Quantum Cosmology and the Creation of the Universe

JONATHAN J. HALLIWELL

Center for Theoretical Physics

Laboratory for Nuclear Science

Massachusetts Institute of Technology

Cambridge

MA 02139

USA

E-mail address: Halliwell@MITLNS.BITNET

CTP preprint # 1928 — December, 1990

Submitted to Scientific American

On staring out into space on a clear night, one can hardly help but wonder, "Where did all this come from ?" Such an esoteric question, for centuries pondered by philosophers and theologians, may seem to lie far outside the traditional domain of respectable scientific investigation. In recent years, however, developments in the cosmology of the very early universe have not only led to the *possibility* of understanding the very beginning of the universe from within the laws of physics, but have underscored the *necessity* of understand it in order to have a complete description of the universe in which we live.

The precise field in which these developments have taken place is quantum cosmology, in which quantum mechanics – usually applied only to the sub-atomic scale – is applied to the entire universe. The field is not a new one. Much of the groundwork was carried out in the 1960's, by Bryce DeWitt, now at the University of Texas at Austin, Charles Misner, of the University of Maryland, and John Archibald Wheeler, then at Austin. But new impetus came to the field in the early 1980's, when a number of cosmologists began to take quantum cosmology very seriously as a realistic framework in which to address the issue of cosmological initial conditions. In particular, quite definite, testable, *laws of initial conditions*, within the context of quantum cosmology have been put forward, most notably by James Hartle, of the University of California at Santa Barbara, in collaboration with Stephen Hawking, of the University of Cambridge, by Andrei Linde, of the Lebedev Physical Institute in Moscow, and by Alexander Vilenkin, of Tufts University. Adjoined with suitable laws governing the evolution of the universe, such laws of initial conditions could conceivably lead to a complete explanation of all cosmological observations. Although this goal is still a long way off, the initial steps in that direction show considerable promise.

Central to modern cosmology is the hot big bang (HBB) model of the universe put forward by George Gamow in 1948. Based on general relativity (Einstein's theory of gravitation), together with some basic physics about the behaviour of matter, this model envisages the universe beginning from a very small, hot, dense initial state some 15 billion years ago and expanding out to the large cold universe in which we live. The HBB model made very definite predictions about the universe we see today, and these predictions have been verified observationally. In particular, it predicted the relative abundances of certain elements, and the existence and exact temperature of the microwave background – the glow of radiation permeating the universe left over from the initial explosion.

Although impressive in its predictions, however, the HBB model leaves many features of the universe unexplained. For example, there is the "horizon problem". The universe we see today consists of a vast number of regions which, in the HBB model, could never have been in causal contact at any stage in their entire history. This makes it particularly difficult to account for the striking uniformity of the universe on very large scales. Then there is the "flatness problem". This is the observational fact that the spatial geometry of the universe is extremely flat, rather than curved like the surface of a sphere. This is rather surprising in the HBB model, because in this model, the universe tends to become more curved as time evolves, and it could only be as flat as it is now if it started out almost exactly flat – flat to within one part in 10^{60}. The necessity to tune the initial conditions so finely is felt to be deeply unnatural.

Perhaps the most significant feature of the universe not adequately accounted for by the HBB model is the origin of large scale structures, such as galaxies. It has been known for a long time that large scale structures in the universe can grow out of small fluctuations in the matter density at very early times in an otherwise homogeneous universe. The requisite form and magnitude of these fluctuations was worked out in the 1970's by Edward R.Harrison, of the University of Massachusetts at Amherst, and independently, by the late Yakov B.Zel'dovich, of the Institute of Physical Problems in Moscow. But the fundamental origin of these fluctuations remained completely unknown – they had to be *assumed* as initial conditions.

This, then, was the story of modern cosmology around the beginning of the 1980's. The above three features of the universe, plus a number of other related ones, could be *accommodated* by the HBB model, possibly by choosing very special initial conditions, but there was no sense in which they were *explained* by the HBB model.

A crucial new element entered the game in 1983, when Alan H.Guth, of MIT, proposed a compelling solution to these problems, known as the inflationary universe scenario (IUS). Like the HBB model, the IUS takes the gravitational field to be classical, described by Einstein's general relativity. The matter content, however, is taken to be a type of matter called a scalar field, whereas in the HBB model, the matter content is a uniformly distributed plasma or dust. Scalar fields arise naturally in many particle physics theories, and are generally the dominant matter content in the extreme conditions encountered in the very early universe. Guth showed that the presence of this scalar field may cause the universe to undergo a very brief but exceedingly rapid period of expansion, called inflation. This expansion is much more rapid than that in the HBB model – exponentially fast, in fact, like the rate of population increase in a population explosion [see "The Inflationary Universe", by Alan H.Guth and Paul J.Steinhardt; *Scientific American*, May, 1984].

Inflation lasts for an almost inconceivably short time – some 10^{-30} seconds; but during this time the universe increases in size by an equally astounding factor of 10^{30}, from an initial size of about 10^{-28} centimetres, to a size of about one metre. At the end of this period of inflation, the decay of the scalar field producing the expansion causes the (initially cold) universe to heat up to a very high temperature, and the subsequent evolution of the universe is exactly as in the HBB model. Inflation is therefore an incredibly brief glitch tacked onto the beginning of the HBB model, but during that time – a blink of the eye even on the the sub-atomic scale – a whole host of cosmological problems are swept under the carpet.

The horizon problem is solved because due to the huge expansion, all of the observed universe emerges from a much smaller region than in the HBB model – sufficiently small that all of it can be in causal contact. The flatness problem is solved because the huge expansion wipes out all spatial curvatures; thus even if the universe really is spatially curved, it appears flat to us because it has been blown up to be so much larger than in the HBB model.

Of perhaps greatest interest is the fact that the IUS offers a possible explanation of the origin of the density fluctuations necessary for the subsequent emergence of large scale structures, such as galaxies. As mentioned above, the dominant matter field present during the inflationary era is a scalar field. This field is taken to be largely homogeneous, but may have small inhomogeneous parts. According to Heisenberg's uncertainty principle of quantum mechanics (discussed at greater length below), these inhomogeneous parts of the field cannot be exactly zero but must be subject to small quantum fluctuations. All types of matter are subject to such fluctuations, and for most purposes they are so small as to be totally insignificant. The point, however, is that the huge expansion the universe goes through during inflation may magnify these initially insignificant microscopic fluctuations into macroscopic fluctuations in density. Indeed, detailed calculations showed that, subject to certain natural assumptions about the scalar field, the resultant density fluctuations were of the type suggested by Harrison and Zel'dovich.

In a single stroke, therefore, inflation appeared to solve a number of important cosmological problems. However, the role of initial conditions in this "solution" to these problems should be understood. The HBB model can explain the observed features of the universe only with extremely unnatural fine-tuning of initial conditions: Finding the present universe would be as unlikely as finding a pencil balanced on its point after an earthquake. The IUS relieves the universe of such extreme dependence on initial conditions, in that it allows the presently observed state of the universe to arise from a much broader, far more plausible set of initial conditions. Nevertheless, the presently observed state of the universe is not relieved from *all* dependence on initial conditions (see Fig.1).

Take, for example, the density fluctuations problem. It is certainly one of the successes of the IUS that it explains their *origin*. But more detailed calculation of these fluctuations involves certain assumptions about the initial state of the inhomogeneous parts of the scalar field before inflation. These assumptions are admittedly very natural, but they are assumptions nevertheless, and the correct density fluctuation spectrum cannot be obtained without them. More importantly, the occurence of inflation itself depends on initial conditions – it occurs only if the scalar field begins with a large, approximately constant energy density. Like it or not, therefore, the success of the IUS is contingent upon certain assumptions about initial conditions.

Where do these assumptions come from? Obviously one can go on asking an almost infinite sequence of such questions, like the overbearingly curious child during the "Why" stage. But the cosmologist seeking a complete explanation of the observed state of the universe is ultimately compelled to ask, "What happened before inflation? How did the universe actually *begin*?"

The present successful description of the universe with IUS followed by HBB involves an expanding universe. Following this expansion backwards in time to the pre-inflationary era, we find that the size of the universe tends towards zero, and the strength of the gravitational field and the energy density of matter tend towards infinity – *i.e.* the universe appears to have emerged from an *initial singularity*, a region of infinite curvature and energy density at which the known laws of physics break down. Such singularities in cosmological models are not artefacts of unfortunate or inaccurate modeling of the universe, but are *generic*. This is a consequence of the famous "singularity theorems", proved in the 1960's by Stephen Hawking, of the University of Cambridge, and Roger Penrose, of the University of Oxford. What these theorems show is that, under very reasonable assumptions, *any* model of the expanding universe extrapolated backwards in time will encounter an initial singularity (see Fig.2).

The physical implication of the singularity theorems is not, however, that a singularity actually occurs in the real universe, but rather, that the theory predicting them – classical general relativity – breaks down at very high curvatures, and must be superseded by some, bigger, better, more powerful theory. What could this more powerful theory be?

A line of reasoning known as dimensional analysis yields some clues. Physics has three fundamental constants which control the *scale* of physical phenomena. They are the velocity of light, c, which sets the scale of relativistic effects, Newton's constant, G, which is a measure of the strength of gravitational effects, and Planck's constant, h, which controls the scale of quantum effects. Around the beginning of this century, Max Planck, the forefather of quantum theory, showed that these three constants may be combined to produce fundamental units of length, time, and mass. They are known, respectively, as the Planck length, denoted ℓ_p, with a value of approximately 10^{-33} centimetres, the Planck time, $t_p \approx 10^{-45}$ seconds, and the Planck mass, $m_p \approx 10^{-5}$ grams.

The Planck length and time are almost unbelievably small – smaller to the atomic scale than the atomic scale is to the laboratory scale. The Planck mass may seem to be an unremarkable number (it is the approximate mass of a cell) but it must be compared to the typical mass scales of elementary particle physics: It is some 10^{19} times greater. The scales indicated by Planck's units are therefore very extreme indeed. Their significance is that they are the length, time, and mass scales at which relativistic, gravitational, and quantum effects become simultaneously comparable. Such scales could never be achieved in any laboratory situation, not even in the most powerful particle physics accelerator. The key point, however, is that these scales *are* approached in the neighbourhood of the initial singularity in the universe. This suggests that physics in the neighbourhood of the initial singularity is best-described using a theory in which relativistic, gravitational, and quantum effects are combined.

General relativity is a theory in which relativistic and gravitational effects are already combined. What is needed, therefore, is a quantized version of general relativity – a *quantum theory of gravity*.

Quantum theory was originally developed in an attempt to account for phenomena whose description appeared to lie beyond the scope of classical mechanics. Classical physics failed to account, for example, for the structure of the atom. Experiments suggested that the atom consisted of a central nucleus with the electrons somehow in orbit about it, like planetary orbits about the sun. Attempts to describe this model mathematically using the classical theories of Newtonian mechanics and Maxwell's electrodynamics led to failure – there was nothing to hold the electrons in orbit, and they would radiate their energy as electromagnetic radiation, spiraling into the nucleus in a fraction of a second.

With this sort of difficulty in mind, quantum mechanics was developed first by Niels Bohr in 1912, and subsequently by Erwin Schrödinger, Werner Heisenberg, Paul Dirac and others in the 1920's. In quantum mechanics, motion is not deterministic, as in classical mechanics, but *probabilistic*. The dynamical variables used to describe a system in classical mechanics, such as position and momentum, cannot in general be ascribed definite values in quantum mechanics. Instead, the fundamental notion in quantum mechanics is that all systems, such as point particles, are fundamentally wavelike in nature. Mathematically, they are taken to be described by a quantity called a *wave function* into which is encoded probabilistic information about positions, momenta, energies *etc.*. The wave function for a particular system is found by solving an equation called the Schrödinger equation.

For the case of a single point particle, the wave function may be thought of as an oscillating field spread throughout physical space. At each point in space it has an amplitude and a wavelength. The square of the amplitude is proportional to the probability of finding the particle at that position; the wavelength, for constant amplitude wave functions, is related to the momentum of the particle. The particle will therefore have a definite position if the wave function is tightly bunched about a particular point in space; and it will have definite momentum if the wavelength and amplitude of the wave function are uniform throughout all of space (see Fig.3). Typical wave functions for a system will not, however, be of either of these types, and there will be a certain amount of indefiniteness, or *uncertainty* in both position and momentum. In particular, because of the mutually exclusive types of wave function required for definite position and definite momentum, position and momentum cannot be definite simultaneously. This is *Heisenberg's uncertainty principle*, and is an elementary consequence of the wavelike nature of particles.

Quantum mechanics exhibits many phenomena that are qualitatively very different to those exhibited by classical mechanics. Two particular quantum phenomena are relevant to this account. The first is the fact that in quantum mechanics a system can never have an energy of exactly zero. The total energy is generally a sum of two positive parts: The first, the kinetic energy, depends on the momentum, and the second, the potential energy, depends on position. Since the uncertainty principle forbids both momentum and position to be definite simultaneously, it is impossible to say that both the kinetic *and* the potential energy are zero exactly. The system always has a state, called the *ground state*, in which the energy is as small as it can be, consistent with the uncertainty principle. These "ground state fluctuations" are important for galaxy formation in the IUS, as mentioned above.

The second feature of interest is the phenomenon of "tunneling". In classical mechanics, a particle traveling with fixed energy cannot get through an energy barrier. For example, a ball at rest in a bowl will never be able to get out. In quantum mechanics, by contrast, because position is generally not a sharply defined quantity but has a spread over a (typically infinite) range, there is a finite probability that the particle will be found on the other side of the barrier. One says that the particle may "tunnel" through the barrier. This purely quantum mechanical effect is responsible for alpha-decay in radioactive atoms: There, the alpha particle is trapped by an energy barrier inside the nucleus, but manages to escape by tunneling. The tunneling process should not, however, be thought of as actually occuring in real time. In fact, in a certain well-defined mathematical sense, the particle is conveniently thought of as penetrating the barrier not in real time, but in "imaginary" time (time multiplied by the square root of minus one).

These distinctly quantum-mechanical effects should not be thought of as being in contradiction with classical physics. Rather, the point of view typically taken is that quantum mechanics is a *broader* theory than classical mechanics and supersedes it as the correct description of nature. On macroscopic scales, it is usually argued that the wavelike nature of particles is highly suppressed, and quantum mechanics reproduces the effects of classical mechanics to a very high degree of precision (although the exact manner in which this "quantum to classical transition" comes about is still a matter of current research). On microsopic scales, however, the predictions of quantum mechanics depart quite radically from those of classical mechanics. For example, the orbiting electron model of the atom discussed earlier is saved from the rather singular behaviour predicted by classical mechanics by virtue of the phenomenon of ground state fluctuations: The electrons have an orbit of minimum energy from which they cannot emit radiation and fall into the nucleus without violating the uncertainty principle.

The important point to take away from this discussion of quantum mechanics, therefore, is this: In situations where classical physics predicts singular

behaviour, the broader viewpoint of quantum mechanics may indicate that classical mechanics is invalid there. A quantum treatment may replace the singular behaviour with regular behaviour, describable well-within the laws of quantum physics. It is this observation that suggests that singularities in classical cosmology could be alleviated by quantum theory.

A modest approach to this issue might involve allowing the gravitational and matter fields to be classical for much of the evolution of the universe and investigating quantum effects only in the immediate vicinity of the initial singularity. A far more comprehensive, far more complete approach is to apply quantum theory to the entire universe at all times and everything it contains – *i.e.* to apply it to all the gravitational and matter fields describing the universe. To do this is to do quantum cosmology.

In quantum cosmology, as for a simple system such as an atom in ordinary quantum mechanics, the fundamental description of the system – the entire universe – is in terms of a wave function, "the wave function of the universe". This wave function is found by solving an equation which is the cosmological analogue of the Schrödinger equation, and is known as the Wheeler-DeWitt equation. Given such a wave function one may in principle extract probabilistic predictions concerning the outcome of observation. In the simplest models of the universe, the analogue of position is the spatial size of the universe, and the analogue of momentum is the rate of expansion of the universe.

Many conceptual and technical difficulties arise in quantum cosmology, above and beyond those in quantum mechanics. The first and most serious difficulty is the lack of a complete, manageable quantum theory of gravity directly applicable to cosmology. Three of the four fundamental forces of nature, the electromagnetic, weak nuclear, and strong nuclear forces have been made consistent with quantum theory – quantized – in a satisfactory fashion. By contrast, all attempts to quantize Einstein's general relativity directly have met with failure. It should be remarked that there exists a theory, called the theory of "superstrings", which is claimed to be a consistent unified quantum theory of all four forces of nature, and thus is, or at least contains, a quantum theory of gravity. Final judgement on superstring theory is yet to be made, however, and although it may well be the sought-after holy grail of a quantum theory of gravity, it is a long way off being a manageable theory directly applicable to cosmology.

From the very beginning, therefore, quantum cosmology rests on very shaky foundations. It is possible to develop a formalism, and perform some rather crude calculations, and this in practice is what one actually does in quantum cosmology. But unlike, for example, the quantized theory of the electromagnetic field, there is no systematic scheme for calculating the quantities of interest to arbitrary precision, and it is difficult to know whether one's calculations have anything to do with a consistent quantum gravity theory, should it exist, let alone with observation.

Not least of all for these reasons, much of the work carried out in quantum cosmology has focussed on *issues* rather than detailed technicalities. Undoubtedly the most difficult, interesting, and important issues concern the application of quantum mechanics to cosmology. There are at least two issues here.

The first issue concerns the *applicability* of quantum mechanics to the entire universe. Quantum mechanics was developed to describe atomic scale phenomena which were totally inexplicable from the perspective of classical mechanics. The beautiful agreement of quantum mechanics with experiment is one of the great triumphs of modern physics, and no physicist in his or her right mind is in any doubt as to its correctness on the atomic scale. One may, however, find a few dissenters if one suggested that quantum mechanics is equally applicable to the everyday macroscopic scale, *e.g.* to tables and chairs. It is very hard to dispute this because, as we have already indicated, the predictions of quantum mechanics coincide very closely with those of classical mechanics on this scale, and genuinely macroscopic quantum mechanical effects are extremely difficult to detect experimentally. Even more contentious is the most extravagant extrapolation one could conceivably make: That quantum mechanics applies to the entire universe at all times and to everything in it. Acceptable or not, this is the fundamental assertion of quantum cosmology.

The second issue, and perhaps the most difficult one, concerns the *interpretation* of quantum mechanics as applied to cosmology. Because quantum mechanics generally prohibits the familiar variables of classical mechanics from taking definite values, it was found necessary in the development of the subject to supplement the formalism with an extra structure to translate the mathematical formalism into statements about what one would actually observe on making a measurement. The foundations of this structure, known as "quantum measurement theory", were laid primarily by Bohr in the 1920's and 30's. Central to Bohr's scheme is the notion that the observer plays a crucial role in the act of measurement. He assumed that the world may be divided into two parts: Microscopic systems such as atoms, governed purely by quantum mechanics, and external macroscopic systems, such as observers and their measuring apparata, governed by classical mechanics. A measurement is an interaction between the observer (or his or her measuring device) and the microscopic system, leading to a permanent recording of the event. During this interaction, the wave function describing the microscopic system undergoes a discontinuous change from whatever initial state it was in, to a final state in which it is definite in the quantity that is being measured. This discontinuous change is referred to, rather dramatically, as "the collapse of the wave function". For example, the

wave function could start out in a state of definite momentum, but if position is being measured it "collapses" into a state of definite position.

Although this scheme, known as the "Copenhagen interpretation" of quantum mechanics, is felt by many to be deeply unsatisfactory from a philosophical point of view, it cannot be denied that it has been thoroughly successful in allowing predictions to be extracted from the theory – predictions which agree with experiment. It is perhaps for this reason that the Copenhagen interpretation stood largely unchallenged for almost half a century. In attempting to apply quantum mechanics to the entire universe, however, one meets with acute difficulties which cannot be brushed off as philosophical niceties. In a theory of the entire universe, of which the observer is part, there should be no fundamental division between observer and observed. Moreover, although it is almost acceptable to believe that it happens at the atomic scale, most physicists feel very uncomfortable at the thought of the wave function of the entire universe collapsing when an observation of the universe is made. Then there is the question of probabilisitc predictions. What do these mean in quantum cosmology? In ordinary quantum mechanics, one envisages an ensemble of many identical systems, such as atoms, and probabilisitic predictions may be tested by making a large number of measurements. But what do probabilities mean when we have just one system which we measure once?

One of the first physicists to take very seriously the notion that one should apply quantum mechanics to the entire universe was the late Hugh Everett III, of Princeton. In his 1957 paper, he presented a framework for the interpretation of quantum mechanics particularly suited to the special needs of cosmology. Everett asserted that the universal wave function should describe macroscopic observers and microscopic systems alike, and that there should be no fundamental division between them. He further asserted that there is no discontinuous change brought about by the collapse of the wave function during the measurement process; only smooth evolution described by the Schrödinger equation. A measurement is just an interaction between different parts of the entire universe, and the wave function should predict what one part of the system "sees" when it observes the other part. But in modeling the measurement process, Everett discovered a truly remarkable thing: The measurement appears to cause the universe, to "split" into sufficiently many copies of itself to take account of all possible outcomes of the measurement.

The reality of these multiple copies of the universe in this severely uneconomical aspect of Everett's "Many Worlds" interpretation of quantum mechanics has been hotly debated. Modern versions of this interpretation, notably that of Murray Gell-Mann of Caltech, and James B.Hartle of the University of California at Santa Barbara, play down the "many worlds" aspect of the theory and talk instead about "decoherent histories". These are possible histories for the universe to which probabilites may be assigned. For practical purposes, it does not matter whether one thinks of all or just one of them as actually happening. Gell-Mann and Hartle also address the problem of understanding probabilistic predictions for the entire universe. They insist that the only probabilities that have any meaning in quantum cosmology are *a priori* probabilities. These are probabilties that are very close to one or zero – definite yes-no predictions. Although most probabilistic predictions are not of this type, they can often be made so by suitable modification of the question one is interested in. Unlike quantum mechanics, therefore, in which the goal is to determine probabilities for the possible outcomes of given observations, the goal in quantum cosmology is to determine those observations for which the theory gives probabilities close to zero or one.

Research on this fascinating topic is still in progress, but for the purposes of practical quantum cosmology, the following understanding has emerged: In certain regions, typically – but not always – when the universe is large, the wave function for the universe indicates that the universe behaves classically to a very high degree of precision. In this case we would say that classical spacetime is a prediction of the theory. Under these conditions, moreover, the wave function provides probabilities for the set of possible classical behaviours of the universe. There will, on the other hand, be other regions in which no such prediction is made. In such regions the notions of space and time quite simply do not exist – there is just a "quantum fuzz", still describable by known laws of quantum physics, but not describable by *classical* laws (see Fig.4). Such regions, as we will discuss below, may surround singularities.

Armed with a quantum theory of cosmology, and a means of interpreting it, one may begin to ask about the beginning of the universe. It should, however, be emphasized that the problem of initial conditions is not actually *solved* by going to quantum cosmology. It certainly puts one in a much better position to address the issue than in classical cosmology, in that there is no longer any worry that one might be trying to impose classical initial conditions in a region in which classical physics is not valid, such as near the initial singularity. But the question of classical initial conditions becomes one of quantum initial conditions. It has become the following question: Of the many possible wave functions permitted by the dynamics (*i.e.* of the many possible solutions to the Wheeler-DeWitt equation), how is just one singled out?

The cosmological situation should be contrasted with the laboratory situation, to which most of physics is directed. There, one has a physical system whose temporal and spatial boundaries are generally quite clearly defined. At those boundaries the experimenter may control, or at least observe, the physical conditions. Given suitable laws of physics, these initial or boundary conditions may be evolved in space and time, thereby predicting the physical state inside

the system under consideration. This may then be compared with experiment or observation.

In cosmology, on the other hand, the system under scrutiny is the entire universe. By definition it has no exterior, no outside world, no "rest of the universe" off onto which specification of boundary or initial conditions may be passed (see Fig.5). It is, therefore, the almost inescapable task of the quantum cosmologist to *propose* laws of initial or boundary conditions for the universe, in very much the same way that it is the task of the theoretical physicist to propose laws of physics to govern the evolution of physical systems. Like such laws of evolution, the ultimate test of any law of initial conditions will be whether or not its predictions agree with observation.

There is a possible escape clause, relieving the quantum cosmologist of his or her onerous task, which is why the word "almost" appears in the above paragraph. In his seminal 1967 paper, Bryce DeWitt expressed the hope that mathematical consistency alone would lead to a unique solution to the Wheeler-DeWitt equation, *i.e.* to a unique wave function for the universe. Such a possibility is entirely without precedent in theoretical physics, and there is unfortunately little indication that DeWitt's hope will ever be fulfilled.

It was the realization, by Hartle and Hawking, by Linde, and by Vilenkin, that one had to face up squarely to the issue of initial conditions in quantum cosmology that led to the revitalization of the subject in the early 1980's alluded to at the beginning of this article. These workers made quite definite proposals that were intended to pick out a particular solution to the Wheeler-DeWitt equation, *i.e.* to single out a unique wave function for the universe.

The proposal made by Hartle and Hawking is known as the "no-boundary" proposal. It defines a particular wave function of the universe using a rather elegant formulation of quantum mechanics developed in the 1940's by the late Richard Feynman of Caltech, referred to as the "path integral" or "sum-over-histories" method. The proposal made, independently by Linde, by Vilenkin and others, is known as the "tunneling proposal", and is a particular way of picking out a solution to the Wheeler-DeWitt equation. Each of these proposals picks out a unique wave function for the universe (contingent, however, on the resolution of a number of technical difficulties recently exposed by James B. Hartle, Jorma Louko and I). In both of these proposals, the wave function indicates that spacetime is classical when the universe is reasonably large (typically larger than a few thousand Planck lengths), in agreement with observation. When the universe is very small, however, classical spacetime is not indicated. In fact, both of these proposals view the beginning of the universe as being very much like a quantum tunneling event – the universe "tunnels" to finite size, starting from zero size *i.e.* it starts from "nothing" (see Fig.6). These proposals therefore have the very desirable feature that the point that was an initial singularity in classical cosmology is surrounded by some kind of fuzzy quantum region.

Given these quantum theories of the creation of the universe, one may finally ask, "How did the universe actually begin?" The response of the quantum cosmologist may be something of a dissappointment. Rather than answer the question, he or she would declare the question disqualified. In the neighbourhood of singularities, such as the initial singularity, the wave functions given by the tunneling and no-boundary proposals indicate that classical general relativity is not valid, and furthermore, that notions of space and time are inappropriate. One cannot, therefore, ask questions involving notions of time or space. The picture that emerges is of the universe appearing from a quantum fuzz with non-zero size and finite (rather than infinite) energy density (see Fig.7).

In the subsequent classical evolution after quantum creation, the wave function provides a probability measure on different possible evolutions, and it is at this point that the no-boundary and tunneling wave functions differ. Workers in the field are not in complete agreement about this, but the current understanding appears to be that the tunneling wave function gives high probability for universes which undergo an initial period of inflation of sufficient duration to solve the horizon and flatness problems; the no-boundary proposal, on the other hand, does predict inflation, but does not appear to predict a ***sufficient*** amount of inflation.

The no-boundary and tunneling proposals also make predictions about the emergence of large scale structure. Recall that in the inflationary universe scenario, although the origin of the initial density perturbation spectrum is explained, the exact form and magnitude depends on certain assumptions about the initial state of the scalar field. To be precise, the necessary assumption is essentially that the inhomogeneous parts of the scalar field start out in their quantum mechanical ground state – the lowest possible energy state consistent with the uncertainty principle. This is a very natural assumption, but it is an assumption nevertheless. In a 1985 paper, Stephen Hawking and I demonstrated this assumption to be a *consequence* of the no-boundary proposal.

In brief, therefore, the achievements of quantum cosmology may be summarized as follows. Firstly, it allows one to determine the conditions under which the notions of classical space and time are valid, and thus the conditions under which classical theories such as general relativity may be used. Secondly, it permits the list of assumptions necessary for conventional cosmology consisting of inflation followed by a hot big bang to be compressed into a simple law of boundary conditions for the wave function of the universe.

It may be that the no boundary proposal or the tunneling proposal is the correct boundary conditions for the wave function of the universe. Unfortunately, the severe problems of actually testing quantum cosmology observationally make this very difficult to check. It could be a very long time, therefore, before we can tell whether either of these proposals are an answer to the question at the beginning of this article, "Where did all this come from?" Nevertheless, through quantum cosmology, we have at least been able to formulate and address the question in a meaningful way.

ACKNOWLEDGEMENTS

I am grateful to numerous friends and colleagues for their comments on the first draft of this article, including Fay Dowker, Luis Garay, Jorma Louko, Guillermo Mena, Nina Shopalovich, Desmond Smith and Eve Sullivan.

This work was supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under contract No.DE-AC02-76ER03069.

FURTHER READING

The literature on quantum cosmology is vast. One of the seminal papers is B.S.DeWitt, "Quantum Theory of Gravity I: The Canonical Theory", *Phys.Rev.* **160**, 1113 (1967). One of the papers that sparked off the more recent interest in the subject is, J.B.Hartle and S.W.Hawking, "The Wave Function of the Universe", *Phys.Rev.* **D28**, 2960 (1983). An extensive bibliography of papers is, J.J.Halliwell, *Int.J.Mod.Phys.* **A5**, 2473 (1990). For a pedagogical introduction to quantum cosmology, see J.J.Halliwell, "Introductory Lectures on Quantum Cosmology", CTP preprint # 1845 (1990), to appear in *Proceedings of the Seventh Jerusalem Winter School on Theoretical Physics: Quantum Cosmology and Baby Universes*, eds. S.Coleman, J.B.Hartle and T.Piran (1991). See also in the same volume, the lectures by J.B.Hartle, "The Quantum Mechanics of Cosmology" (Santa Barbara preprint UCSBTH90-31, 1990).

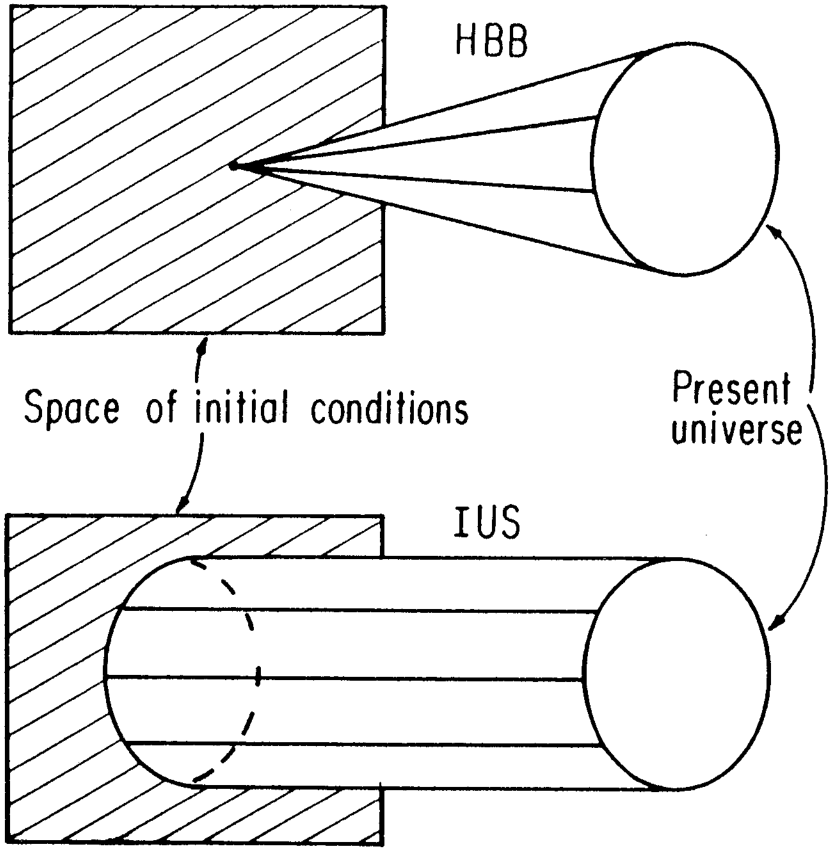


Figure 1.

A schematic figure illustrating the sets of initial conditions that can lead to the present universe in the HBB model and in the IUS. In the HBB model the present universe can arise, but only from an exceedingly small set of initial conditions. The IUS permits the present universe to arise from a much larger, far more plausible set of initial conditions, but it does not mean that the present universe could arise from *any* initial state.

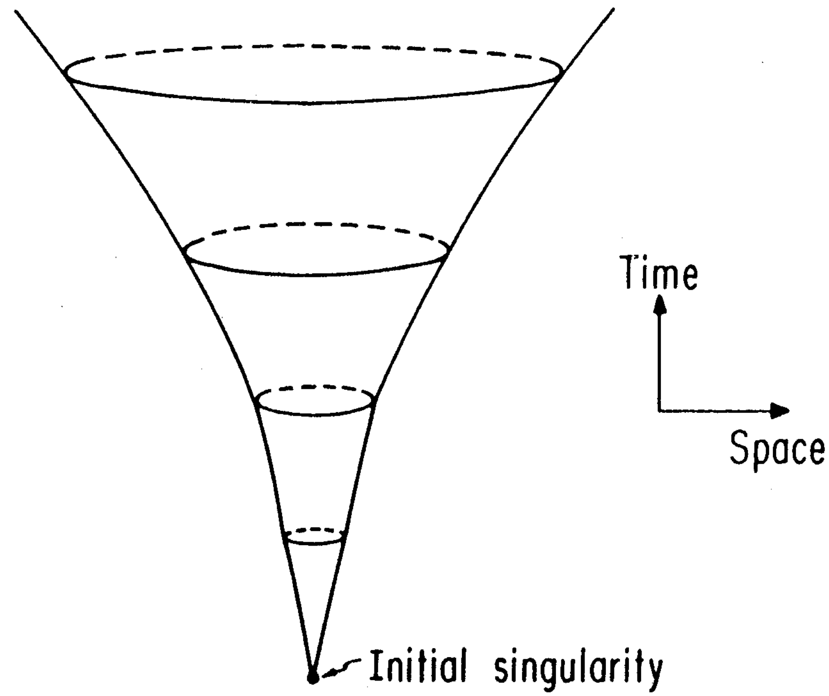


Figure 2.

The singuarity theorems of Hawking and Penrose indicate that any reasonable model of the expanding universe will generally encounter a singularity when evolved backwards in time. At such a point, the curvature of the spacetime becomes infinite, the known laws of physics break down, and predictability is completely lost.

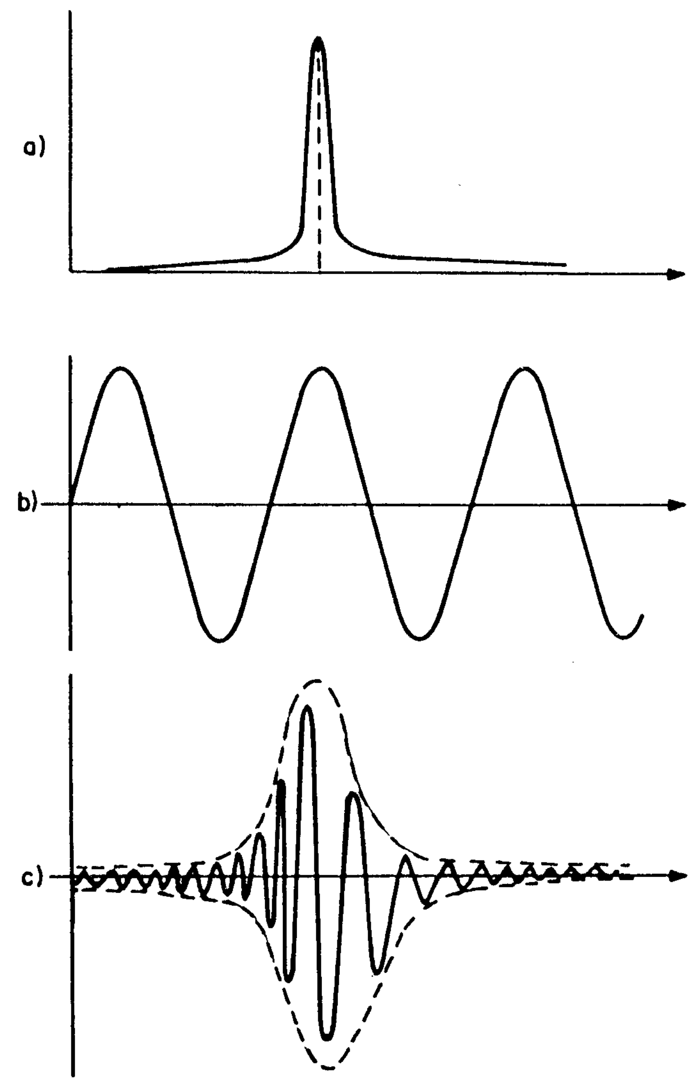
Figure 3

Plots of different types of wave functions in quantum mechanics. The plot is of the wave function against position in real physical space.

Figure (3a) shows the wave function for a particle in a state of definite position. The wave function is very sharply peaked about a particular point in space. The uncertainty in momentum, however, is essentially infinite, in that there is equal probability of finding it with any wavelength.

Figure (3b) shows the wave function for a particle in a state of definite momentum. The wave functions has a definite wavelength and constant amplitude throughout all of space. The particle position, however, is completetly uncertain, in that there is equal probability of finding it to be anywhere.

Figure (3c) shows a wave function for a particle in a state referred to as a "coherent" state, which is the perfect compromise between definite position and definite momentum. There is uncertainty in both position and momentum, but the uncertainty is as small as it can be consistent with Heisenberg's uncertainty principle.



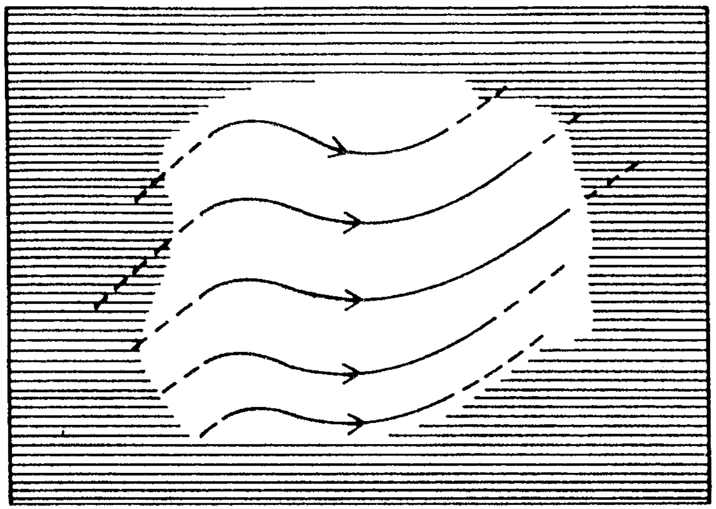


Figure 4.

A schematic picture showing the behaviour of a typical wave function, as a function of, for example, the size of the universe, the energy density of matter *etc.* In certain regions the wave function indicates that the notion of classical spacetime is an appropriate one. Precisely, the wave function becomes strongly peaked about sets of classical histories for the universe. These are denoted by the bold lines in the figure. There will, however, be other regions in which classical spacetime is not indicated to be a valid notion. These regions of "quantum fuzz" are denoted by the shaded region in the figure. They are still describable by quantum laws of physics, but not describable by classical laws.

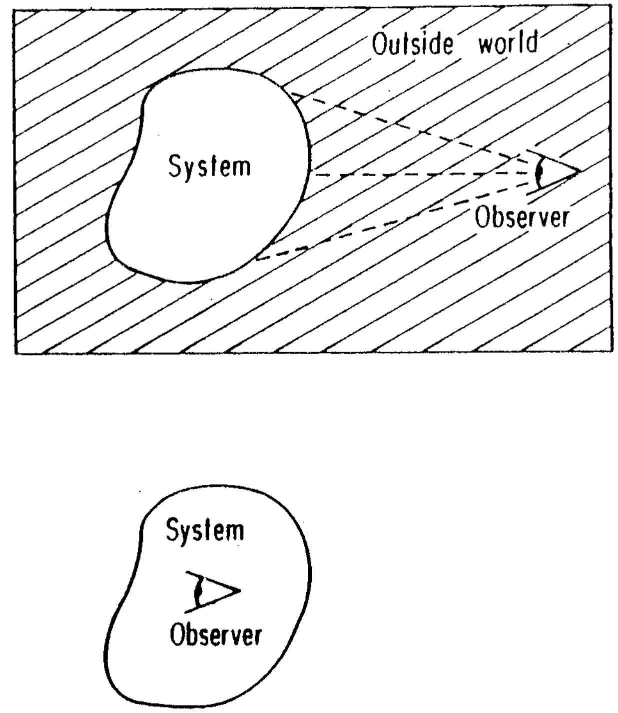


Figure 5.

The difference between laboratory physics and cosmology. In laboratory physics the observer, who sits outside the system, may control (or at least observe) the external conditions, and these are used as the boundary conditions when determining what is going on inside the system. In cosmology, on the other hand, not only is the observer on the inside, but there is no outside world off onto which the specification of boundary or initial conditions can be passed.

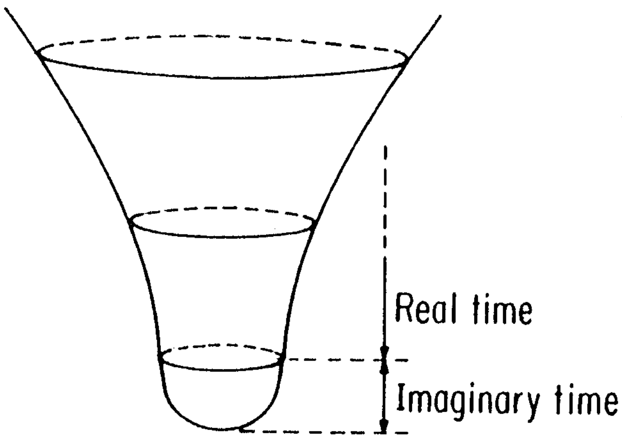


Figure 6.

The universe tunneling from nothing, as in the no-boundary and tunneling proposals for the wave function of the universe. The universe "starts" at zero size in a perfectly smooth way, and evolves in imaginary time to a size of a few Planck lengths. Taking that small but finite size as its initial condition, the universe evolves thereafter in real, physical time. Although everything now appears to be smooth and non-singular, the singularity theorems are not violated because the "smoothing off" of Figure 2 is in imaginary time, but the singularity theorems refer to real time only.

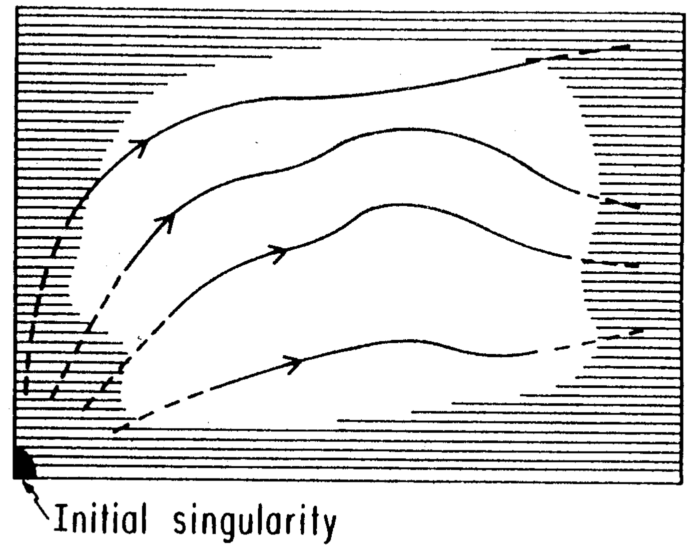


Figure 7.

A schematic picture showing the possible classical (real time) histories for the universe indicated by the no-boundary and tunneling wave functions. The initial singularity is surrounded by a quantum fuzz, denoted by the shaded region, (which actually corresponds to the imaginary time region in Figure 6), where no notions of space and time exist. An observer in the present looking back in time would see the classical histories for the universe (denoted by the bold lines) emerging from this quantum fuzz at finite size in a non-singular way.