A path towards quantum gravity phenomenology

Daniel Sudarsky
Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, A. Postal 70-543, México D.F. 04510, México
E-mail: sudarsky@nucleares.unam.mx

Abstract. In this article I review the arguments that were recently put forward by a couple of colleagues and myself [1] about the shortcomings of the standard explanations of the quantum origin of cosmic structure in the inflationary scenario. The crux of the problem is that the inhomogeneity and anisotropy of our universe seem to emerge from an exactly homogeneous and isotropic initial state through processes that do not break those symmetries. We have argued that some novel aspect of physics must be called upon to able to fully address the problem. The approach we followed is inspired on Penrose's proposals that quantum gravity might lead to a real, dynamical collapse of the wave function. This point of view leads us to consider the possibility of extracting information about the nature of quantum gravity from the observations of the very early universe.

1. Introduction

One of the most important advances in physical cosmology are the precision measurements of the anisotropies in the CMB[2] together with their explanation within the context of the inflationary scenarios[3]. However after the first glances at the explanations one notices something odd: The description of the relevant part of our Universe starts with an initial set of conditions which are totally homogeneous and isotropic both in the background space-time and in the quantum state that is supposed to describe the “fluctuations”, and it is quite easy to see that the subsequent evolution through dynamics that do not break these symmetries can only lead to an equally homogeneous and anisotropic universe. The arguments normally used in order to deal with this issue, are phrased in terms of “the quantum to classical transition”, without focussing on the required concomitant breakdown of homogeneity and isotropy in the state. The various alternatives have been critically discussed in [1].

Moreover, if we were to think in terms of first principles, we would start by acknowledging that the correct description of the problem at hand would involve a full theory of quantum gravity coupled to a theory of all the matter quantum fields, and that there, the issue would be whether we start with a quantum state that is homogeneous and isotropic or not?. Even if these notions do not make sense within that level of description, a fair question is whether or not, the inhomogeneities and anisotropies we are interested on, can be traced to aspects of the description that have no contra-part in the approximation we are using. Recall that such description involves the separation of background vs. fluctuations and thus must be viewed only as an approximation, that allows us to separate the nonlinearities in the system—as well as those aspects that are inherent to quantum gravity—from the linear part of problem represented by the fluctuations, which are treated in terms of linear quantum field theory. If one chose to ignore the
problem and view it as something inherent to such approximation, one could not argue that one has an understanding of the origin of the CMB spectrum. Furthermore, even without addressing the symmetries issue one should worry in principle about the justification of a treatment of the gravitational degrees of freedom in terms of what amounts to a standard quantum field theoretical approach, which one knows is in principle inadequate, even if the problems of such treatment of quantum gravity do not seem to surface directly in our calculations.

It is of course not at all clear that the problem of the transition from a symmetric state (i.e. with homogeneity and isometry) to a state that is not symmetric, should be related to quantum gravity, but since that is the only sphere of fundamental physics for which we have so far failed to find a satisfactory conceptual understanding it seems worthwhile to consider their possible connection. Penrose’s [4], ideas regarding the fundamental changes, that he argues are needed in quantum mechanics and their connection to quantum gravity, are taken as further justification for this views. In the treatment developed in [1] the approach is to bring up as aspect that we view as part of the quantum gravity realm, to the forefront in order to modify – in a minimalistic way– the semiclassical treatment, to deal with the unsatisfactory part of thereof.

2. The quantum origin of the seeds of cosmic structure

We must warn the reader that most colleagues who have been working in this field for a long time take the view, that there is no problem at all in the transition from a homogeneous and isotropic early state of the universe, a late state that is neither. It is however a fact that the arguments invoked in this regard tend to differ from one inflationary cosmologist to another [5]. Very few do acknowledge that there seems to be something unclear at this point [6]. One can see that the situation at hand, is quite different from any other situation usually treated using quantum mechanics. Note for instance that, while in analyzing ordinary situations, quantum mechanics offers us, at least one self consistent assignment at each time of a state of the Hilbert space to our physical system (we are of course thinking of the Schroedinger picture), the same is not true for the standard analysis of the current problem. In trying to the consider such assignment – of a state at each time – when presented with any of the proposed justifications offered to deal with the issue one must be ready to accept one of the following: i) our universe was not really in that symmetric state (corresponding to the vacuum of the quantum field), ii) our universe is still described by a symmetric state, iii) at least at some points in the past the description of the state of our universe could not be done within quantum mechanics, iv) quantum mechanics does not correspond to the full description of a system at all times, or v) our own observations of the universe mark the transition from a symmetric to an asymmetric state. None of these represent a satisfactory alternative, in particular, if we want to claim, that we understand the evolution of our universe and its structure – including ourselves – as the result of the fluctuations of quantum origin in the very early stages of our cosmology. If we want to identify the situation at hand, say, with the well known Schroedinger cat problem, we find that not only we play the role pf the cat, but moreover, that the cat plays a role in giving raise to the conditions that a prerequisite for its own coming into being. For a more thorough discussion we refer the reader to [1].

The interesting part of this situation is that one is forced to call upon some novel physical process to fill in the shortcomings of the justification of the steps that are used to take us from that early and symmetric state, to the asymmetric state of our universe today, or that which is studied through the CMB. In [1] we have shown that the requirement that one should obtain results compatible with current observations is already sufficient to restrict in important ways some specific aspects of these novel physics. Thus, when considering that the origin of such new physics is associated with some aspects of quantum gravity, one is already in a position of setting phenomenological constraints on at least this aspect of the quantum theory of gravitation.
Next we give a short description of this analysis. We will be using units where \(c = 1\) but will keep \(\hbar\) and \(G\) explicitly throughout the manuscript. The starting point is as usual the action of a scalar field coupled to gravity.

\[
S = \int d^4x \sqrt{-g} \left[ -\frac{1}{16\pi G} R[g] - 1/2 \nabla_a \phi \nabla_b \phi g^{ab} - V(\phi) \right],
\]

where \(\phi\) stands for the inflaton and has units of \((\text{Mass}/\text{Length})^{1/2}\), while its potential \(V\) has units of \(\text{Mass}/(\text{Length})^3\). The coordinates \(\eta, x^i\) will have units of length but the metric \(a\) will be dimensionless. One then splits both, metric and scalar field into a spatially homogeneous "background" part and an inhomogeneous part "fluctuation", i.e. the scalar field is written \(\phi = \phi_0 + \delta \phi\), while the perturbed metric can, after appropriate gauge fixing and by focusing on the scalar perturbation be written

\[
d s^2 = a(\eta)^2 \left[ -(1 + 2\Psi) d\eta^2 + (1 - 2\Psi) \delta_{ij} dx^i dx^j \right],
\]

where \(\Psi\) represents the relevant perturbation and is called the Newtonian potential.

The equations governing the background metric (corresponding to setting \(\Psi = 0\)) and the homogeneous scalar field \(\phi_0(\eta)\) are, the scalar field equation,

\[
\ddot{\phi}_0 + 2\frac{\dot{a}}{a} \dot{\phi}_0 + a^2 \partial_\phi V[\phi] = 0,
\]

where the dot represents derivative with respect to the conformal time \(\eta\), and Friedman’s equation

\[
3 \frac{\dot{a}^2}{a^2} = 4\pi G (\dot{\phi}_0^2 + 2 a^2 V[\phi_0]).
\]

The background solution corresponds to the standard inflationary cosmology during the inflationary era has a scale factor \(a(\eta) = \frac{1}{\eta^{6/7}},\) with \(H_0^2 \approx (8\pi/3)G V\) while the scalar \(\phi_0\) field in the slow roll regime so \(\dot{\phi}_0 = -(a^3/3\dot{a}) V'\). This era is supposed to give rise to a reheating period whereby the universe is repopulated with ordinary matter fields, and then to a standard hot big bang cosmology leading up to the present era. The scale factor \(a(\eta)\) has of course a different functional form during these latter periods but we will ignore such details on the account that most of the change in the value of \(a\) occurs during the inflationary regime. The overall normalization of the scale factor will be set so \(a = 1\) at the "present cosmological time". The inflationary regime would end for a value of \(\eta = \eta_0\), negative and very small in absolute terms.

The perturbation of the scalar field leads to a perturbation of the energy momentum tensor, and thus Einstein’s equations at lowest order lead to

\[
\nabla^2 \Psi = 4\pi G \dot{\phi}_0 \delta \dot{\phi} \equiv s \delta \dot{\phi}
\]

where we introduced the abbreviation \(s = 4\pi G \dot{\phi}_0\). This will be our main equation. Next, we write the quantum theory of the field \(\delta \phi\). It is convenient to consider instead the field \(y = a \delta \phi\). We consider the field in a box of side \(L\), and decompose the real field \(y\) into plane waves

\[
y(\eta, \bar{x}) = \frac{1}{L^3} \sum \xi \left( a_{k} y_k(\eta) e^{i\bar{k} \cdot \bar{x}} + a_{k}^\dagger y_k(\eta) e^{-i\bar{k} \cdot \bar{x}} \right),
\]

where the sum is over the wave vectors \(\bar{k}\) satisfying \(k_i L = 2\pi n_i\) for \(i = 1, 2, 3\) with \(n_i\) integers.
We rewrite the field and momentum operators as
\[ \hat{y}(\eta, \vec{x}) = \frac{1}{\sqrt{2}} \sum_k e^{i\vec{k} \cdot \vec{x}} \hat{\gamma}_k(\eta), \quad \hat{\pi}(\eta, \vec{x}) = \frac{1}{\sqrt{2}} \sum_k e^{i\vec{k} \cdot \vec{x}} \hat{\pi}_k(\eta), \] (7)
where \( \hat{\gamma}_k(\eta) \equiv y_k(\eta)\hat{a}_k + \hat{g}_k(\eta)\hat{a}_k\) and \( \hat{\pi}_k(\eta) \equiv g_k(\eta)\hat{a}_k + \hat{g}_k(\eta)\hat{a}_k\) with
\[ y_k^{(\pm)}(\eta) = \frac{1}{\sqrt{2k}} \left( 1 \pm \frac{i}{\eta k} \right) \exp(\pm i\eta k), \quad g_k^{(\pm)}(\eta) = \pm i \sqrt{\frac{k}{2}} \exp(\pm i\eta k). \] (8)

Given that we are interested in considering a kind of self induced collapse which operates in close analogy with a “measurement”, we must work with Hermitian operators. Thus we decompose both \( \hat{\gamma}_k(\eta) \) and \( \hat{\pi}_k(\eta) \) into their real and imaginary parts \( \hat{\gamma}_k(\eta) = \hat{\gamma}_k^R(\eta) + i\hat{\gamma}_k^I(\eta) \) and \( \hat{\pi}_k(\eta) = \hat{\pi}_k^R(\eta) + i\hat{\pi}_k^I(\eta) \) where
\[ \hat{\gamma}_k^{R,I}(\eta) = \frac{1}{\sqrt{2}} \left( y_k(\eta)\hat{a}_k^{R,I} + \hat{g}_k(\eta)\hat{a}_{k}^{R,I} \right), \quad \hat{\pi}_k^{R,I}(\eta) = \frac{1}{\sqrt{2}} \left( g_k(\eta)\hat{a}_k^{R,I} + \hat{g}_k(\eta)\hat{a}_{k}^{R,I} \right). \] (9)

The operators \( \hat{\gamma}_k^{R,I}(\eta) \) and \( \hat{\pi}_k^{R,I}(\eta) \) are hermitian by construction. The next step is to specify what we mean by “the collapse of the wave function”. We will do this minimalistically by giving a simple parametrization of the state of the field mode after a collapse in terms of its collapse time. Then we will follow the field evolution through collapse to the end of inflation. We will the collapse to be analogous to some sort of imprecise measurement of the operators \( \hat{\gamma}_k^{R,I}(\eta) \) and \( \hat{\pi}_k^{R,I}(\eta) \) which, are hermitian operators and thus reasonable observables. These field operators contain complete information about the field.

Let \( |\Xi\rangle \) be any state in the Fock space of \( \hat{y}\), and assign to each such state the following quantity: \( d_k^{R,I} = \langle \hat{a}_k^{R,I} \rangle_\Xi \). The expectation values of the modes of the fundamental field operators are then expressible as
\[ \langle \hat{\gamma}_k^{R,I} \rangle_\Xi = \sqrt{2} \Re(y_k d_k^{R,I}), \quad \langle \hat{\pi}_k^{R,I} \rangle_\Xi = \sqrt{2} \Re(g_k d_k^{R,I}). \] (10)

For the vacuum state \( |0\rangle \) we have of course: \( \langle \hat{\gamma}_k^{R,I} \rangle_0 = 0, \langle \hat{\pi}_k^{R,I} \rangle_0 = 0 \), while their corresponding uncertainties are
\[ (\Delta \hat{\gamma}_k^{R,I})_0^2 = (1/2)|y_k|^2(hL^3), \quad (\Delta \hat{\pi}_k^{R,I})_0^2 = (1/2)|g_k|^2(hL^3). \] (11)

We note that in the standard treatments we would have to argue at this point that the situation corresponds somehow to an anisotropic and inhomogeneous universe, while it is clear that none of the relevant quantities has any dependence on the ”spatial coordinates” \( x^i \).

**The collapse**

Now we will specify the rules according to which collapse happens. Again, at this point our criteria will be simplicity and naturalness.

What we have to describe is the state \( |\Theta\rangle \) after the collapse. We need to specify \( d_k^{R,I} = \langle \Theta | \hat{a}_k^{R,I} \rangle_\Theta \). In the vacuum state, \( \hat{\gamma}_k \) and \( \hat{\pi}_k \) individually are distributed according to Gaussian distributions centered at 0 with spread \( (\Delta \hat{\gamma}_k)_0^2 \) and \( (\Delta \hat{\pi}_k)_0^2 \) respectively. In this first collapse model, we will make the following assumption about the (distribution of) state(s) \( |\Theta\rangle \) after collapse:
\[ \langle \hat{\gamma}_k^{R,I} (\eta^k) \rangle_\Theta = x_k^{R,I} \sqrt{(\Delta \hat{\gamma}_k^{R,I})_0^2} = x_k^{R,I} |y_k(\eta^k)| \sqrt{hL^3/2}, \] (12)
\[ \langle \hat{\pi}_k^{R,I} (\eta^k) \rangle_\Theta = x_k^{R,I} \sqrt{(\Delta \hat{\pi}_k^{R,I})_0^2} = x_k^{R,I} |g_k(\eta^k)| \sqrt{hL^3/2}. \] (13)
where $x_{k,1}, x_{k,2}$ are selected randomly from within a Gaussian distribution centered at zero with spread one. From these equations we solve for $d_k^{R,l}$. Here we must emphasize that our universe, corresponds to a single realization of the random variables, and thus each of the quantities $x^{R,l}_{k,1,2}$ has a single specific value. Later, we will see how to make relatively specific predictions, despite these features.

Now recall that our general following approach the R.H.S of our basic formula Eq.(5) becomes an expectation value so that equation turns into

$$\nabla^2 \Psi = s \langle \delta \phi \rangle. \quad (14)$$

Before the collapse occurs, the expectation value on the right hand side is zero. Let us now determine what happens after the collapse: To this end, take the Fourier transform of Eq.(14) and obtain

$$\Psi_k(\eta) = \frac{-s}{k^2} \langle \delta \phi_k \rangle_\Theta = \frac{-s}{k^2} \sqrt{\hbar} k \frac{1}{2a} F(k), \quad (15)$$

with

$$F(k) = \frac{1}{2}[A_k(x_{k,1}^0 + ix_{k,1}^1) + B_k(x_{k,2}^0 + ix_{k,2}^1)], \quad (16)$$

where

$$A_k = \sqrt{\frac{1 + z_k^2}{z_k}} \sin(\Delta_k), \quad B_k = \cos(\Delta_k) + (1/z_k) \sin(\Delta_k), \quad (17)$$

and where $\Delta_k = k\eta - z_k$ with $z_k = \eta_k^2 k$.

Next we turn to the observational quantities. We will, for the most part, disregard the changes to dynamics that happen after re-heating and due to the transition to standard (radiation dominated) evolution. The quantity that is measured is the “Newtonian Potential” $\Psi$, by the understanding that they are the result of gravitational red-shift in the CMB photon frequency $\nu$ so $\frac{\partial \Psi}{\partial T} = \frac{\partial \nu}{\partial T} = \frac{\partial \nu}{\partial \nu} = \frac{\Psi}{\nu} \approx \Psi$.

The quantity that is presented as the result of observations is $\text{OB}_l = l(l + 1)(2l + 1)^{-1} \sum_m |\alpha_{lm}^{\text{obs}}|^2$. The observations indicate that (ignoring the acoustic oscillations, which is anyway an aspect that is not being considered in this work) the quantity $\text{OB}_l$ is essentially independent of $l$ and this is interpreted as a reflection of the “scale invariance” of the primordial spectrum of fluctuations.

Then, as we noted the measured quantity is the “Newtonian potential” on the surface of last scattering: $\Psi(\eta_D, \vec{x}_D)$, from where one extracts

$$a_{lm} = \int \Psi(\eta_D, \vec{x}_D) Y_{lm}^* d^2 \Omega. \quad (18)$$

To evaluate the expected value for the quantity of interest we use (15) and (15) to write

$$\Psi(\eta, \vec{x}) = \sum_k \frac{sU(k)}{k^2} \sqrt{\frac{\hbar}{L^3}} \frac{1}{2a} F(\vec{k}) e^{i k \cdot \vec{x}}, \quad (19)$$

where we have added the factor $U(k)$ to represent the aspects of the evolution of the quantity of interest associated with the physics of period from re-heating to de-coupling.

After some algebra we obtain

$$a_{lm} = s \sqrt{\frac{\hbar}{L^3}} \frac{1}{2a} \sum_k \frac{U(k) \sqrt{k}}{k^2} F(\vec{k}) 4\pi j_l((|\vec{k}| R_D) Y_{lm}(\vec{k}), \quad (20)$$
where \( \hat{k} \) indicates the direction of the vector \( \vec{k} \). It is in this expression that the justification for the use of statistics becomes clear. The quantity we are in fact considering is the result of the combined contributions of an ensemble of harmonic oscillators each one contributing with a complex number to the sum, leading to what is in effect a 2 dimensional random walk whose total displacement corresponds to the observational quantity. To proceed further we must evaluate the most likely value for such total displacement. This we do with the help of the imaginary ensemble of universes, and the identification of the most likely value with the ensemble mean value. Now we compute the expected magnitude of this quantity. After taking the continuum limit we find,

\[
|\alpha_{lm}|_{L,M}^2 = \frac{s^2h}{2\pi a^2} \int \frac{U(k)^2 C(k)}{k^4} j_l^2(||\vec{k}|)R D)k^3 dk,
\]

where

\[
C(k) \equiv 1 + (2/z_k^2) \sin(\Delta_k)^2 + (1/z_k) \sin(2\Delta_k).
\]

The last expression can be made more useful by changing the variables of integration to \( x = kR_D \) leading to

\[
|\alpha_{lm}|_{L,M}^2 = \frac{s^2h}{2\pi a^2} \int \frac{U(x/R_D)^2 C(x/R_D)}{x^4} j_l^2(x)x^3 dx,
\]

which in the exponential expansion regime where \( \mu \) vanishes and in the limit \( z_k \to -\infty \) where \( C = 1 \), and taking for simplicity \( U(k) = U_0 \) to be independent of \( k \), (neglecting for instance the physics that gives rise to the acoustic peaks), we find:

\[
|\alpha_{lm}|_{L,M}^2 = \frac{s^2U_0^2h}{2\pi a^2} \frac{1}{l(l+1)}.
\]

Now, since this does not depend on \( m \) it is clear that the theoretical expectation for the observational quantity is,

\[
OB_l = \frac{l(l+1)}{2(l+1)} \sum_m |\alpha_{lm}|^2 = (\pi/3)G\frac{(V')^2}{V}U_0^2 = (2\pi/3)\epsilon\bar{V}U_0^2,
\]

where we have used the standard definition of the dimensionless slow roll parameter \( \epsilon \equiv (1/2)(M_{Pl}^2/h)(V'/V)^2 \) which is normally expected to be rather small and the dimensionless potential \( \bar{V} \equiv Vh^3/M_{Pl}^4 \). The form of the spectrum is thus in agreement with the so called scale invariant spectrum obtained in ordinary treatments and in the observational studies. Thus, if \( U \) could be prevented from becoming too large during re-heating, the quantity of interest would be proportional to \( \epsilon \) a possibility that was not uncovered in the standard treatments. That is , the present analysis offers a path to get rid of the “fine tuning problem” for the inflationary potential, i.e. even if \( Vh^3 \sim M_{Pl}^4 \), the temperature fluctuations in the CMB could remain rather small (at the level of \( 10^{-5} \) as observed in the CMB).

Now, let us focus on the effect of the finite value of times of collapse \( \eta_k \). That is, we consider the general functional form of \( C(k) \). The first thing we note is that in order to get a reasonable spectrum there seems to be only one simple option: That \( z_k \) be essentially independent of \( k \) that is the time of collapse of the different modes should depend on the mode’s frequency according to \( \eta_k^2 = z/k \). This is a rather strong conclusion which could represent relevant information about whatever the mechanism of collapse is.

Next we consider a simple proposal about the collapse mechanism which following Penrose’s ideas is assumed to be tied to Quantum Gravity, and examine it with the above results in mind.
3. A version of ‘Penrose’s mechanism’ for collapse in the cosmological setting

Penrose has for a long time advocated that the collapse of quantum mechanical wave functions might be a dynamical process independent of observation, and that the underlying mechanism might be related to gravitational interaction. More precisely, according to this suggestion, collapse into one of two quantum mechanical alternatives would take place when the gravitational interaction energy between the alternatives exceeds a certain threshold. In fact, much of the initial motivation for the present work came from Penrose’s ideas and his questions regarding the quantum history of the universe.

A very naive realization of Penrose’s ideas in the present setting could be obtained as follows: Each mode would collapse by the action of the gravitational interaction between it’s own possible realizations. In our case, one could estimate the interaction energy $E_I(k, \eta)$ by considering two representatives of the possible collapsed states on opposite sides of the Gaussian associated with the vacuum. We interpret $\Psi$, literally as the Newtonian potential and consequently the right hand side of Eq. (5), (after a rescaling by $a^{-2}$ to replace the laplacian expressed in the comoving coordinates $x$ to a laplacian associated with coordinates measuring physical length) should be identified with matter density $\rho$. Therefore, $\rho = a^{-2} \dot{\phi}_0 \delta \phi$. Then we have:

$$E_I(\eta) = \int \Psi^{(1)}(x, \eta) \rho^{(2)}(x, \eta) a^3 d^3x = a \int \Psi^{(1)}(x, \eta) \dot{\phi}_0 \delta \phi^{(2)}(x, \eta) d^3x.$$  (26)

Note that in this section we are ignoring the overall sign of this energy which being a gravitational binding energy would naturally be negative. We next express this energy in terms of the Fourier expansion leading to:

$$E_I(\eta) = (a/L^3) \Sigma_{k, k'} \Psi^{(1)}_{k}(\eta) \dot{\phi}_k \delta \phi^{(2)'}_{k'}(\eta) \int e^{i\kappa (k-k')} d^3x = (a/L^3) \dot{\phi}_0 \Sigma_{k} \Psi^{(1)}_{k}(\eta) \delta \phi^{(2)'}_{k}(\eta),$$  (27)

where (1), (2) refer to the two different realizations chosen. Recalling that $\Psi_k = (-s/k^2) \delta \phi$, we find

$$E_I(\eta) = -4\pi G (a/L^3) \dot{\phi}_0 \Sigma_{k}(1/k^2) \delta \phi^{(1)}_{k}(\eta) \delta \phi^{(2)'}_{k}(\eta) \approx \Sigma_k (\pi \hbar G/ak)(\dot{\phi}_0)^2.$$  (28)

Where we have used equation (11), to estimate $\delta \phi^{(1)}_{k}(\eta) \delta \phi^{(2)'}_{k}(\eta)$ by $|< \delta \phi, |^2 = \hbar k L^3(1/2a)^2$.

This result can be interpreted as the sum of the contributions of each mode to the interaction energy of different alternatives. According to all the considerations we have made, we view each mode’s collapse as occurring independently, so the trigger for the collapse of mode $k$ would be, in accordance to Penrose’s ideas, the condition that this energy $E_I(k, \eta) = (\pi \hbar G/ak)(\dot{\phi}_0)^2$ reaches the ‘one-graviton level’, namely, that it equals the value of the Planck Mass $M_p$. Now we use the specific expressions for the scale factor $a = \frac{1}{\pi H}$ and the slow rolling of the background scalar field $\dot{\phi}_0 = (1/3)(\dot{a}/a^3)V'$ to find

$$E_I(k, \eta) = \frac{\pi \hbar G}{9H^2}(a/k)(V')^2.$$  (29)

Thus the condition determining the time of collapse $\eta_k$ of the mode $k$ becomes,

$$\eta_k = \frac{\pi}{9} (\hbar V')^2 (H I M_p)^{-3} = \frac{\epsilon}{8\sqrt{6\pi}} (V')^{1/2} \equiv z^c,$$  (30)

which is independent of $k$, and thus, leads to a roughly scale invariant spectrum of fluctuations in accordance with observations. Note that the energy of mode $k$ in Eq. (29) is an increasing function of conformal time $\eta$, during the slow roll regime, indicating that during the very early times $E_I(k, \eta) < M_{Planck}$ so the various alternatives can coexists as quantum superpositions.
according to Penrose’s ideas. However as $a$ grows this “interaction energy” reaches the level in which the alternatives can not longer coexist and a collapse is triggered.

We can look closer into this issue and ask when do the relevant modes collapse?. In order to do this we use the value for $\zeta$ and recall that the time of collapse is determined by $\eta_k = \zeta \pi / k$, and thus the scale factor at the time of collapse of the modes with wave number $k$ was $a_k\pi = (H\eta_k)^{-1} = (12/\epsilon)kl_p(V)^{-1}$ where $l_p$ stands for the Planck length. As the value of the scale factor $a$ at the last scattering surface was $a \approx 10^{-4}$ (recall that the scale factor $a$ has been set so its value today is 1), the modes that are relevant to say scales of order $10^{-3}$ of the size of the surface of last scattering (corresponding to a fraction of a degree in today’s sky) have $k \approx 10^{-10}$ly$^{-1}$.

Thus, taking $\epsilon \times \tilde{V}$ to be of order $10^{-5}$, we have for those modes $a_k^{\pi} \approx 10^{-45}$ corresponding to $N_e = 103$ e-folds of total expansion, or something like 80 e-folds before the end of inflation in standard type of inflationary scenarios. Thus in this scheme inflation must have at least 90 e-folds for it to include the complete description of the regime we are considering and to account also for the collapse of the modes that are of the order of magnitude of the surface of last scattering itself. The usual requirements of inflation put the lowest bound at something like 60 e-folds of inflation so the present requirement is not substantially stronger.

We next compared the collapse time with the so called, time of “Horizon crossing” $\eta_k^H$ for mode $k$, corresponding to the time in which physical wavelength reaches the Hubble distance $H^{-1}$. These latter time is thus determined from: $a_k^H = a(\eta_k^H) = k/(2\pi H) = kl_p(3/2\pi^3)^{1/2}(V)^{-1/2}$. Therefore the ratio of scale factors at collapse and at horizon crossing for a given mode is $a_k^H/a_k^{\pi} = (16/\epsilon)(6\pi^3)^{1/2}(V)^{-1/2}$, which would ordinarily be a very large number, indicating that the collapse time would be much later that the time of “Horizon exiting” or crossing out, of the corresponding mode.

Thus we find that a naive realization of Penrose’s ideas seems, at first sight, to be a good candidate to supply the element that we argued is missing in the standard accounts of the emergence of the seeds of cosmic structure from quantum fluctuations during the inflationary regime in the early universe.

4. An alternative collapse scheme and the fine tuning problem

Here we consider some more speculative ideas which have not yet undergone a careful study. Nevertheless, it seems worthwhile to present them at this point to illustrates the power of the new way of looking at some of the relevant issues. We have earlier considered a collapse scheme that seemed very natural, but there is another scheme that could be considered even more natural in light of the point view explored in the previous section: that the uncertainties in the matter sources of the gravitational field are the triggers of the collapse. We note that it is only the conjugate momentum to the field $\pi_k(\eta)$ that acts as source of the “Newtonian potential” in Eq. (5) and contributes to the gravitational interaction energy in Eq. (26), thus it seems natural to assume that it is only this quantity what is subject to the collapse (i.e is only this operator that is subjected to a “self induced measurement”) while the field $y_k(\eta)$ itself is not. In this case, the analysis is almost identical: The collapse is defined by

$$\langle y_k^{R,I}(\eta_k^R) \rangle_\Theta = 0, \quad \langle \pi_k(\eta_k^R) \rangle_\Theta = x_k^{R,I} \sqrt{\langle \Delta \pi_k(y)^{R,I} \rangle_0^2} = x_k^{R,I} |g_k(\eta_k^R)| \sqrt{hL^3/2}, \quad (31)$$

where $x_k$ are selected randomly from within a Gaussian distribution centered at zero with spread one. Again from these equations we solve for $d_R^{R,I}$, and proceed as before. The only difference so far is that the function $C(k)$ containing information about the collapse changes slightly (see [1]) to:

$$C'(k) = 1 + (1 - 1/z_k^2) \sin(\Delta_k)^2 - (1/z_k) \sin(2\Delta_k). \quad (32)$$
Compare the above expression with that corresponding to the first collapse scheme Eq. (22). However the point we want to make is that this scheme seems to be a rather serious candidate to alleviate the fine tuning problem, that, as we have mentioned in the discussion around Eq. (25), seems to affect most inflationary scenarios. The point is that according to quite general analysis, [7] the quantity:
\[ \zeta \equiv \Psi + aH\frac{\delta \phi}{\dot{\phi}_0} \]
remains constant through the cosmological evolution even if the equation of state changes, so it seems natural to expect that in our context the corresponding quantity
\[ \zeta \equiv \Psi + aH\langle \delta \phi \rangle_{\Theta}/\dot{\phi}_0 = \Psi + H\langle y \rangle_{\Theta}/\dot{\phi}_0 \] (33)
would be conserved from immediately after the collapse (a process through which the classical equations would not hold) through the reheating and up to the late times associated with the observation. (we are essentially relying on Ehrenfest’s theorem). However for the collapse mode we have considered in this section the last term in the equations vanishes just after the collapse so the value of the Newtonian potential would be that of the estimation we have made before, and, as indicated in the discussion of the last part of section 2, this indicates a substantial amelioration of the fine tuning problem. This seems a very interesting possibility, but clearly it must be investigated much more profoundly before any compelling claims in this regard could be made.

5. Additional considerations

Another important observation follows directly from the basic point of view adopted in this analysis: The source of the fluctuations that lead to anisotropies and inhomogeneities lies in the quantum uncertainties of the scalar field, which collapse due to some unknown quantum gravitational effect. Once collapsed, these density inhomogeneities and anisotropies feed into the gravitational degrees of freedom leading to nontrivial perturbations in the metric functions, in particular the so called newtonian potential. However, the metric itself is not a source of the quantum gravitational induced collapse (in following with the equivalence principle the local metric perturbations have no energy). Therefore, as the scalar field does not act as a source for the gravitational tensor modes – at least not at the lowest order considered here – the tensor modes can not be excited. The scheme thus naturally leads to the expectation of a zero– or at least a strongly suppressed– contribution of the tensor modes to the CMB fluctuation spectrum. See however[8] for another scheme which also leads to strong suppression of tensor modes.

Finally we address a recent article [9] in which colleagues reiterate, in response to [1], that the problem we have alluded to, does not exist. It is illuminating to consider their claims and views in this regard: 1) That the situation is analogous to the spontaneous symmetry breaking in field theories, 2) that the environment selects as preferential “pointer” basis the the one that diagonalizes the field operators rather than the momenta operators because of couplings of other fields to the field operators, and 3) that ”the initial symmetric vacuum state evolves into a symmetric superposition of inhomogeneous states out of which one component is selected”. The first point is rather complex to discuss and will be addressed elsewhere, so lets focus here on the last two. Point 2) can be seen to be in fact incorrect, because most of the known interactions are of the guage-theory type and couple both to the momentum and the other ”spatial” field gradients rather than to the fields themselves. However the clearest problem lies in point 3) where the authors do acknowledge that the unitary evolution leads at late times to a “symmetric superposition of inhomogeneous states” which is, according to quantum mechanics nothing but a fully symmetric state. Then somehow, without any recur to a physical process or mechanism, at some unspecified time, one of the components of this state gets “selected”, why? when ? by whom? are questions that apparently need no answer.
6. Conclusions

We have reviewed a serious shortcoming of the inflationary account of the origin of cosmic structure, and have given a brief account of the proposals to deal with them which were first reported in [1]. These lines of inquiry have lead to the recognition that something else seems to be needed for the whole picture to work and that it could be pointing towards an actual manifestation quantum gravity. We have shown that not only the issues are susceptible of scientific investigation based on observations, but also that a simple account of what is needed, seems to be provided by the extrapolation of Penrose’s ideas to the cosmological setting. Interestingly the scheme does in fact lead to some deviations from the standard picture where the metric and scalar field perturbations are quantized. For one, as we pointed out above, one is lead to expect no excitation of the tensor modes. This approach also opens new avenues to address the fine tuning problem that affects most inflationary models, because one can follow in more detail the objects that give rise to the anisotropies and inhomogeneities, and by having the possibility to consider independently the issues relative to formation of the perturbation, and their evolution through the reheating era. Some of these aspects can, in principle, be tested, indicating that what initially could have been thought to be essentially a philosophical problem, leads instead to truly physical issues.

Our main point is however that in our search for physical manifestations of new physics tied to quantum aspects of gravitation, we might be ignoring the most dramatic such occurrence: The cosmic structure of the Universe itself.

Acknowledgments

It is a pleasure to acknowledge very helpful conversations with J. Garriga, E. Verdaguer and A. Perez. This work was supported in part by DGAPA-UNAM IN108103 and CONACyT 43914-F grants.

References

[1] A Perez, H Sahlmann and D Sudarsky, Class and Quant Gravity 23 2317-54 (2006).

[2] A E Lange et. al. Phys. Rev. D 63 042001, (2001);
G Hinshaw et. al., Astrophys. J. Supp., 148, 135 (2003);
K M Gorski et. al. Astrophys. J. 464 L11, (1996);
C L Bennett et. al. Astrophys. J. Suppl. 148, 1,(2003).

[3] A Guth and S Y Pi Phys. Rev. D 32, 1899, (1985);
S W Hawking Nucl. Phys. B 224, 180, (1983);
J J Halliwell and S W Hawking, Phys. Rev. D , 31, 1777,(1985).

[4] “The Emperor’s New Mind”, R Penrose, The Emperor’s New Mind, (Oxford University Press 1989).

[5] J J Halliwell, Phys. Rev. D 39, 2912,(1989);
C Kiefer Nucl. Phys. Proc. Suppl. 88, 255 (2000);
D Polarski and A. A. Starobinsky [arXiv:gr-qc/9504030] (1996);
W H Zurek, Environment Induced Superselection In Cosmology in Moscow 1990, Proceedings, Quantum gravity (QC178:S4:1990), 456-72 (see High Energy Physics Index 30 (1992) No. 624);
R Branderberger H Feldman and V. Mukhavov,Phys. Rep., 215, 203, (1992);
R Laflamme and A Matala. Int. J. Mod. Phys. D 2, 171 (1993);
M Castagnino and O Lombardi, Int. J. Theor. Phys. 42, 1281, (2003);
F C Lombardo and D Lopez Nacir, Phys. Rev. D 72, 063506 (2005);
J Martin, Lect. Notes Phys. 669, 199 (2005);
L P Grishchuk and J Martin, Phys. Rev. D 56, 1924 (1997).

[6] T Padmanabhan, Cosmology and Astrophysics Through Problems, (Cambridge University Press 1996);
J B Hartle,“Generalized Quantum mechanics for Quantum Gravity”, arXive: gr-qc/0510126.

[7] J Garriga and V F Mukhanov Phys.Lett. B 458 219-225,(1999).

[8] J Khoury, P J Steinhardt and N Turok Phys. Rev. Lett. 91 161301, (2003).

[9] C Keifer, I Lohmar, D Polarski and A. S. Starobinsky, arXive: astro-ph/0610700