The Beginning of Time

In this lecture, I would like to discuss whether time itself has a beginning, and whether it will have an end. All the evidence seems to indicate, that the universe has not existed forever, but that it had a beginning, about 15 billion years ago. This is probably the most remarkable discovery of modern cosmology. Yet it is now taken for granted. We are not yet certain whether the universe will have an end. When I gave a lecture in Japan, I was asked not to mention the possible re-collapse of the universe, because it might affect the stock market. However, I can re-assure anyone who is nervous about their investments that it is a bit early to sell: even if the universe does come to an end, it won't be for at least twenty billion years. By that time, maybe the GATT trade agreement will have come into effect.

The time scale of the universe is very long compared to that for human life. It was therefore not surprising that until recently, the universe was thought to be essentially static, and unchanging in time. On the other hand, it must have been obvious, that society is evolving in culture and technology. This indicates that the present phase of human history can not have been going for more than a few thousand years. Otherwise, we would be more advanced than we are. It was therefore natural to believe that the human race, and maybe the whole universe, had a beginning in the fairly recent past. However, many people were unhappy with the idea that the universe had a beginning, because it seemed to imply the existence of a supernatural being who created the universe. They preferred to believe that the universe, and the human race, had existed forever. Their explanation for human progress was that there had been periodic floods, or other natural disasters, which repeatedly set back the human race to a primitive state.

This argument about whether or not the universe had a beginning, persisted into the 19th and 20th centuries. It was conducted mainly on the basis of theology and philosophy, with little consideration of observational evidence. This may have been reasonable, given the notoriously unreliable character of cosmological observations, until fairly recently. The cosmologist, Sir Arthur Eddington, once said, 'Don't worry if your theory doesn't agree with the observations, because they are probably wrong.' But if your theory disagrees with the Second Law of Thermodynamics, it is in bad trouble. In fact, the theory that the universe has existed forever is in serious difficulty with the Second Law of Thermodynamics. The Second Law, states that disorder always increases with time. Like the argument about human progress, it indicates that there must have been a beginning. Otherwise, the universe would be in a state of complete disorder by now, and everything would be at the same temperature. In an infinite and everlasting universe, every line of sight would end on the surface of a star. This would mean that the night sky would have been as bright as the surface of the Sun. The only way of avoiding this problem would be if, for some reason, the stars did not shine before a certain time.

In a universe that was essentially static, there would not have been any dynamical reason, why the stars should have suddenly turned on, at some time. Any such "lighting up time" would have to be imposed by an intervention from outside the universe. The situation was different, however, when it was realised that the universe is not static, but expanding. Galaxies are moving steadily apart from each other. This means that they were closer together in the past. One can plot the separation of two galaxies, as a function of time. If there were no acceleration due to gravity, the graph would be a straight line. It would go down to zero separation, about twenty billion years ago. One would expect gravity, to cause the galaxies to accelerate towards each other. This will mean that the graph of the separation of two galaxies will bend downwards, below the straight line. So the time of zero separation, would have been less than twenty billion years ago.

At this time, the Big Bang, all the matter in the universe, would have been on top of itself. The density would have been infinite. It would have been what is called, a singularity. At a singularity, all the laws of physics would have broken down. This means that the state of the universe, after the Big Bang, will not depend on anything that may have happened before, because the deterministic laws that govern the universe will break down in the Big Bang. The universe will evolve from the Big Bang, completely independently of what it was like before. Even the amount of matter in the universe, can be different to what it was before the Big Bang, as the Law of Conservation of Matter, will break down at the Big Bang.

Since events before the Big Bang have no observational consequences, one may as well cut them out of the theory, and say that time began at the Big Bang. Events before the Big Bang, are simply not defined, because there's no way one could measure what happened at them. This kind of beginning to the universe, and of time itself, is very different to the beginnings that had been considered earlier. These had to be imposed on the universe by some external agency. There is no dynamical reason why the motion of bodies in the solar system can not be extrapolated back in time, far beyond four thousand and four BC, the date for the creation of the universe, according to the book of Genesis. Thus it would require the direct intervention of

God, if the universe began at that date. By contrast, the Big Bang is a beginning that is required by the dynamical laws that govern the universe. It is therefore intrinsic to the universe, and is not imposed on it from outside.

Although the laws of science seemed to predict the universe had a beginning, they also seemed to predict that they could not determine how the universe would have begun. This was obviously very unsatisfactory. So there were a number of attempts to get round the conclusion, that there was a singularity of infinite density in the past. One suggestion was to modify the law of gravity, so that it became repulsive. This could lead to the graph of the separation between two galaxies, being a curve that approached zero, but didn't actually pass through it, at any finite time in the past. Instead, the idea was that, as the galaxies moved apart, new galaxies were formed in between, from matter that was supposed to be continually created. This was the Steady State theory, proposed by Bondi, Gold, and Hoyle.

The Steady State theory, was what Karl Popper would call, a good scientific theory: it made definite predictions, which could be tested by observation, and possibly falsified. Unfortunately for the theory, they were falsified. The first trouble came with the Cambridge observations, of the number of radio sources of different strengths. On average, one would expect that the fainter sources would also be the more distant. One would therefore expect them to be more numerous than bright sources, which would tend to be near to us. However, the graph of the number of radio sources, against there strength, went up much more sharply at low source strengths, than the Steady State theory predicted.

There were attempts to explain away this number count graph, by claiming that some of the faint radio sources, were within our own galaxy, and so did not tell us anything about cosmology. This argument didn't really stand up to further observations. But the final nail in the coffin of the Steady State theory came with the discovery of the microwave background radiation, in 1965. This radiation is the same in all directions. It has the spectrum of radiation in thermal equilibrium at a temperature of 2 point 7 degrees above the Absolute Zero of temperature. There doesn't seem any way to explain this radiation in the Steady State theory.

Another attempt to avoid a beginning to time, was the suggestion, that maybe all the galaxies didn't meet up at a single point in the past. Although on average, the galaxies are moving apart from each other at a steady rate, they also have small additional velocities, relative to the uniform expansion. These so-called "peculiar velocities" of the galaxies, may be directed sideways to the main expansion. It was argued, that as you plotted the position of the galaxies back in time, the sideways peculiar velocities, would have meant that the galaxies wouldn't have all met up. Instead, there could have been a previous contracting phase of the universe, in which galaxies were moving towards each other. The sideways velocities could have meant that the galaxies didn't collide, but rushed past each other, and then started to move apart. There wouldn't have been any singularity of infinite density, or any breakdown of the laws of physics. Thus there would be no necessity for the universe, and time itself, to have a beginning. Indeed, one might suppose that the universe had oscillated, though that still wouldn't solve the problem with the Second Law of Thermodynamics: one would expect that the universe would become more disordered each oscillation. It is therefore difficult to see how the universe could have been oscillating for an infinite time.

This possibility, that the galaxies would have missed each other, was supported by a paper by two Russians. They claimed that there would be no singularities in a solution of the field equations of general relativity, which was fully general, in the sense that it didn't have any exact symmetry. However, their claim was proved wrong, by a number of theorems by Roger Penrose and myself. These showed that general relativity predicted singularities, whenever more than a certain amount of mass was present in a region. The first theorems were designed to show that time came to an end, inside a black hole, formed by the collapse of a star. However, the expansion of the universe, is like the time reverse of the collapse of a star. I therefore want to show you, that observational evidence indicates the universe contains sufficient matter, that it is like the time reverse of a black hole, and so contains a singularity.

In order to discuss observations in cosmology, it is helpful to draw a diagram of events in space and time, with time going upward, and the space directions horizontal. To show this diagram properly, I would really need a four dimensional screen. However, because of government cuts, we could manage to provide only a two dimensional screen. I shall therefore be able to show only one of the space directions.

As we look out at the universe, we are looking back in time, because light had to leave distant objects a long time ago, to reach us at the present time. This means that the events we observe lie on what is called our past light cone. The point of the cone is at our position, at the present time. As one goes back in time on the diagram, the light cone spreads out to greater distances, and its area increases. However, if there is sufficient matter on our past light cone, it will bend the rays of light towards each other. This will mean that, as one goes back into the past, the area of our past light cone will reach a maximum, and then start to decrease. It is this focussing of our past light cone, by the gravitational effect of the matter in the universe, that is the signal that the universe is within its horizon, like the time reverse of a black hole. If one can determine that there is enough matter in the universe, to focus our past light cone, one can then apply the singularity theorems, to show that time must have a beginning.

How can we tell from the observations, whether there is enough matter on our past light cone, to focus it? We observe a number of galaxies, but we can not measure directly how much matter they contain. Nor can we be sure that every line of sight from us will pass through a galaxy. So I will give a different argument, to show that the universe contains enough matter, to focus our past light cone. The argument is based on the spectrum of the microwave background radiation. This is characteristic of radiation that has been in thermal equilibrium, with matter at the same temperature. To achieve such an equilibrium, it is necessary for the radiation to be scattered by matter, many times. For example, the light that we receive from the Sun has a characteristically thermal spectrum. This is not because the nuclear reactions, which go on in the centre of the Sun, produce radiation with a thermal spectrum. Rather, it is because the radiation has been scattered, by the matter in the Sun, many times on its way from the centre.

In the case of the universe, the fact that the microwave background has such an exactly thermal spectrum indicates that it must have been scattered many times. The universe must therefore contain enough matter, to make it opaque in every direction we look, because the microwave background is the same, in every direction we look. Moreover, this opacity must occur a long way away from us, because we can see galaxies and quasars, at great distances. Thus there must be a lot of matter at a great distance from us. The greatest opacity over a broad wave band, for a given density, comes from ionised hydrogen. It then follows that if there is enough matter to make the universe opaque, there is also enough matter to focus our past light cone. One can then apply the theorem of Penrose and myself, to show that time must have a beginning.

The focussing of our past light cone implied that time must have a beginning, if the General Theory of relativity is correct. But one might raise the question, of whether General Relativity really is correct. It certainly agrees with all the observational tests that have been carried out. However these test General Relativity, only over fairly large distances. We know that General Relativity can not be quite correct on very small distances, because it is a classical theory. This means, it doesn't take into account, the Uncertainty Principle of Quantum Mechanics, which says that an object can not have both a well defined position, and a well defined speed: the more accurately one measures the position, the less accurately one can measure the speed, and vice versa. Therefore, to understand the very high-density stage, when the universe was very small, one needs a quantum theory of gravity, which will combine General Relativity with the Uncertainty Principle.

Many people hoped that quantum effects, would somehow smooth out the singularity of infinite density, and allow the universe to bounce, and continue back to a previous contracting phase. This would be rather like the earlier idea of galaxies missing each other, but the bounce would occur at a much higher density. However, I think that this is not what happens: quantum effects do not remove the singularity, and allow time to be continued back indefinitely. But it seems that quantum effects can remove the most objectionable feature, of singularities in classical General Relativity. This is that the classical theory, does not enable one to calculate what would come out of a singularity, because all the Laws of Physics would break down there. This would mean that science could not predict how the universe would have begun. Instead, one would have to appeal to an agency outside the universe. This may be why many religious leaders, were ready to accept the Big Bang, and the singularity theorems.

It seems that Quantum theory, on the other hand, can predict how the universe will begin. Quantum theory introduces a new idea, that of imaginary time. Imaginary time may sound like science fiction, and it has been brought into Doctor Who. But nevertheless, it is a genuine scientific concept. One can picture it in the following way. One can think of ordinary, real, time as a horizontal line. On the left, one has the past, and on the right, the future. But there's another kind of time in the vertical direction. This is called imaginary time, because it is not the kind of time we normally experience. But in a sense, it is just as real, as what we call real time.

The three directions in space, and the one direction of imaginary time, make up what is called a Euclidean space-time. I don't think anyone can picture a four dimensional curve space. But it is not too difficult to visualise a two dimensional surface, like a saddle, or the surface of a football.

In fact, James Hartle of the University of California Santa Barbara, and I have proposed that space and imaginary time together, are indeed finite in extent, but without boundary. They would be like the surface of the Earth, but with two more dimensions. The surface of the Earth is finite in extent, but it doesn't have any boundaries or edges. I have been round the world, and I didn't fall off.

If space and imaginary time are indeed like the surface of the Earth, there wouldn't be any singularities in the imaginary time direction, at which the laws of physics would break down.

And there wouldn't be any boundaries, to the imaginary time space-time, just as there aren't any boundaries to the surface of the Earth. This absence of boundaries means that the laws of physics would determine the state of the universe uniquely, in imaginary time. But if one knows the state of the universe in imaginary time, one can calculate the state of the universe in real time. One would still expect some sort of Big Bang singularity in real time. So real time would still have a beginning. But one wouldn't have to appeal to something outside the universe, to determine how the universe began. Instead, the way the universe started out at the Big Bang would be determined by the state of the universe in imaginary time. Thus, the universe would be a completely self-contained system. It would not be determined by anything outside the physical universe, that we observe.

The no boundary condition, is the statement that the laws of physics hold everywhere. Clearly, this is something that one would like to believe, but it is a hypothesis. One has to test it, by comparing the state of the universe that it would predict, with observations of what the universe is actually like. If the observations disagreed with the predictions of the no boundary hypothesis, we would have to conclude the hypothesis was false. There would have to be something outside the universe, to wind up the clockwork, and set the universe going. Of course, even if the observations do agree with the predictions, that does not prove that the no boundary proposal is correct. But one's confidence in it would be increased, particularly because there doesn't seem to be any other natural proposal, for the quantum state of the universe.

The no boundary proposal, predicts that the universe would start at a single point, like the North Pole of the Earth. But this point wouldn't be a singularity, like the Big Bang. Instead, it would be an ordinary point of space and time, like the North Pole is an ordinary point on the Earth, or so I'm told. I have not been there myself.

According to the no boundary proposal, the universe would have expanded in a smooth way from a single point. As it expanded, it would have borrowed energy from the gravitational field, to create matter. As any economist could have predicted, the result of all that borrowing, was inflation. The universe expanded and borrowed at an ever-increasing rate. Fortunately, the debt of gravitational energy will not have to be repaid until the end of the universe.

Eventually, the period of inflation would have ended, and the universe would have settled down to a stage of more moderate growth or expansion. However, inflation would have left its mark on the universe. The universe would have been almost completely smooth, but with very slight irregularities. These irregularities are so little, only one part in a hundred thousand, that for years people looked for them in vain. But in 1992, the Cosmic Background Explorer satellite, COBE, found these irregularities in the microwave background radiation. It was an historic moment. We saw back to the origin of the universe. The form of the fluctuations in the microwave background agree closely with the predictions of the no boundary proposal. These very slight irregularities in the universe would have caused some regions to have expanded less fast than others. Eventually, they would have stopped expanding, and would have collapsed in on themselves, to form stars and galaxies. Thus the no boundary proposal can explain all the rich and varied structure, of the world we live in. What does the no boundary proposal predict for the future of the universe? Because it requires that the universe is finite in space, as well as in imaginary time, it implies that the universe will re-collapse eventually. However, it will not re-collapse for a very long time, much longer than the 15 billion years it has already been expanding. So, you will have time to sell your government bonds, before the end of the universe is nigh. Quite what you invest in then, I don't know.

Originally, I thought that the collapse, would be the time reverse of the expansion. This would have meant that the arrow of time would have pointed the other way in the contracting phase. People would have gotten younger, as the universe got smaller. Eventually, they would have disappeared back into the womb.

However, I now realise I was wrong, as these solutions show. The collapse is not the time reverse of the expansion. The expansion will start with an inflationary phase, but the collapse will not in general end with an anti inflationary phase. Moreover, the small departures from uniform density will continue to grow in the contracting phase. The universe will get more and more lumpy and irregular, as it gets smaller, and disorder will increase. This means that the arrow of time will not reverse. People will continue to get older, even after the universe has begun to contract. So it is no good waiting until the universe re-collapses, to return to your youth. You would be a bit past it, anyway, by then.

The conclusion of this lecture is that the universe has not existed forever. Rather, the universe, and time itself, had a beginning in the Big Bang, about 15 billion years ago. The beginning of real time, would have been a singularity, at which the laws of physics would have broken down. Nevertheless, the way the universe began would have been determined by the laws of physics, if the universe satisfied the no boundary condition. This says that in the imaginary time direction, space-time is finite in extent, but doesn't have any boundary or edge. The predictions of the no boundary proposal seem to agree with observation. The no boundary hypothesis also predicts that the universe will eventually collapse again. However, the contracting phase, will not have the opposite arrow of time, to the expanding phase. So we will keep on getting older, and we won't return to our youth. Because time is not going to go backwards, I think I better stop now.