

The very early universe : Problems and perspectives*

Jayant V. Narlikar

Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005

Received 1984 December 7

1. Introduction

One of the most exciting and actively pursued frontier areas today is the study of the very early universe. 'Very' tells us that the epoches under consideration are considerably earlier than those first discussed by George Gamow and his colleagues Ralph Alpher and Robert Herman back in the 1940s. The early universe of Gamow was concerned with the time span 1-200s after the big bang. During this time span the baryonic component of the universe underwent nucleosynthesis : the protons and neutrons combined to form light atomic nuclei like ^4He , ^3H , etc. The bulk of the matter part of the universe at this stage was of course in the form of free protons. The next important component was made of the ^4He nuclei, about 25% by mass. The remaining parts were extremely tiny mass fractions of ^2H , ^3He , ^6Li , ^7Li , etc. And of course there was plenty of radiation.

The discovery of the microwave background in 1965 and the realization (during the 1960s and the early 1970s) that the hot early universe provided the only scenario for explaining the observed abundance of nuclei led to considerable confidence in the big bang picture. This in itself would probably not have triggered off the present flurry of activity in the field of the very early universe. The stimulus came from another, entirely unexpected, quarter. The particle physicists who, as high priests of physics had hitherto looked upon cosmology as a somewhat dubious part of physics, suddenly realized that the very early universe offered the only scenario for testing the authenticity of their holy grail : the grand unified theories (GUTs) and the more esoteric ideas about supersymmetry (SUSY). So the investigations of the universe were swiftly elevated from speculative parascience to the most fundamental area of theoretical physics.

In my talk today I propose to describe some highlights of this frontier area and the problems that have appeared on the footsteps of its successes.

*Presidential address delivered at the tenth meeting of the Astronomical Society of India 1984 November at Bombay.

2. The timetable of the big bang universe

Figure 1 outlines the time-temperature sequence in the universe right from its origin in a big bang at time $t = 0$. The key to the timetable lies in the relation

$$t_{\text{second}} = 2.4 g^{-1/2} T_{\text{MeV}}^{-2} = 2.4 \times 10^{-6} g^{-1/2} T_{\text{GeV}}^{-2} \quad \dots(1)$$

relating the time elapsed since the big bang to the temperature T of the universe as measured from the thermodynamic equilibrium of the various species of relativistic particles in it. (See Narlikar 1983 for details of derivation of the above formula). The

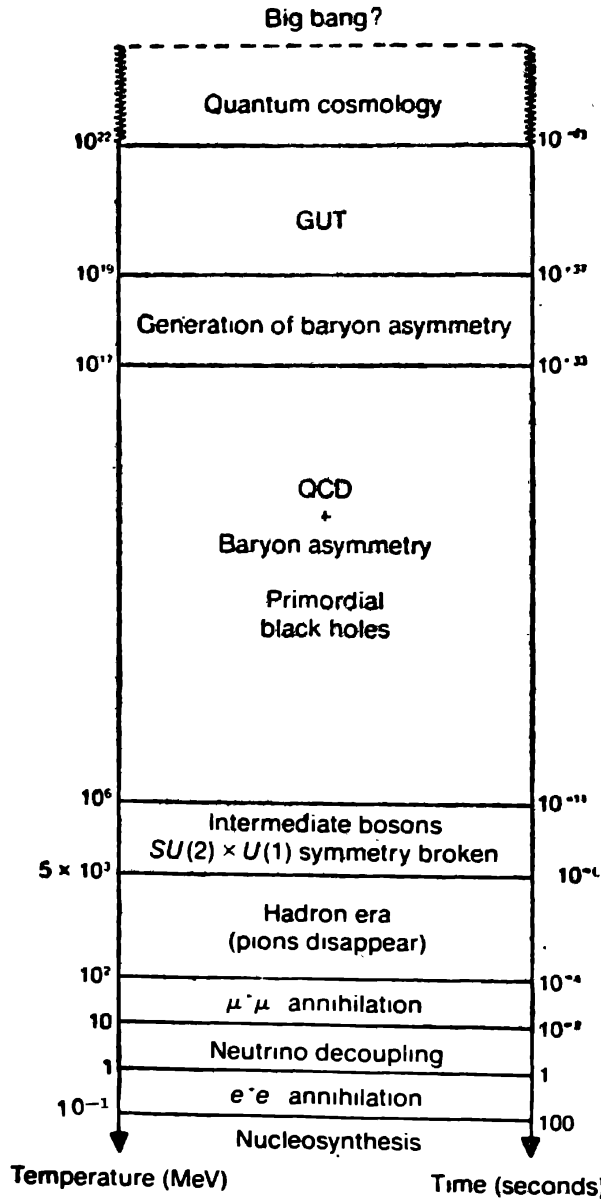


Figure 1. The time-temperature sequence in the early universe from the big bang epoch $t = 0$ to the epoch ($t \cong 200$ s) by which time the primordial nucleosynthesis is essentially over.

constant g denotes the effective number of spin states of the primordial particles. The number of fermionic spin states has to be multiplied by the factor $7/8$ to get their effective number while the number of bosonic spin states is included unchanged in g . The temperature is expressed in the units of MeV (or GeV)—a fact that reflects the influence of particle physics on cosmology.

The formula (1) is made of two separate relations which bring out the vital role of cosmology in this game. Denoting the linear scale factor of expansion of the universe by S , we have the following two relations from Einstein's cosmological equations :

$$S \propto t^{1/2}, \quad T \propto \frac{1}{S}. \quad \dots(2)$$

The advantage of expressing T in MeV, GeV etc. is that the formula (1) tells us at a glance what the typical energy of a relativistic particle at any given time was. Gamow's early universe at the age of 1s had the relativistic fermions e^- , e^+ , ν_e , $\bar{\nu}_e$, ν_μ , $\bar{\nu}_\mu$; the only relativistic bosons being the photons. Thus $g = 9$ and the formula (1) gives $T \cong 0.9$ MeV. At this temperature the baryons were nonrelativistic, as were the mesons and the muons.

The GUTs however require energies of the order of 10^{15} GeV or more—some twelve orders of magnitude higher than the particle energies possible in man-made accelerators. The gap is simply too big to be bridged with the present technology or even with technologies of the foreseeable future. The GUTs are therefore untestable in any terrestrial laboratory, and as such they would have degenerated to mere speculations ... but for the very early universe.

A glance at formula (1) tells us that for $g \cong 100$, say, the particle energies of 10^{15} GeV were attained in the universe when it was only $\sim 2.4 \times 10^{-37}$ s old. Thus in order to study the full dynamical effects of the grand unified theories we have to consider the universe that young. Likewise the SUSY energies are $\sim 10^{17}$ GeV and they translate to a time $t \sim 10^{-40}$ s.

Figure 1 summarizes this temporal sequence. The higher the relevant particle energy of interaction the earlier the epoch at which that interaction was significant. The notion of symmetry breaking at lower energies is thus translated to the notion of symmetry breaking at later epochs when the universe was cooler and comparatively less energetic. The GUT symmetry is therefore broken at $t \sim 10^{-37}$ s, while the electroweak symmetry is broken 'much later' at $t \sim 10^{-10}$ s.

The shortest timescale, however, is connected with quantum gravity. It is none other than the Planck time

$$t_p = \sqrt{\frac{G\hbar}{c^5}} \cong 5 \times 10^{-44} \text{ s}. \quad \dots(3)$$

The dynamics of the universe as determined by the *classical* gravitational equations of Einstein is suspect prior to this era. This of course raises doubts about the very existence of the big bang epoch at $t = 0$. I will return to this question towards the end of my talk.

3. Relics of the very early universe

If GUTs and SUSY are to be tested against the background of the very early universe, the only way to do this is to ask for predictions of survivors of the epoch when those theories played an active role. This was the philosophy behind Gamow's calculations of the early universe. The relics of the epoch 1–200 s when those calculations were relevant are the light nuclei and the microwave background. What relics do we expect from the era $t \sim 10^{-37}$ s ?

One of the relics that could possibly have to do with this very early era is the ratio of photons to baryons in the present universe. This ratio, given the observational uncertainties, appears to lie in the range $\sim 10^8$ – 10^{10} . As there could not have been any significant change in this ratio since the above era, its value must somehow have been fixed then. Why should this value lie in this particular range ?

Before GUTs or SUSY came on the scene the somewhat naive estimates of this ratio threw up two difficulties. First, there was nothing then known or expected in fundamental physics that suggested how a net baryon number could be created. Assuming that the universe started with a symmetric composition of matter and antimatter, there were no means of generating a net baryon number. Secondly, even if one assumed that baryons and antibaryons are present today in equal numbers, these when compared to the number of photons were too few ! The early calculations (see Steigman 1976) gave the photon to baryon ratio as high as $\sim 10^{18}$ (and of course the same ratio for photons to antibaryons because of matter antimatter symmetry),

GUTs changed all that. In the simple SU(5) model for example, the basic ingredients are six quarks and six leptons. There are 24 bosons in theory some of which (called tentatively the X-bosons) mediate in an interaction that changes quarks to leptons and vice versa. Figure 2 illustrates a way opened out by this possibility, of the decay of a proton.

Proton decay has been predicted with timescales $\geq 10^{31}$ yr by many versions of GUTs. Although the timescale is long, the stochastic nature of the decay process makes it possible to observe proton decay events in large piles of matter (since it is essentially made of nucleons). It is too early to say whether proton decay has definitely been observed in some of the many experiments currently set up in different parts of the world, or to rule out some GUT models.

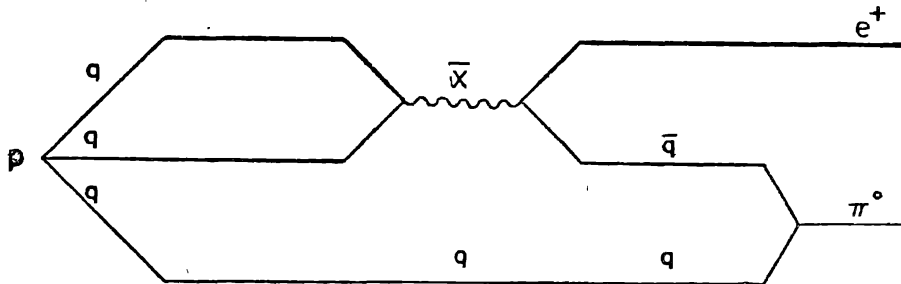


Figure 2. One of the possible modes for the decay of a proton. Two quarks in the proton combine to form anti-X boson that decays into a positron and antiquark. The antiquark combines with the third quark to form a pion.

Nevertheless it is now possible to visualize a scenario in the very early universe when the X-bosons created or destroyed baryons and antibaryons. To generate a net baryon number, however, two more ingredients were needed in this cooking recipe. First, the rate at which baryons are produced should be different from the rate at which antibaryons are produced. Further, there should be lack of thermodynamic equilibrium in order to create an asymmetry between the baryon generating reactions and their inverses, the baryon destroying reactions (and like wise for antibaryons).

By treading judiciously in the parameter space of the theory it is possible to produce a photon to baryon ratio in the range 10^4 – 10^{12} , encompassing the observed range. Depending on your personal prejudices, you can either look upon this result as a remarkable achievement that has solved an outstanding problem in cosmology or as a parameter-fitting exercise that offers nothing more than a plausible scenario for the very early universe.

GUTs also tell us about the existence of many species of neutrinos and their likely mass ranges. Neutrinos have several consequences for the present universe. Some are as follows.

(a) If there are too many species of massive neutrinos that survive to the present day, they will push up the mass density of the universe to uncomfortably high values. For example, if the mass of ν_i a neutrino of i -th species to survive is m_i in electron volts, then the ratio of the neutrino density ρ_ν to the closure density ρ_c is given by

$$\Omega_\nu \equiv \frac{\rho_\nu}{\rho_c} = \sum_{i=1}^N \frac{m_i}{150} \left(\frac{T_0}{3} \right)^3 h_0^{-2}. \quad \dots(4)$$

Here N is the total number of neutrino species and their antiparticles, T_0 is the present temperature of the microwave background, while the Hubble constant is $100 h_0 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The parameter h_0 lies in the range 0.5 to 1.

An immediate problem of having too many massive neutrinos surviving to this day is illustrated in figure 3: for high values of Ω_ν the age of the universe is too low. For $\Omega_\nu = 1$ the age is less than $6.6 h_0^{-1}$ billion years. Even for $h_0 = 1/2$ the age is probably too low to accommodate globular cluster ages (15–18 billion years). For $N = 6$ (three neutrinos ν_e, ν_μ, ν_τ and their antiparticles) equation (4) tells us that $\Omega_\nu = 1$ for $m_i = 25$ (eV). Evidently there is a problem here if the present claims of neutrino masses ~ 30 eV (Lubimov 1983) are borne out.

(b) The number of neutrino species present at this time of primordial nucleosynthesis affects the abundance of primordial helium. The value 0.25 for the mass fraction of primordial helium is for two neutrino species, ν_e , and ν_μ . For each new species the fraction goes up by ~ 0.02 . Already there are indications that the primordial mass fraction of ^4He may be as low as 0.22 (Piembert *et al.* 1978). Thus introduction of further species will exacerbate the discrepancy.

(c) One of the problems of cosmology concerns galaxy formation. How did massive structures like galaxies, clusters of galaxies, and super-clusters evolve in a

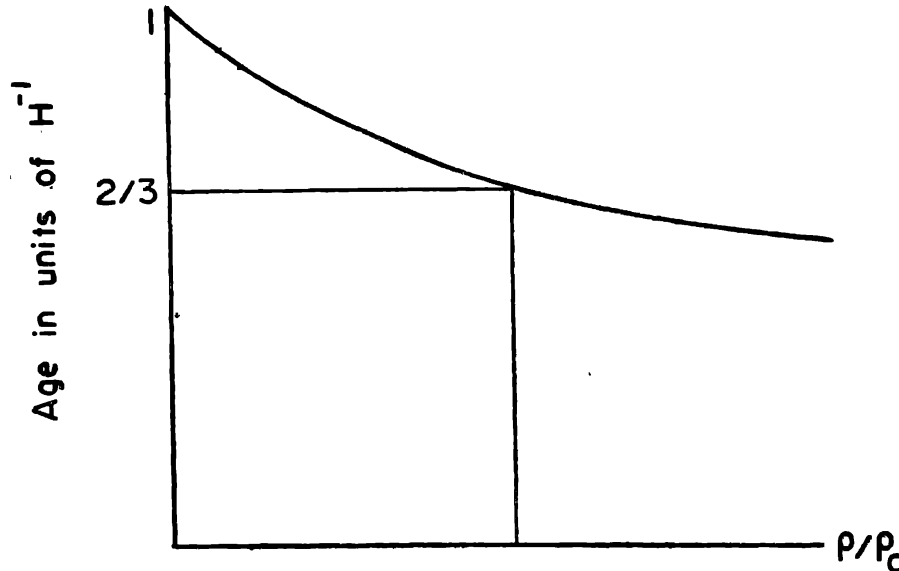


Figure 3. The age of the Friedmann universe expressed in units of H^{-1} as a function of $\Omega =$ present matter density/ ρ_c .

universe that was initially quite smooth in composition? When numerous attempts to construct plausible scenarios failed, it was hoped that fluctuations in a neutrino-dominated universe would provide the answer. In a recent review Hut & White (1984) have discussed the many problems facing such a scenario. For neutrinos with rest masses ~ 30 eV, as suggested by the Russian experiment, this scenario produces a galaxy distribution that is too highly clumped compared to what is observed. The discrepancy could be resolved if apart from the electron neutrino two more massive neutrinos exist with masses in the range 0.1 MeV–0.25 GeV. Such massive neutrinos cannot of course survive to present epoch and must decay. The numerous astrophysical constraints on decay times and masses do not rule out massive neutrinos but as Hut & White conclude ‘there seems no natural reason why the parameters of real neutrinos should lie in the narrow ranges required to get a viable cosmology’.

A related problem of relic monopoles may also be briefly mentioned here. Although pure Maxwellian electrodynamics does not allow the existence of magnetic monopoles, GUTs do. In fact monopoles with masses as high as $\sim 10^{16}$ GeV are believed to arise in the very early universe. Can they be destroyed after creation? If not, calculations tell us that the monopole mass density in the present universe should be as high as $> 10^{14} \rho_c$. This result is too absurd to be correct. But then how were the monopoles eliminated?

4. Inflation

Three years ago Guth (1981) proposed a novel idea which has now taken root in the early-universe mythology, if not in quantitative detail, at least qualitatively. (I shall comment on this remark later.) Guth’s idea was to suggest that the phenomenon of symmetry breaking whereby a GUT breaks into the strong and the

electroweak interactions does not take place instantaneously. The change produced by symmetry breaking is like a phase transition from say gaseous state to a liquid state. In Guth's scenario, the phase transition takes place in the very early universe more like that of a supercooled steam. Such steam is known to be unstable and condenses into droplets.

Taking this analogy to the GUTs phase transition, Guth found that the corresponding process generates considerably large negative pressures that cause the 'droplets' to expand. In other words, because of the gravitational coupling of energy to spacetime structure via Einstein's equations, parts of the universe expand rapidly. The scale factor, in contrast to its $t^{1/2}$ -type time dependence, behaves like

$$S \propto \exp at. \quad \dots(5)$$

Such a rapid expansion was termed by Guth as 'inflation'.

What does the inflationary universe achieve? It resolves some of the outstanding difficulties of the big bang cosmology. One of these difficulties relates to the *flatness* of the present space. To understand this difficulty let us note that the spatial geometry of the Friedmann universe can be of three types as indicated by a parameter k . For $k = 1$ the space is closed; it has constant positive curvature, a finite volume but no boundary. For $k = 0$ the space is flat, *i.e.*, it has Euclidean geometry. For $k = -1$ the space is open with constant negative curvature. Figure 4 illustrates expansion properties of the three types of models.

Notice that for large values of t the scale factor behaves in markedly different manner in the three cases. The universe expands and then contracts for $k = 1$. It expands for ever, though with different rates, for $k = 0, -1$.

However, close to $t = 0$ the three curves merge into each other. If we take $k = 0$ (the flat space) as the reference standard, minute differences in the expansion rate in the very early stages get exaggerated at later stages. When we take this fact into account in assessing the present observations we begin to appreciate the flatness problem.

Basically, the ratio

$$\Omega_0 = \frac{\rho_0}{\rho_c} \quad \dots(6)$$

of the present density to closure density decides which value of k is correct. If $\Omega_0 = 1, k = 0$ while $\Omega_0 > 1$ (< 1) implies $k = +1$ (-1). The observational uncertainty in Ω_0 is some what like this :

$$0.1 \lesssim \Omega_0 \lesssim 5. \quad \dots(7)$$

Although the uncertainty of a factor 50 at present seems large, it arose out of an extremely small magnitude at the very early stage. The curves in figure 4 lying *now* within the range (7) were lying in a range of Ω at $t \sim 10^{-37}$ s as small as

$$|\Omega - 1| \lesssim 10^{-52}. \quad \dots(8)$$

What made the universe so finely tuned to the 'flat' value ?

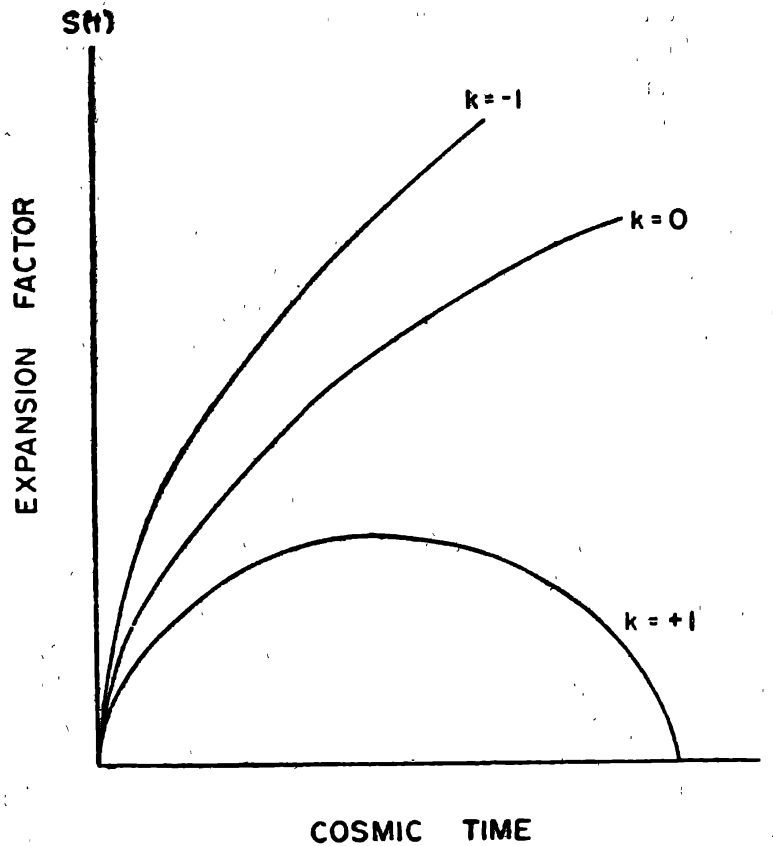


Figure 4. The three curves above illustrate the three modes of expansion of the Friedmann models for $k = 1, 0, -1$.

In the inflationary universe it is argued that the subregion expanding exponentially quickly attains a value of Ω so finely tuned to 1 as to be practically equal to 1 even at the present stage. In other words, the difference $|\Omega - 1|$ is not only as small as that given by equation (8) but is in fact much smaller. Thus it is no surprise that our observations reveal the range (7): the correct value is $\Omega_0 = 1$.

Inflation also helps in wiping away inhomogeneities in the expanding subregion. It thereby solves the so called 'horizon problem'. The horizon in the big bang arises as illustrated in figure 5. For $S \propto t^{1/2}$ as given by equation (2), there is a limit to which an observer can see at any given time. The range r_H is limited by

$$r_H \lesssim 2 ct, \quad \dots (9)$$

where c = speed of light. At $t = 10^{-37}$ s, $r_H \cong 1.5 \times 10^{-27}$ cm. The horizon limits the distance over which physical influences travel at any given time. Thus at $t = 10^{-37}$ s, this limit was $\sim 10^{-27}$ cm. Now at this stage the temperature of the universe was $\sim 10^{15}$ GeV $\sim 10^{28}$ K. This temperature has now fallen to ~ 3 K by the inverse scale law (2). The region of size 10^{-27} cm has thus expanded to ~ 10 cm.

This is an extremely small region by present cosmological standards, the Hubble radius being $10^{28} h_0^{-1}$ cm. The microwave background shows uniformity over this

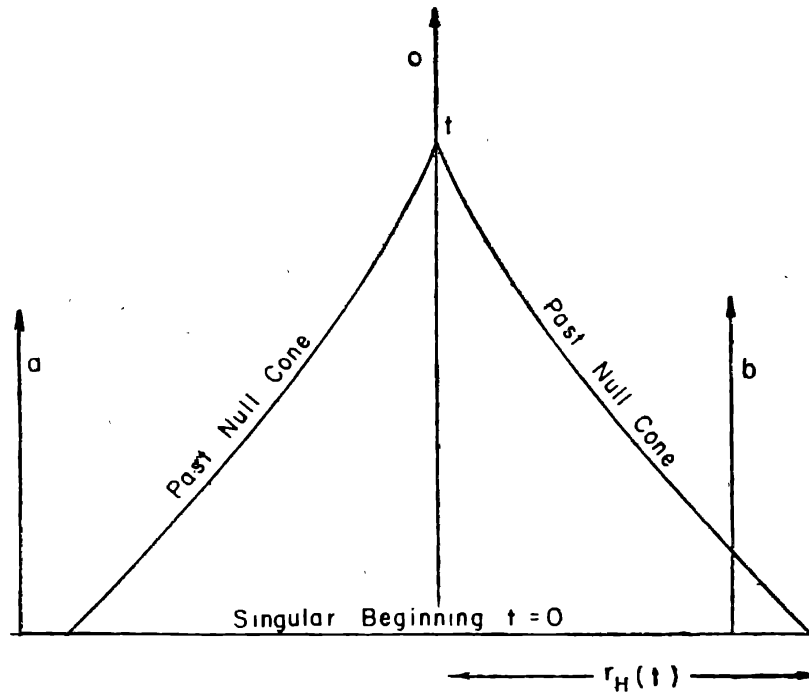


Figure 5. The past light cone of an observer O at epoch t intersects the $t = 0$ hypersurface in a sphere whose radius is indicated by $r_H(t)$ in the above two-dimensional plot. Only those particles like b whose world lines at $t = 0$ were within $r_H(t)$ could be seen by O , and could influence O . This sphere is called the particle horizon of O . Observers like a lying outside the particle horizon of O cannot influence O yet.

distance in spite of the fact that at the crucial GUTs era uniformity was possible over a very small fraction (10^{-28}) of this size. In other words, how, in spite of the limiting horizons, did the universe manage to establish uniformity over such vast regions?

Inflation solves this problem by blowing up the small horizon-bound subregion into a vastly bigger size. The exponential function has this facility of generating a large magnification over modest times. For the same reason, it is argued that the awkward magnetic monopoles would not survive in a preinflated uniform universe.

In spite of its apparent successes this scenario has many problems of its own. Although it shows that the present universe must be extremely finely tuned to the $k = 0$ value (with $\Omega = 1$ to a high degree of accuracy) the details seem to need parameters with narrow ranges of values no less contrived than the fine tuning it was meant to explain. For example it is assumed that the λ -term generated by GUTs must exactly cancel the opposite λ -term of cosmological origin that is supposed to have existed prior to the inflationary phase. The cancellation has to be finely tuned to one part in 10^{108} . Why should this happen (Wilczek 1982)?

A further difficulty is that the inflationary universe still generates fluctuations or inhomogeneities that are too large to explain galaxy formation. It is not possible for me to go into details of how this conclusion is arrived at (see *e.g.* Wilczek 1982; Hawking 1982; Guth & Pi 1982; Starobinskii 1982; Bardeen *et al.* 1983). It remains possible of course to think of an 'epicycle' to patch up this difficulty.

The word 'epicycle' is used here to remind us of the attempts by the ancient Greek astronomers to explain the trajectories of planets with the help of epicycles, that is, circles whose centres moved on other circles whose centres moved on other circles and so on. The reality in this case was far from what was believed; but the successive addition of epicycles led to the erroneous belief that the theory was on the right lines. Likewise, the inflationary scenario which began with many attractive qualitative features has now, after several metamorphoses, found that quantitative details require patchwork. The original Guth version of inflation has by now picked up so many epicycles that it has lost its basic charm and simplicity. I have no answer however to a staunch defender of the inflationary universe who asks 'who told you that the universe is simple anyway?'

5. Quantum cosmology

It seems to me that some of these problems arise because the physicists have not been daring enough! The really important era in the history of the universe probably occurred even earlier than 10^{-37} s. We have already seen that prior to $t_p \cong 5 \times 10^{-44}$ s, classical gravity cannot be trusted. Can quantum gravity offer us something new, thereby solving some of these problems just as quantum electrodynamics solved many outstanding problems of classical electrodynamics?

Unfortunately, quantizing gravity has not been that easy. There have been numerous formal approaches, all of which are still being developed without producing any practical dividend. Against this background I wish to report briefly some work done here (Narlikar & Padmanabhan 1983) that shows a positive and definitive outcome.

Basically Padmanabhan & I adopted a limited objective that has particular relevance to cosmology. In the expanding universe model the main dynamical degree of freedom is carried by the scale factor. To capture the essence of difference between classical and quantum cosmology we should therefore quantize the scale factor. To do so covariantly, however, one should not restrict oneself to homogeneous isotropic models with S a function of t only. Rather one should quantize the conformal part of the metric. Fortunately this can be done without difficulty for two reasons; one physical, the other mathematical.

The physical reason is that the conformal transformation preserves the light cone structure and hence the global causality conditions. In the absence of this simple consequence it would be extremely hard to interpret the working of physical processes during quantum transitions of spacetime geometry. The mathematical reason is that quantization of the conformal degrees of freedom can be achieved exactly by Feynman's path integral technique. Without going into specific details I quote results from this work that have a bearing on cosmology.

(i) It is found that there is growing quantum uncertainty as we probe the history of the universe closer and closer to the $t = 0$ epoch in such a way that models sharing the classical characteristic of the big bang origin are found to occur with vanishing likelihood.

(ii) The particle horizons in the majority of the models do not exist. Thus there is no inhibition to the establishment of homogenization early on in the universe during the quantum era.

(iii) If the universe arose out of an empty Minkowski universe by quantum conformal fluctuations then with overwhelmingly large probability it will go into the flat ($k = 0$) Friedmann model. Thus the flatness problem is resolved.

I present these results not as the last word but as the beginning of a line of investigation that may ultimately demonstrate that the problems of the very early universe will not be fully appreciated (let alone solved) without inputs from quantum cosmology.

6. Concluding remarks

Does God play with loaded dice favouring nonsingular universes as quantum cosmology would have us believe? If this is so the universe did not have a beginning although it may have passed through phases of extremely high density when quantum rules held sway.

A universe without a beginning and without an end is not new to cosmology. In the 1950s and till 1965, before the discovery of the microwave background, the steady state theory offered such an alternative to the big bang cosmology. Indeed, many of the concepts of the steady state theory such as baryons nonconservation, the exponential expansion, the smoothing of inhomogeneities by inflation are resurfacing in the scenarios of the very early universe (Narlikar 1984). Doing away with 'the beginning' is another of the steady state legacies which seems now to be taken up by quantum cosmology.

Whatever the eventual outcome of the current interaction between cosmologists and particle physicists no one can deny that the problems and conundrums of the very early universe make it a fascinating topic of study today.

References

- Bardeen, J. M., Steinhardt, P. J. & Turner, M. S. (1983) *Phys. Rev.* **D28**, 679.
 Guth, A. (1981) *Phys. Rev.* **D23**, 347.
 Guth, A. & Pi, S. -Y. (1982) *Phys. Rev. Lett.* **49**, 1110.
 Hawking, S. W. (1982) *Phys. Lett.* **115B**, 295.
 Hut, P. & White, S. D. M. (1984) *Nature* **310**, 637.
 Lubimov, V. A. (1983) in *Proc. Int. Conf. High Energy Physics, Brighton*.
 Narlikar, J. V. (1983) *Introduction to Cosmology*, Jones & Bartlett.
 Narlikar, J. V. (1984) *J. Ap. Astr.* **5**, 67.
 Narlikar, J. V. & Padmanabhan, T. (1983) *Phys. Rep. C*, **100**, 151.
 Primbert, M., Torres-Piembert, S. and Rayo, J. F. (1978) *Ap. J.* **220**, 16.
 Starobinskii, A. A. (1982) *Phys. Lett.* **117B**, 175.
 Steigman, G. (1976) *A. Rev. Astr. Ap.* **14**, 339.
 Wilczek, F. (1982) *Proc. Nuffield Workshop on the Early Universe, Cambridge*.