \kappa \langle T_b^a \rangle$. But suppose we had found the metric g_{ab} for the exact T_b^a and averaged the metric over scales large compared to the mean free

But what coordinate system should we use to solve the exact equation $G_b^a = \kappa T_b^a$? The validity of standard GR and general covariance at small scales imply that you could have used any coordinate system you like. If you then do the averaging, your final result will indeed be Friedmann *geometry* but expressed in some strange coordinate system. To recover Friedmann geometry in the standard Friedmann coordinates, you should work with a sub-class of all possible coordinate systems while solving the exact equations $G_b^a = \kappa T_b^a$ and doing the averaging. In other words, if you want the exact metric, averaged over large scales, to exhibit the symmetries of the Friedmann universe explicitly, then you need to restrict the general covariance at small scales.¹⁷

This shows that the existence of an absolute frame of rest at large scales actually selects out a class of coordinate systems with special properties at small scales. But we see no experimental evidence for such a selection at small scales. To the extent we can determine experimentally, there is no trace of an absolute rest frame or even a preferred class of frames in sub-cosmic level physics. The laboratory scale experiments looking for an absolute frame of rest (without using the CMB) have repeatedly drawn a blank, but WMAP or PLANCK has no difficulty in determining it using the CMB, even locally. Roughly speaking, physics at cosmic scales breaks the general covariance (and even Lorentz invariance) operationally by providing us with an absolute standard of rest; but as we move to smaller and smaller scales, we are left with no trace of the cosmic frame of rest in any other phenomena, and the diffeomorphism invariance of equations holds. This enhancement of symmetry¹⁸ as we proceed to smaller scales is definitely a peculiarity of our universe which cries out for an explanation; once again, we have no clue how to go about it.

In summary, the fact that the *entire* universe is filled by sources (definitely the CMB, even if you ignore clustered matter) which maintains the large-scale homogeneity and isotropy is extremely peculiar. It is this substratum which allows us to define a cosmic rest frame, *purely from observations*. You could certainly describe the Friedmann geometry in any coordinate system you like and the physics will not change; this is assured by the fact that general relativity respects general covariance. But it also remains a fact that observers can measure – and indeed they have measured – your absolute velocity with respect to a cosmic rest frame. In fact, paradoxical though it might seem, the Friedmann model provides a *generally covariant* procedure for constructing an *absolute* frame of rest!

6. The arrow of time, expansion and spontaneous classicalization

I will now raise a question which, at the outset, may sound somewhat strange. *Why does the universe expand and, thereby, give us an arrow of time?* To appreciate the significance of this question, recall that Eq. (9) is invariant under time reversal $t \rightarrow -t$. (After all, Einstein's equations themselves are time reversal invariant.) To match the observations, we have to choose a solution with $\dot{a} > 0$ at some fiducial time $t = t_{\text{fid}} > 0$ (say, at the current epoch), thereby breaking the time-reversal invariance of the system. This, by itself, is not an issue for a laboratory system. We know that a particular solution to the dynamical equations describing the system need not respect all the symmetries of the equations. But, for the universe, this *is* indeed an issue.

To see why, let us first discuss the case of $(\rho + 3p) > 0$ for all t . The choice $\dot{a} > 0$, at any instant of time, implies that we are *postulating* that the universe is expanding at that instant. Then Eq. (9) tells us that the universe will expand at all times in the past and will have a singularity ($a = 0$) at some finite time in the past (which we can take to be $t = 0$ without loss of generality). The structure of Eq. (9) prevents us from specifying the initial conditions at $t = 0$. So, if you insist on specifying the *initial* conditions and integrating the equations forward in time, you are forced to take $\dot{a} > 0$ at some $t = \epsilon > 0$, thereby breaking the time reversal symmetry. The universe expands at present 'because' we chose it to expand at some instant in the past. This expansion, in turn, gives us an arrow of time with either t or a can be used as a time coordinate. But why do we have to choose the solution with $\dot{a} > 0$ at some instant? This is the essence of the so-called *expansion problem* [6]. An alternative way of posing the same question is the following: How come a cosmological arrow of time emerges from equations of motion which are time-reversal invariant?

In a laboratory, we can usually take another copy of the system we are studying and explore it with a time-reversed choice of initial conditions, because the time can be specified by degrees of freedom *external* to the system. We cannot do it for the universe because we do not have extra copies of it handy and – equally importantly – there is nothing external to it to specify the time. So the problem, as described, is specific to cosmology.

So far we assumed that $(\rho + 3p) > 0$, thereby leading to a singularity. Since meaningful theories must be nonsingular, we certainly expect a future theory of gravity – possibly a model for quantum gravity – to eliminate the singularity [effectively leading to $(\rho + 3p) < 0$]. Can such a theory solve the problem of the arrow of time? This seems unlikely. To see this, let

path, etc. Will such an average metric match with the solution to $G_b^a = \kappa \langle T_b^a \rangle$ at appropriate length scales? The entire GR works on the assumption that such an averaging is valid in spite of the non-linear nature of Einstein's equations. If you do not assume this, you cannot solve Einstein's equations reliably in any region occupied by normal matter.

¹⁷ You can do this calculation explicitly in the Schwarzschild–deSitter geometry containing a mass M and the cosmological constant Λ . If you express the metric in static, spherically symmetric coordinates and average the metric over scales large compared to the gravitational radius of M , you will indeed recover deSitter *geometry* but in the static coordinates. You have to make a peculiar coordinate transformation of the Schwarzschild–deSitter geometry before averaging if the averaging has to reproduce the deSitter universe in the standard Friedmann coordinates.

¹⁸ I mean the *mathematical* symmetry of the *equations* describing the physics; *not* the *physical* symmetry of the matter distribution. Obviously the matter distribution is more symmetric at large scales than at small scales; this is precisely what prevents us from identifying a special reference frame at small scales, but allows us to do it at large scales!

us ask what kind of dynamics we would expect in such a ‘final’ theory. The classical dynamics will certainly get modified at the Planck epoch, but, away from it, we expect some effective equations (possibly with quantum corrections) to govern the evolution of an (effective) expansion factor. The solutions could, for example, have a contracting phase (followed by a bounce) or could start from a Planck-size universe at $t = -\infty$, just to give two non-singular possibilities. While we do not know these equations or their solutions, we can be confident that they will still be time-reversal invariant because quantum theory, as we know it, is time-reversal invariant. So except through a choice for initial conditions (now possibly at $t = -\infty$), we still cannot explain how the cosmological arrow of time emerges. Since quantum gravity is unlikely to produce an arrow of time, it is a worthwhile pursuit to try and understand this problem in the (semi)classical context.¹⁹

Given all these, it will be nice if we can find a simpler way by which the equations of motion that are time-reversal invariant can lead to an evolution which singles out an arrow of time. At first sight one might think this is impossible but one can manage to do it with unbounded Hamiltonians. I will describe the idea with a simple example. Consider an ‘inverted’ oscillator $q(t)$ obeying the equations of motion

$$\ddot{q} = \omega^2 q \quad (23)$$

This equation is clearly invariant under $t \rightarrow -t$ so one would have thought that no arrow of time will emerge from the dynamics, unless we impose it in the initial conditions. The general solution Eq. (23) is

$$q(t) = q(0) \cosh \omega t + \dot{q}(0) \omega^{-1} \sinh \omega t \quad (24)$$

For a generic initial condition, there is no relationship between $q(0)$ and $\dot{q}(0)$. So at late times (i.e. $t \gg \omega^{-1}$), we find that one branch of the solution is selected out:

$$q(t) \approx \frac{1}{2}(q(0) + \dot{q}(0)\omega^{-1})e^{\omega t} \propto e^{\omega t} \quad (25)$$

leading to an “expansion” and an arrow of time! The solution in Eq. (24) is time-reversal invariant in the sense that $q(t) = q(-t)$ if we let $\dot{q}(0) \rightarrow -\dot{q}(0)$ when we do $t \rightarrow -t$. But once we have chosen a generic solution with some uncorrelated $q(0)$ and $\dot{q}(0)$, the *late time dynamics* picks out an arrow of time correlating the increase of $q^2(t)$ with the increase of t . (Of course, there are special initial conditions like, e.g., $q(0) = -\dot{q}(0)\omega^{-1}$ or $q(0) = 0 = \dot{q}(0)$ for which this will not happen, but these are special choices and not generic.)

One can easily show that this behavior arises for a wide class of Hamiltonians that are unbounded. It is not necessary that the potential energy is unbounded. If the kinetic energy term has the ‘wrong’ sign, so that the Lagrangian has a form like $L = -(1/2)\dot{q}^2 - V(q)$ with a V which is positive and unbounded from above, say, we will again end up with an instability and the late time evolution of q will give an arrow of time.

Interestingly enough, it was known for decades [7] that the expansion factor $a(t)$ does have such a wrong sign in the kinetic energy term in the Hilbert action and hence represents an unstable mode. The above interpretation suggests that *it is this cosmic instability which we call expansion*, and for timescales larger than the Planck time, it picks out an arrow of time. Clearly, the same feature will occur even in any effective theory describing the (semi)classical gravity once $a(t)$ acquires an unstable dynamics. This is guaranteed to happen because any sensible quantum cosmological model will approach the Friedmann model at times larger than the Planck time and – in the Friedmann limit – $a(t)$ is an unstable model.

This instability has relevance for another peculiarity related to the early, quantum regime of the universe which does not seem to have attracted the attention it deserves. It is generally believed that the very early phase of the universe needs to be described quantum mechanically using the principles of quantum gravity, because during these earliest moments, the length scale associated with the curvature will be comparable to the Planck scale. This, however, raises a conundrum not encountered elsewhere in physics. *How come the universe, which started out as a quantum mechanical system, became classical spontaneously as it evolved?!*

To see this issue in proper context, you need to introduce some theoretical structure which can tell you the ‘level of classicality’ of a system. Many such definitions can be given, each suited for specific systems (for a sample of ideas and references to previous work, see [8–10]). Most of them use the idea that a classical system follows a sharply peaked trajectory in *phase space* of the form $p = f(q)$ while a quantum system will not exhibit correlations between q and p . You can choose any sensible descriptor of classicality and ask whether a quantum mechanical system can become classical spontaneously. Just for illustration, consider the Wigner function $W(q, p, t)$ built from a wave function $\psi(t, q)$ describing a quantum system evolving under the action of a Hamiltonian $H(q, p)$. We would like to know what kind of Hamiltonians will ensure that, as $t \rightarrow \infty$, the Wigner function gets sharply peaked on a classical trajectory in the phase space.

The answer is surprising: If the Hamiltonian is bounded, then the system cannot evolve spontaneously to classical behaviour. In other words, if you want to start with a quantum universe and ensure that the evolution takes it to a classical

¹⁹ A more complicated “solution” to the arrow of time issue, which is sometimes suggested, is as follows: Consider a very inhomogeneous initial condition in some $t = t_i$ hypersurface and let us assume that certain regions behave like ‘local’ Friedmann models with $\dot{a} > 0$ and other regions have $\dot{a} < 0$ so that no global arrow of time can be defined from the expansion. Next, assume that the dynamics of these patches are independent of each other and we just happen to be in a patch with $\dot{a} > 0$, thereby ‘solving’ the problem. This scenario has several difficulties. For a generic initial condition, the patches will *not* evolve independently in a nonlinear theory of gravity. Even defining a ‘local expansion factor’ for a ‘local patch’ without assuming special symmetries is impossible. Such scenarios are often invoked in the context of inflationary models but they do not have rigorous mathematical justification.

universe, the effective Hamiltonian describing such an evolution cannot be bounded; rather, it should exhibit an instability from a conventional point of view. As I said before, many of the toy Hamiltonians describing the reduced phase space of gravity do have one degree of freedom – precisely the one corresponding to the expansion factor $a(t)$ – which comes up with a wrong sign for the kinetic energy. The above result tells you that the wrong sign is indeed the right sign if the universe has to become classical on its own.

This result is very special to cosmology. Hamiltonians with negative kinetic energy terms are taboo in laboratory scale physics for good reasons; you do not want run-away situations in the lab. But one can easily accommodate such an instability in the behavior of the universe. *The cosmic expansion itself is just a run-away solution fed by the instability.* This fact also ties up two concepts: (i) the origin of the cosmic arrow of time and (ii) the spontaneous quantum to classical transition made by the universe during its evolution. Normal systems in the lab do not spontaneously evolve into a more and more classical state as time evolves. But this is precisely what systems with unbounded Hamiltonians do. For example, the inverted harmonic oscillator in Eq. (23), treated as a quantum system, does become more and more classical as it evolves.

A more general description of this result is as follows: Decompose the spatial 3-metric in the form $g_{\alpha\beta} = a^2 h_{\alpha\beta}$ with $\det h = 1$ as a gauge condition. Then, using the form of the Einstein–Hilbert Lagrangian, one can show that while the kinetic energy term for $a(t)$ has the wrong sign, the other degrees of freedom, represented by $h_{\alpha\beta}$, have the correct sign. In other words, it is only the overall scale factor of the 3-metric which has an instability. Hence it is this degree of freedom which turns classical first during the evolution. If these features are preserved in the effective quantum-corrected description of gravity, then we can hope to have an explanation for a broader question: Why is the classical universe described by a single dynamical degree of freedom $a(t)$, rather than by, say a Bianchi type-I model with three degrees of freedom?

Unfortunately, the results obtained so far with regards to this question are limited to simple toy models. There is no assurance that they will hold in a more general context of quantum gravity. But if they do, it tells you that the quantum gravitational description of spacetime should contain the seeds for an instability which is rather unexpected. The standard models for high energy physics and many candidate models of quantum gravity shy away from working with unbounded Hamiltonians because they are mathematical nightmares. What is more, this instability does not seem to do much at smaller scales in the universe: we do not see spontaneous nucleation of mini-universes all over space today. It is difficult to embed a quantum cosmological model within the broader context of quantum gravity such that everything will work out fine for the large-scale universe but physics at smaller scales will not have any unwanted instabilities. Once again, it appears that the universe is a rather special system and not just a specific solution to a certain more general gravitational theory.

7. An alternative paradigm for cosmology

Let me summarize the discussion so far before proceeding further: The conventional approach to cosmology, which describes the universe using a specific solution to the gravitational field equations, leads to the following issues.

1. Einstein's equation, in general, does not put any constraint on T_b^a except requiring $\nabla_a T_b^a = 0$. In the context of cosmology, this allows you to accommodate any evolutionary history for the universe with a suitable choice for the equation $p = p(\rho)$. The fact that we will be able to probe the matter sector only up to a finite energy scale in the lab at any given time, while the energy scale close to the big bang can be arbitrarily high (if $p > 0$, $\rho > 0$), makes the early evolution of the universe under-determined both in principle and in practice.
2. Observations suggest that the evolution of our universe is well approximated by the differential equation Eq. (13) containing three constant parameters $\rho_{\text{inf}}, \rho_{\text{eq}}, \rho_{\Lambda}$ with two – apparently unconnected – epochs of accelerated expansion. These three parameters, which constitute the signature of our universe, do not seem to have any conceptual or numerical relationship. In other words, our universe is built using three unrelated, ad-hoc numbers.
3. It is possible to construct a rather strange combination of these three densities and define a quantity I (see Eq. (21)) which has the numerical value 4π to the precision of 1 part in 1000. This (four) “pi in the sky” demands an explanation, which is difficult to conceive of within the context of the conventional approach because the matter sector is completely unconstrained. (“Everything is allowed in cosmology.”)
4. The cosmos, at very large scales, provides us with an absolute frame of rest (in which the CMB is isotropic) and an absolute time coordinate (in the form of the temperature of the CMB). You can measure your absolute motion using the dipole anisotropy of the CMB, which acts like a cosmic aether. However, we see no experimental trace of the existence of such an absolute standard of rest at sub-cosmic scales (if we do not use the CMB). In other words, sub-cosmic scale physics appears to be invariant under a much larger group (viz., the general coordinate transformation group) than the very large scales which define a cosmic aether and a preferred coordinate system. This enhancement of symmetry at small scales vis-a-vis the largest scales is intriguing.
5. Cosmic evolution introduces an arrow of time even though Einstein's equations, like the rest of physics, are invariant under time reversal. (This arrow of time arises through dynamics rather than through statistical coarse graining, etc.) While the unstable mode, corresponding to the expansion factor $a(t)$, might have something to do with this, the details are still uncertain. In particular, this explanation requires the effective Hamiltonian describing the cosmos to be unbounded; it is not clear how such a feature could be embedded in a quantum gravitational model without affecting small scale physics. Once again, we see a conceptual tension between the description of sub-cosmic scale physics and the dynamics of the large-scale universe.

- The universe is probably the only system known to us which made a spontaneous transition from quantum mechanical behaviour to classical behaviour – i.e. treated as a dynamical system, it became more and more classical as time went on. Systems with bounded Hamiltonians cannot do this. This suggests a relation between spontaneous transition to classicality and the arrow of time, possibly again through the “wrong sign” for the kinetic energy associated with the expansion factor. While this could very well be the reason, the details remain to be worked out.

I believe the conundrums described above provide sufficient motivation to look for an alternative paradigm to describe the cosmos. In this last part, I will outline such a paradigm and how it addresses at least one of the crucial issues, viz., the problem of the cosmological constant.

The conventional approach begins by assuming the validity of GR to describe the evolution of spacetime and then obtain a specific solution to the field equation to describe the evolution of the large-scale universe. There is, however, considerable evidence to suggest that the field equations of gravity themselves have only the conceptual status similar to the equations describing an elastic solid or a fluid (see, e.g., [3,16,21]). In the alternative perspective which emphasizes this feature, gravity is the thermodynamical limit of the statistical mechanics of the underlying spacetime degrees of freedom (the ‘atoms of space’). The field equations are obtained from a thermodynamic variational principle which is similar to extremizing a thermodynamic potential to obtain the equilibrium state of the normal matter. The validity of the GR field equations then correspond to a maximum entropy configuration of the atoms of space.

The evolution of spacetime itself can be described in a *purely thermodynamic language* in terms of suitably defined degrees of freedom in the bulk and boundary of a 3-volume. It can be shown [16] that the evolution of geometry, interpreted as the heating and cooling of null surfaces, is described by the equation:

$$\int_{\mathcal{V}} \frac{d^3x}{8\pi L_p^2} \sqrt{h} u_a g^{ij} \mathcal{E}_\xi p_{ij}^a = \epsilon \frac{1}{2} k_B T_{\text{avg}} (N_{\text{sur}} - N_{\text{bulk}}) \tag{26}$$

where

$$N_{\text{sur}} \equiv \int_{\partial\mathcal{V}} \frac{\sqrt{\sigma} d^2x}{L_p^2}; \quad N_{\text{bulk}} \equiv \frac{|E|}{(1/2)k_B T_{\text{avg}}} \tag{27}$$

are the degrees of freedom in the surface $\partial\mathcal{V}$ and bulk \mathcal{V} of a 3-dimensional region and T_{avg} is the average Davies–Unruh temperature [17,18] of the boundary. The h_{ab} is the induced metric on the $t = \text{constant}$ surface, $p_{bc}^a \equiv -\Gamma_{bc}^a + \frac{1}{2}(\Gamma_{bd}^d \delta_c^a + \Gamma_{cd}^d \delta_b^a)$, and $\xi^a = Nu^a$ is the proper-time evolution vector corresponding to observers moving with four-velocity $u_a = -N\nabla_a t$. The factor $\epsilon = \pm 1$ ensures the correct result for either sign of the Komar energy E . The time evolution of the metric in a region (described by the left hand side of Eq. (26)), can be interpreted [19] as the heating/cooling of the spacetime and arises because $N_{\text{sur}} \neq N_{\text{bulk}}$. In any *static* spacetime [20], on the other hand, $\mathcal{E}_\xi(\dots) = 0$, leading to “holographic equipartition”: $N_{\text{sur}} = N_{\text{bulk}}$.

Equation (26) translates the gravitational dynamics into the thermal evolution of the spacetime. The validity of Eq. (26) for all observers (i.e. foliations) ensures the validity of Einstein’s equations. I stress that, even though Eq. (26) describes a time evolution, it is obtained from an extremum condition for a thermodynamic variational principle and represents the thermodynamic equilibrium between matter degrees of freedom and microscopic degrees of freedom of the spacetime.

In the specific context of cosmology, one can write a similar but simpler equation of the form [15]:

$$\frac{dV_H}{dt} = L_p^2 (N_{\text{sur}} - N_{\text{bulk}}) \tag{28}$$

where $V_H = (4\pi/3)H^{-3}$ is the volume of the Hubble sphere, $N_{\text{sur}} = A_H/(L_p^2) = 4\pi H^{-2}/L_p^2$ is the number of microscopic degrees of freedom of the spacetime on the Hubble sphere, $N_{\text{bulk}} = -E/[(1/2)k_B T]$ is the equipartition value for the bulk degrees of freedom corresponding to the Komar energy E contained in the Hubble sphere, and $T = (H/2\pi)$ is the Hubble temperature. (I have assumed $E < 0$ to describe the current accelerated phase of the universe; otherwise one needs to flip suitable signs to keep $N_{\text{bulk}} > 0$.) This equation is equivalent to the second equation in Eq. (9). In fact one can also rewrite the first equation in Eq. (9) in thermodynamic language as an energy balance relation:

$$\rho V_H = TS \tag{29}$$

where $S = A_H/(4L_p^2) = \pi H^{-2}/L_p^2$ is the entropy associated with area of the Hubble sphere making TS in the right hand side the heat energy of the boundary surface. This equation tells you that the total energy within the Hubble sphere is equal to the heat energy of the boundary surface.

But if this is indeed the correct description of the microscopic degrees of freedom of the spacetime, then the field equations of GR, representing some kind of thermodynamic equilibrium between matter and the microscopic degrees of freedom of the spacetime, cannot be universally valid. Recall that, in the case of standard fluid mechanics, we may have to abandon the thermodynamic description in two different contexts. First, if we probe the fluid at scales comparable to

the mean free path, you need to take into account the discreteness of molecules, etc., and the fluid description breaks down. Second, a fluid simply might not have reached local thermodynamic equilibrium at the scales (which can be large compared to the mean free path) we are interested in. In the first case, the ‘fluid’ description itself breaks down; in the second case we do have a continuum description of the fluid, but it needs to be studied using non-equilibrium kinetic theory.

Something analogous happens in the description of gravity. The microscopic degrees of freedom of the spacetime could have reached the maximum entropy configuration at sub-cosmic scales, making the standard field equations of gravity valid at these scales (say at scales $10^6 L_p \lesssim x \lesssim H^{-1}$). Equation (26) (which is identical to $G_b^a = \kappa T_b^a$ but expressed in a thermodynamic language) holds at these scales. For scales close to L_p , the discrete nature of spacetime has to be taken into account and this is similar to probing a fluid at scales comparable to the mean free path; we do not yet know how to do this, which is the usual problem of quantum gravity. But it is also possible that the microscopic degrees of freedom of the spacetime have not reached the maximum entropy configuration at very large scales comparable to the horizon scale (which is much larger than the scale of the Hubble radius in the RD and MD phases). At these scales we again expect Eq. (28) and Eq. (29) to be modified because the microscopic degrees of freedom of the spacetime are not in the maximum entropy configuration. This is similar to the situation in non-equilibrium thermodynamics for normal fluids. In such a description, the symmetry of Einstein’s equations, viz. general covariance emerges when the microscopic degrees of freedom of the spacetime reach the maximum entropy configuration at the intermediate scales. At very large scales, this ‘equilibrium’ has not yet been achieved and the universe, at very large scales, picks out a cosmic frame of rest. (Of course, we also do not know how to introduce the concept of general covariance in a meaningful way close to Planck scales; but that is a different – and more well-known – story.)²⁰

So what could be the additional ingredient we need to introduce into the standard GR? I believe [21] this has to do with the concept of information stored in the spacetime and its accessibility by different observers.

A key feature of gravity is its ability to control the amount of information accessible to any given observer. A well-known example of this idea arises in the physics of black holes. More generally, the lack of access to spacetime regions leads to a configurational entropy related to the microscopic degrees of freedom of the spacetime. Over decades, we have come to realize [13] that information is a physical entity and that anything which affects the flow and accessibility of information will have direct physical significance. One consequence of such a paradigm is that matter and geometry will be more closely tied together (through the information content) than in the conventional approach. You should not be able to build an ad-hoc universe with randomly chosen values for the three densities $\rho_{\text{inf}}, \rho_{\text{eq}}, \rho_{\Lambda}$. We would expect, for example, a relationship connecting the late time accelerated expansion with the early inflationary phase through the information content of spacetime.

It turns out that, by applying this idea to the cosmos in a specific manner, you can solve the deepest mystery about our universe, viz., the small numerical value ($\Lambda L_p^2 \approx 10^{-122}$) of the cosmological constant, Λ . The I in Eq. (21) measures (in a way defined more precisely below) the amount of information accessible to an eternal observer in our universe. If $\rho_{\Lambda} = 0$, such an observer can observe all of spacetime and can acquire an infinite amount of information. But when $\rho_{\Lambda} \neq 0$, the information accessible to the observer is finite and is related to ρ_{Λ} through Eq. (21). So, if you have an independent way of fixing I , then you can use Eq. (21) to express ρ_{Λ} in terms of the other two densities. The value of I is fixed by the following fact: It turns out that, the spacetime becomes effectively 2-dimensionally close to Planck scales, irrespective of the dimension exhibited by the spacetime at large scales [14]. This, in turn, implies that the basic unit of information stored in the microscopic degrees of freedom of the spacetime is given by $A/L_p^2 = 4\pi$ where $A = 4\pi L_p^2$ is the area of the 2-sphere with radius L_p . This unit of quantum gravitational information is what appears in Eq. (22) and allows us to determine the numerical value of the cosmological constant. Let me now fill in some details of this result.²¹

8. Cosmic Information and the cosmological constant

Let us begin by recalling how the existence of a non-zero cosmological constant prevents an eternal observer O (i.e. an observer whose world line extends to $t \rightarrow \infty$ and who makes observations at very late times) from acquiring information from the far reaches of our universe. Let $x(a_2, a_1)$ be the comoving distance traveled by light between the epochs $a = a_1$ and $a = a_2$ with $a_2 > a_1$ in the standard Friedmann model with expansion factor $a(t)$. This is given by:

$$x(a_2, a_1) = \int_{t_1}^{t_2} \frac{dt}{a(t)} = \int_{a_1}^{a_2} \frac{da}{a^2 H(a)} \tag{30}$$

²⁰ A very figurative analogy is provided by a large chunk of ice containing a point source of heat inside. The heat melts ice around it creating a region containing water, which expands, maintaining thermodynamic equilibrium. The degrees of freedom in the form of water have a higher degree of symmetry (rotational invariance), compared to the degrees of freedom locked up in the ice-lattice. In the case of the universe, the expansion leads to the emergence of space [15] with the microscopic degrees of freedom of the spacetime reaching the maximum entropy configuration. This region exhibits a higher degree of symmetry (the general covariance of Eq. (26), which as I said, is just $G_b^a = \kappa T_b^a$), compared to the larger scales of the universe.

²¹ These results are based on unpublished work done in collaboration with H. Padmanabhan.

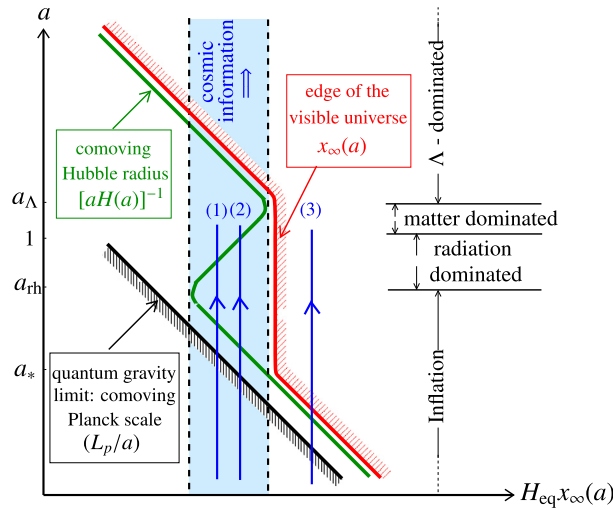


Fig. 1. The different length scales in a universe with an inflationary phase and a non-zero cosmological constant. The red curve is the maximum comoving size of a region from which signals can reach an observer at very late times. The information in the shaded region to the right of the red curve is not accessible to an observer even if she waits till eternity. The green curve is the comoving Hubble radius. The slanted black curve is the comoving scale corresponding to the Planck length and the shaded region below this black line is dominated by quantum gravitational effects. The vertical lines are different proper length scales which cross the Hubble radius and the horizon. The two lines, marked 1 and 2, leave the Hubble radius during inflation and re-enter it during the radiation/matter-dominated epoch. These remain within the horizon of the observer at the origin (red curve) and are visible to her at, say, $a = a_{th}$. The line marked 3 corresponds to a proper length scale which goes out of, not only the Hubble radius, but also the horizon and thus become inaccessible to the observer at, say, $a = a_{th}$. So the relevant part of the cosmic information is contained within the blue vertical band, between the two vertical lines which are tangential to the comoving Hubble radius at its turning points. The arrows at the top of the band denote the direction of flow of the cosmic information.

So the comoving $[x_{\infty}(a)]$ and proper $[r_{\infty}(a)]$ sizes of the regions of the universe at an epoch a , from which O can receive signals at very late times, are given by [12]:

$$x(\infty, a) \equiv x_{\infty}(a) = \int_a^{\infty} \frac{d\bar{a}}{\bar{a}^2 H(\bar{a})}; \quad r_{\infty}(a) = ax_{\infty}(a) \equiv a \int_a^{\infty} \frac{d\bar{a}}{\bar{a}^2 H(\bar{a})} \tag{31}$$

The behaviour of $x_{\infty}(a)$ and $r_{\infty}(a)$ depend crucially on whether the cosmological constant is zero or non-zero. If $\Lambda = 0$ and the universe is dominated by, say, matter at late times, then $H(a) \propto a^{-n}$, with $n > 1$ at late times. Then, both these integrals will diverge at the upper limit as $t \rightarrow \infty$, irrespective of the behaviour of the universe at earlier epochs. So, in a universe with $\Lambda = 0$, information from the infinite expanse of space will be accessible to the eternal observer at late times; there is no blocking of information.

If $\Lambda \neq 0$ and $H(a) \rightarrow H_{\Lambda}$ = constant at late times, then the situation is different. In that case, both the integrals in Eq. (31) are finite at the upper limit and an eternal observer have only access information from a finite region of space at an epoch a , irrespective of how long she waits. The amount of accessible Cosmic Information (which I call ‘Cosmln’) is now reduced from an infinite amount to a finite value, say I_c , as a direct consequence of the fact that $\Lambda \neq 0$. It is, therefore, reasonable to expect that the numerical value of Λ should be related to I_c with I_c decreasing with increasing Λ . I will now derive this relation.

Consider a universe (like ours) with three distinct phases of evolution: (i) At very early times, the universe was in a state of inflation with $H(a) = H_{inf} = \text{constant}$. (ii) At some point $a = a_{th}$, the inflation ends; the universe reheats and becomes radiation-dominated. The radiation-dominated phase goes on till $a = a_{eq}$ which is the epoch of radiation–matter equality. For $a_{eq} \lesssim a \lesssim a_{\Lambda}$, the universe is matter-dominated. (iii) Later on, at some point, $a \gtrsim a_{\Lambda}$, the cosmological constant starts to dominate the expansion of the universe. We will rescale the expansion factor such that $a_{eq} = 1$, and use Eq. (13) to describe the evolution of the universe. I will also assume instant reheating at $a = a_{th}$ for simplicity.²²

The geometrical features related to $x_{\infty}(a)$ and other relevant length scales are depicted in Fig. 1. The green curve is the comoving Hubble radius $d_H(a)/a \equiv 1/aH(a)$. It decreases (as $1/a$) during the inflationary phase, reaching a minimum at $a = a_{th}$; it then increases as a^2 in the radiation-dominated phase and as $a^{3/2}$ in the matter-dominated phase, attaining a maximum around $a \approx a_{\Lambda}$; in the Λ -dominated phase, it again decreases as $1/a$. The red curve denotes $x_{\infty}(a)$ obtained by evaluating the integral in Eq. (31) and represents the visibility limit. During the Λ -dominated phase, this curve closely tracks the comoving Hubble radius ($x \approx a_{\Lambda}^{3/2}/a$) but soon becomes vertical to a high degree of approximation. During the

²² For our universe, observations tell us that $a_{th} \approx 7.4 \times 10^{-25}$, $a_{\Lambda} \approx 2.8 \times 10^3$, if we set $a_{eq} = 1$.

matter- and radiation-dominated phases (i.e. during $a_\Lambda \gtrsim a \gtrsim a_{\text{rh}}$) the $x_\infty(a)$ is approximately constant – varying just by a factor 3 (from $\sim a_\Lambda^{1/2}$ at $a = a_\Lambda$ to $\sim 3a_\Lambda^{1/2}$ at $a = a_{\text{rh}}$) when a varies by a factor $\sim 10^{28}$. During the inflationary phase, $x_\infty(a)$ again tracks d_H/a asymptotically, with an approximate behaviour $x_\infty(a) \approx [3a_\Lambda^{1/2} - a_{\text{rh}}] + a_{\text{rh}}^2/a$. As I mentioned, the region of space from which an eternal observer can acquire information is finite for all finite a if the cosmological constant is non-zero.

Our next task is to quantify the amount of cosmic information that is actually accessible to the eternal observer. To do this, recall that a comoving scale $x = \text{constant}$ is associated with a proper length scale $r = a(t)x$. The proper length scales (e.g., the wavelengths of modes) will get stretched exponentially during the inflation, and will exit the Hubble radius. (One can think of $\lambda(t) = a(t)\lambda_0$ as a physical length being stretched by expansion. You can equally well think of it as the proper length associated with a geodesic labeled by a comoving scale $x_0 = \lambda_0$. I will use the former terminology since it is more familiar to cosmologists.) After remaining outside the Hubble radius for some time, some of them will re-enter the Hubble radius during the matter/radiation-dominated epoch. (Two such modes are shown as (1) and (2) in Fig. 1.) In contrast, the mode marked as (3) will leave the Hubble radius but will *never* re-enter it. Such modes cross not only the Hubble radius but also the horizon (red line) and become invisible to the eternal observer at, say, the epoch of reheating $a = a_{\text{rh}}$. So the modes relevant to us are confined to those between the two dotted horizontal lines which are tangential to the Hubble radius at its turning points. The total number of such modes (or, equivalently, geodesics) is a measure of the information content I_c .

Let us compute how many modes cross the Hubble radius during the inflationary phase between $a = a_*$ and $a = a_{\text{rh}}$. Since the deSitter space is invariant under time translation, the *rate* at which the modes leave the Hubble radius will be a constant. So the number of modes $I(a_2, a_1)$ which cross the Hubble radius during an interval $a_1 < a < a_2$ will be proportional to $H(t_2 - t_1)$. So the total number of modes which cross the Hubble radius during the inflationary epoch will be proportional to $N_e \equiv H\Delta t$, where Δt is the relevant duration in the inflationary phase. (Here N_e denotes the number of e-foldings in the interval Δt .) Therefore, the CosmIn is given by:

$$I_c \propto N_e \tag{32}$$

and all we need to do is to fix the proportionality constant using a suitable measure. This measure can be introduced as follows: The number of modes dN within the comoving Hubble volume $V_H(a) = (4\pi/3)(aH)^{-3}$ with wave numbers in the range d^3k is given by $dN = V_H(a)d^3k/(2\pi)^3 = V_H(a)dV_k/(2\pi)^3$ where $dV_k = 4\pi k^2 dk$. A mode with the comoving wave number k will leave the Hubble radius when $k = k(a) \equiv aH(a)$. So the modes with wave numbers within the range $(k, k + dk)$, where $dk = [d(aH)/da] da$, will exit the Hubble radius in an interval $(a, a + da)$. Therefore, the number of modes that cross the Hubble radius during the interval $a_1 < a < a_2$ is given by

$$N(a_2, a_1) = \int_{a_1}^{a_2} \frac{V_H(a)}{(2\pi)^3} \frac{dV_k[k(a)]}{da} da = \frac{2}{3\pi} \ln \left(\frac{a_2 H_2}{a_1 H_1} \right) \tag{33}$$

(This result, of course, is applicable for any $a(t)$; not just in the inflationary phase.) During inflation, when $a(t) \propto \exp(H_{\text{inf}}t)$, this expression reduces to $(2/3\pi) \ln(a_2/a_1)$, which tells us that the proportionality constant in Eq. (32) is $(2/3\pi)$. We thus find the value of CosmIn to be:

$$I_c = \frac{2}{3\pi} N_e = \frac{2}{3\pi} \ln \left(\frac{a_{\text{rh}}}{a_*} \right) \tag{34}$$

We can relate the ratio a_{rh}/a_* to the three densities $\rho_\Lambda, \rho_{\text{eq}}$ and ρ_{inf} which will give $a_{\text{rh}}/a_* \propto (\rho_{\text{inf}}/\rho_{\text{eq}})^{1/4} (\rho_{\text{eq}}/\rho_\Lambda)^{1/6}$. To determine the proportionality constant, we need to evaluate the turning point of the $d_H(a)/a$ curve near $a = a_\Lambda$ which requires solving a cubic equation. Doing this (for details, see [2]), we find that the proportionality constant has the value $(4/27)^{1/6} = 2^{1/3}/3^{1/2}$. Substituting into Eq. (34), we achieve our first goal, viz. relating the *non-zero* value of the cosmological constant to the *finite* amount of cosmic information accessible to an eternal observer (I_c):

$$\rho_\Lambda = \frac{4}{27} \frac{\rho_{\text{inf}}^{3/2}}{\rho_{\text{eq}}^{1/2}} \exp(-9\pi I_c) \tag{35}$$

As expected, the cosmological constant vanishes when the information content is infinite ($I_c \rightarrow \infty$) and vice-versa.

Equation (35) will determine ρ_Λ in terms of ρ_{inf} and ρ_{eq} provided we know the value of CosmIn from some other physical consideration. (I share the hope that ρ_{inf} and ρ_{eq} will be eventually determined from high energy physics in terms of the inflationary model and the dark matter content of the universe.) To do this, notice that the modes which exit the Hubble radius during the inflationary epoch correspond to sub-Planckian scales in the early part of inflation. In Fig. 1, the black line indicates the *comoving* length scale L_p/a corresponding to the Planck length L_p . The region below this line corresponds to *proper* length scales smaller than the Planck length, and will be dominated by quantum gravitational effects. The modes containing the cosmic information cross the comoving Planck length during the earlier stages of evolution and hence carry the imprint of quantum gravitational effects. So we expect I_c to be determined by quantum gravitational considerations.

It can be shown that the spacetime becomes effectively two-dimensional near Planck scales [14] and hence the unit I_{QG} of quantum gravitational information content of spacetime is given by the degrees of freedom contained in a 2-sphere of radius L_P , viz., $I_{\text{QG}} = 4\pi L_P^2/L_P^2 = 4\pi$. Therefore the natural numerical value for I_c can be taken to be:

$$I_c = I_{\text{QG}} = 4\pi \quad (36)$$

Substituting into Eq. (35), we get a remarkable formula for the cosmological constant

$$\rho_\Lambda = \frac{4}{27} \frac{\rho_{\text{inf}}^{3/2}}{\rho_{\text{eq}}^{1/2}} \exp(-36\pi^2) \quad (37)$$

If we use the typical values $\rho_{\text{inf}} = (1.2 \times 10^{15} \text{ GeV})^4$, $\rho_{\text{eq}} = (0.86 \text{ eV})^4$, we get $\rho_\Lambda = (2.2 \times 10^{-3} \text{ eV})^4$ which agrees well with the observed value! That is, the idea that the cosmic information content accessible to an eternal observer, I_c , is equal to the basic quantum gravitational unit of information $I_{\text{QG}} = 4\pi$, determines the numerical value of the cosmological constant correctly. Let me conclude with a few comments on this result:

- the relation $I_c = I_{\text{QG}} = 4\pi$, also determines the relevant number of e -foldings in the inflationary epoch which carries the cosmic information. This is given by $N_e = (3\pi/2)I_c = 6\pi^2 \approx 59$, which provides an adequate amount of inflation;
- Equation (35) can be inverted to express the cosmic information content I_c in terms of the three densities. As I mentioned earlier, if we use the values for ρ_Λ and ρ_{eq} known from observations and take $\rho_{\text{inf}} = (10^{15} \text{ GeV})^4$, we find that:

$$I_c = \frac{1}{9\pi} \ln \left(\frac{4}{27} \frac{\rho_{\text{inf}}^{3/2}}{\rho_\Lambda \rho_{\text{eq}}^{1/2}} \right) \approx 4\pi \left(1 \pm \mathcal{O}(10^{-3}) \right) \quad (38)$$

That is, the current observations show that the CosmIn indeed has a value 4π to the precision of one part in a thousand! Because of the logarithmic dependence on the cosmic parameters in Eq. (38), this result is fairly stable and renders a purely observational support for the claim $I_c = I_{\text{QG}} = 4\pi$.

- Theoretically, one would like to determine the value of ρ_Λ in terms of other parameters. Observationally, we can determine the values of ρ_{eq} and ρ_Λ very well today but have no direct handle on ρ_{inf} . Using Eq. (37), we can predict the value of ρ_{inf} in terms of the well-determined parameters ρ_{eq} and ρ_Λ . We then get $\rho_{\text{inf}}^{1/4} = 1.2 \times 10^{15} \text{ GeV}$, which is again a remarkable result.²³ This is probably the *only* model with quantum gravitational inputs which leads to a falsifiable prediction.

I believe the result obtained above is just a tip of the iceberg. It tells you that microscopic quantum gravitational physics can leave its trace in cosmic scales and, in fact, probably can be tested only in cosmology. Further exploration will require a systematic computation of corrections to Eq. (28) and Eq. (29) in a context similar to non-equilibrium thermodynamics.

Acknowledgements

My research is partially supported by the J. C. Bose Research grant of DST, India. I thank Sumanta Chakraborty, Shyam Date, Tamara Davis, Dawood Kothawala, Kinjalk Lochan and Aseem Paranjape for comments on an earlier draft.

References

- [1] T. Padmanabhan, Cosmological Constants – Part 1, Lecture in the mini-series *Cosmology and the Constants of Nature* from the *Philosophy of Cosmology* project; a University of Oxford and Cambridge Collaboration 2014, Available at: <https://www.youtube.com/watch?v=Pcd911qSTBk>.
- [2] T. Padmanabhan, H. Padmanabhan, *Int. J. Mod. Phys. D* 23 (2014) 1430011, arXiv:1404.2284.
- [3] T. Padmanabhan, *Gen. Rel. Gravit.* 42 (2010) 2743–2750, arXiv:1001.3380.
- [4] See, e.g., T. Padmanabhan, *Gravitation: foundations and frontiers*, in: CUP, 2010, Eqn. (7.52) in p. 304.
- [5] A. Mazumdar, The origin of dark matter, matter–anti-matter asymmetry, and inflation, arXiv:1106.5408.
- [6] One source which mentions “expansion problem” explicitly is J.A. Peacock, *Cosmological Physics*, Cambridge University Press, 1999, 324 p.
- [7] B.S. DeWitt, Has one of the earliest explorations in this direction, *Phys. Rev.* 160 (1967) 1113.
- [8] K. Lochan, K. Parattu, T. Padmanabhan, *Gen. Rel. Gravit.* 47 (2015) 1841, arXiv:1404.2605.
- [9] G. Mahajan, T. Padmanabhan, *Gen. Rel. Gravit.* 40 (2008) 661–708, arXiv:0708.1233; *Gen. Rel. Gravit.* 40 (2008) 709–747, arXiv:0708.1237.
- [10] S. Singh, S. Kumar Modak, T. Padmanabhan, *Phys. Rev. D* 88 (2013) 125020, arXiv:1308.4976.
- [11] T. Padmanabhan, *Phys. Rev. D* 66 (2002) 021301, arXiv:hep-th/0204150.
- [12] These expressions were studied in the past, in the context of holographic bounds. See, e.g., T. Banks, W. Fischler, astro-ph/0307459, and references therein.

²³ In the calculation leading to Eq. (38), I assumed that the reheating is instantaneous; ambiguities in the reheating dynamics can change this result by a factor of about 5, leading to the prediction $\rho_{\text{inf}}^{1/4} \approx (1-5) \times 10^{15} \text{ GeV}$.

- [13] R. Landauer, Information is physical, in: Proc. Workshop on Physics and Computation PhysComp '92, IEEE Comp. Sci. Press, Los Alamitos, 1993, pp. 1–4.
- [14] T. Padmanabhan, S. Chakraborty, D. Kothawala, Gen. Rel. Gravit. 48 (2016) 55, arXiv:1507.05669;
For a sample of similar earlier results, see e.g., S. Carlip, R. Mosna, J. Pitelli, Phys. Rev. Lett. 107 (2011) 021303, arXiv:1103.5993;
AIP Conf. Proc. 1196 (2009) 72, arXiv:0909.3329;
G. Calcagni, Phys. Rev. Lett. 104 (2010) 251301, arXiv:0912.3142;
J. Ambjorn, J. Jurkiewicz, R. Loll, Phys. Rev. Lett. 95 (2005) 171301, arXiv:hep-th/0505113;
L. Modesto, P. Nicolini, Phys. Rev. D 81 (2010) 104040, arXiv:0912.0220;
G. Calcagni, D. Oriti, J. Thurigen, Class. Quant. Grav. 30 (2013) 125006, arXiv:1208.0354;
V. Husain, S.S. Seahra, E.J. Webster, Phys. Rev. D 88 (2013) 024014, arXiv:1305.2814.
- [15] T. Padmanabhan, Res. Astron. Astrophys. 12 (2012) 891, arXiv:1207.0505.
- [16] T. Padmanabhan, Gen. Relat. Gravit. 46 (2014) 1673, arXiv:1312.3253.
- [17] P.C.W. Davies, J. Phys. A 8 (609) (1975).
- [18] W.G. Unruh, Phys. Rev. D 14 (1976) 870.
- [19] K. Parattu, B.R. Majhi, T. Padmanabhan, Phys. Rev. D 87 (2013) 124011, arXiv:1303.1535.
- [20] T. Padmanabhan, Phys. Rev. D 81 (2010) 124040, arXiv:1003.5665.
- [21] T. Padmanabhan, Exploring the nature of gravity, arXiv:1602.01474;
Int. J. Mod. Phys. D 25 (2016) 1630020, arXiv:1603.08658.