Black Holes, Cosmology and Space-Time Singularities

Nobel Lecture, December 8, 2020 by
Roger Penrose
University of Oxford, Oxford, United Kingdom.

IN 1908 HERMANN MINKOWSKI introduced the idea of space-time, which is a 4-dimensional space that encapsulates pretty well all of Einstein's 1905 theory of special relativity. At first, Einstein didn't like the idea very much. He initially thought it was just mathematical sophistry or something like that, but subsequently he picked up on it, realizing the power of the 4-dimensional geometrical perspective, and it became central to its generalization to Einstein's *general* theory of relativity. In Fig. 1, I have indicated three coordinate axes for ordinary 3-dimensional space, but in Fig. 2, we move on to introduce a time axis to describe 4-dimensional space-time.

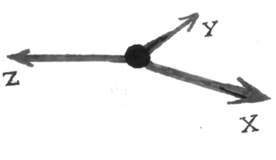

Figure 1.

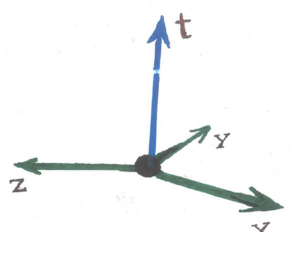

Figure 2.

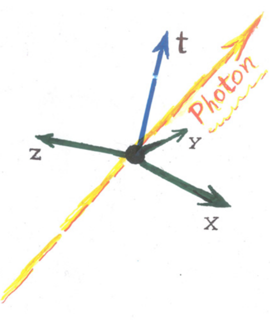

Figure 3.

Now, the most important thing about this description is to represent the speed of light, so in Fig. 3 we have a light ray (or photon history) where, in order that it does not simply almost lie along the spatial "floor", we need to choose space and time units so that the light ray can be represented as tilted at some reasonable angle to this "floor", such as at ~45°. In Fig. 4 we can depict the *null cone*, which represents all the space-time directions of light rays through our chosen origin point. These cones are very important for the structure of space-time, and we shall frequently be concerned with the null cones themselves rather than with any particular light ray (Fig. 5). Moreover, we need not be concerned with the choice of axes either (Fig. 6). What is physically important is this null cone itself, at each point of the space-time.

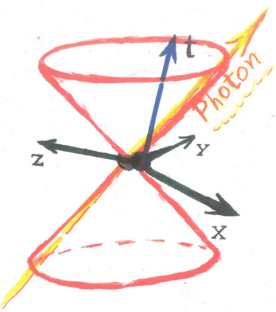

Figure 4.

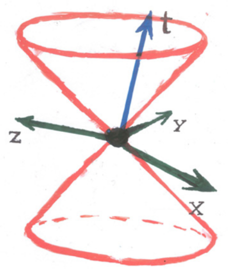

Figure 5.

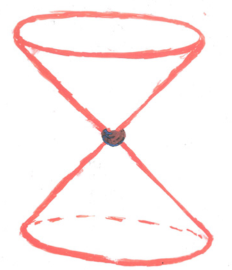
Figure 6.

In general relativity we likewise have a null cone at each point of a (now curved) space-time, representing the local speed of light at each space-time point, but the cones can be more or less all over the place (Fig. 7). We can imagine a space-time point p with all the light rays coming out of *p*. In Einstein's general theory of relativity, these light rays are geometrically what are called null geodesics. In what follows, we shall refer to these *null geodesics* simply as rays. When these rays are extended outwards into the future, away from p in the space-time, we get what is called the future *light cone* of *p* (Fig. 8). The null cones would be *tangent* to this light cone wherever it goes. But, as you can see at the back, at the top right-hand side of Fig. 8, the light rays may start to cross each other, and this sort of thing can make light cones complicated. However, it is important for what I am going to discuss shortly that you appreciate how to deal with such things. In general situations, you certainly are likely to get such crossover regions, caustics and things like that, and how to deal with them is a central feature of the following discussion.

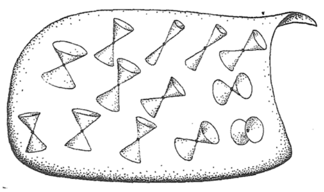

Figure 7.

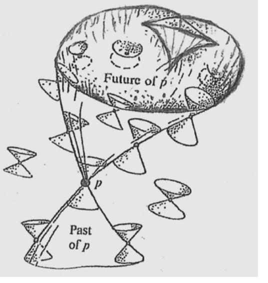

Figure 8.

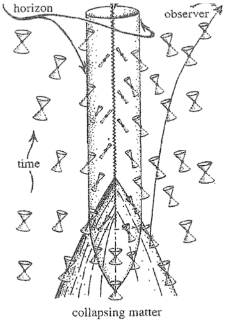

Figure 9.

Now, let us consider the next picture (Fig. 9), where we see a space-time depiction of what is, in effect, the Oppenheimer-Snyder (O-S) evolution of the collapse of a dust cloud to what we now refer to as a *black hole* – although that particular aspect of Fig. 9 was not something that was properly appreciated in 1939, when J. Robert Oppenheimer and his student Hartland Snyder published their paper [1]. What they were concerned with was the behaviour of the material body, indicated in the central lower part of Fig. 9, depicted as falling inwards (as we move up the picture to express the passage of time). This gravitationally collapsing material was to be what is referred to as "dust", which simply means a fluid material with no pressure. These authors had been considering how a very massive individual star might behave, according to Einstein's general relativity, in the star's late stages, when nuclear energy and other relevant resources would have become exhausted so that the material's pressure would become too small to be able to prevent the star's collapse. Whereas such an idealization of the gravitational collapse of an individual star need not be regarded as particularly realistic, such a picture might also be applied to much larger collections of material where, in its outer regions at least, the "dust" approximation is not at all unreasonable.

It should be mentioned that this collapse phenomenon had been encountered much earlier, initially by Subrahmanyan Chandrasekhar [2], in his 1931 discussion of white dwarfs, where he had calculated that such a star, when it had cooled off, would not be able to sustain itself (by its electron degeneracy pressure) if its mass were greater than around 1½ times the mass of the sun, and his considerations are now, not unreasonably, considered to constitute the "birth" of the idea of the plausible existence of what we currently refer to as "black holes". Nevertheless, in my considerations here, I shall refer only to the O-S picture, simply because with the very much larger collections of material that will be our main concern, gravity will be the dominant physical interaction, and such things as nuclear forces and degeneracy pressure become unimportant, so that the simpler O-S picture suffices for our purposes.

The collapse is taken to be exactly spherically symmetrical, in the O-S situation, so that all the matter falls directly inwards towards the central point, where the material's density becomes *infinite*, in this idealization, so the space-time *curvature* must also become infinite, in accordance with Einstein's general relativity. According to the picture of Fig. 9, depicted above where the infalling material is represented, we see this singularity to be extending further upwards, i.e. maintaining its existence with respect to an external time measure, so even in the vacuum region of the picture, we appear to have to maintain the singularity, long after the material has seemed to be absorbed into the singularity. It does not appear that Oppenheimer and Snyder were particularly concerned with such an aspect of their collapse picture – i.e., with the apparent future of the vacuum region following the collapse – but it is, indeed, a matter that we shall shortly need to address in a serious way.

Although the O-S picture of such a gravitational collapse might well not be considered to be altogether realistic under general circumstances, the space-time picture of Fig. 9 can nevertheless be regarded as providing us with a reasonable first impression of the kind of situation that might well arise, at least in the early stages of a large-scale gravitational collapse, before the effects of departure from spherical symmetry and the presence of pressure may indeed begin to have significant implications. It is thus important to know what aspects of this picture might or might not be expected to be maintained under more general circumstances.

The O-S idealized description of a gravitational collapse was fairly well appreciated at the much later time when quasars were discovered in the early 1960s, these bring extra-ordinarily powerful very distant sources of radio signals, and theorists then started to speculate as to whether there might be something like an O-S collapse involved, yet without the understandably gross simplifications of spherical symmetry and pressure-free material, assumed in the O-S situation.

My own acquaintance with the kind of space-time geometry that is required for this O-S collapse picture had occurred somewhat earlier than this, namely in 1959, when I was a mathematics Research Fellow at St John's College Cambridge. In January 1959, I had not yet become aware of the O-S paper, but I went to a lecture by David Finkelstein, given at King's College London, accompanied by my good friend and mentor Dennis Sciama, from whom I had learnt a good deal of relevant physics. He drove me there from Cambridge, having encouraged me that the talk would be interesting to me. Finkelstein's seminar described how you can smoothly pass through what had then been referred to as the "*Schwarzschild singularity*", a feature occurring at a certain radius out from the centre, according to the well-known spherically symmetrical *Schwarzschild solution* of Einstein's vacuum equations, for the external gravitational field of a spherically symmetrical body. This radius is:

$$r = 2m$$

in units where the speed of light c and the gravitational constant G are both chosen to have the value unity:

$$c = 1, G = 1.$$

(Without this choice of units, we have $r = 2Gm/c^2$.) In the coordinate description that Schwarzschild had used, we find that a metric component indeed becomes infinite, and this cannot be avoided in the seemingly very reasonable time-symmetrical description adopted by Schwarzschild.

Karl Schwarzschild had found this very basic solution of Einstein's vacuum equations in 1916, not long after Einstein introduced his general theory of relativity in 1915. Schwarzschild had solved the equations of Einstein's general relativity for the vacuum gravitational field outside a spherically symmetrical body. He also provided a solution for the material of the body itself, but that was not an altogether realistic material, and is not important for our discussion here.

In fact, *the top* part of Fig. 9, depicting the situation arising after all the actual matter has disappeared, apparently all having been absorbed into the central space-time singularity, depicts essentially what Finkelstein described, where you see a portion of the Schwarzschild vacuum space-time, but in a coordinate description that allows a smooth extension to within the $r = 2m$ Schwarzschild radius. Such a situation was described by Finkelstein in a coordinate description that allows for a non-singular extension of the space-time to within the Schwarzschild radius of $r = 2m$, thereby allowing a description for which the infalling material can be seen as actually falling smoothly through the $r = 2m$ Schwarzschild radius that had appeared to be a singularity in Schwarzschild's original description.

Indeed, the description shown in Fig. 9 was not the one presented by Schwarzschild's choice of coordinates, his reasonable-seeming choice providing a *static, time-symmetric* picture, which would not allow the inward-tilting null cones that we see at this radius in the upper part of Fig. 9. Instead, as remarked previously, the expression that Schwarzschild obtained becomes *infinite* at the Schwarzschild radius r = 2m, so that the term "Scharzschild singularity" is now seen to be inappropriate. Finkelstein, in his talk, provided a coordinate change to obtain an elegant time-asymmetric form [3] that extended the solution inwards to values of the radius r that lie in the full range

$$0 < r < \infty,$$

the metric form remaining perfectly smooth across $r = 2m$. This "Schwarzschild radius" now describes the upper cylinder of Fig. 9, where the inward tilting null cones become tangential to this $r = 2m$ cylinder. Clearly, the tilting of the cones presents a time-asymmetric description at that radius, which explains why Schwarzschild's time-symmetric assumption leads to what appears to be a singularity, rather than the *horizon* that was made evident by Finkelstein's choice of coordinates, and which describes the situation of the upper part of Fig. 9.

It should be pointed out that in the early days of general relativity, various other theorists – going back to Painlevé in 1921 [4] – had also found coordinate changes that could eliminate this Schwarzschild singularity, but most did not fully appreciate the physical implications of the "horizon" character of this Schwarzschild radius. Most noteworthy among those who did properly appreciate this physical situation was Abbé Georges Lemaître [5], who understood that infalling material could cross $r = 2m$ into an interior region, consistently with the situation depicted in Fig.9.

I came away from Finkelstein's lecture thinking that, whereas by a suitable coordinate choice you can – somewhat remarkably – get rid of this "Schwarzschild singularity", located at a certain distance $2m$ out from the centre, you nevertheless still have a genuine singularity in the middle (at $r = 0$) where space-time curvatures become infinite, so no coordinate change can help. Accordingly, I began to wonder whether there might perhaps be a general theorem, or something like that, which showed that whatever you might do to complicate the space-time metric, to describe a similar but very irregular collapse situation, you might still have to get a genuine singularity. I had not heard of any such theorem, nor did I have any idea how one might prove such a thing, so I started to think to myself: "what might I know, plausibly relevant to general relativity, that maybe other people in the field don't generally know much about – so possibly this could be helpful to me to achieve something along these lines, that theorists working in the field seem not be familiar with?"

The area that I thought might be helpful was the theory of 2-*component spinors*. I should explain that I finally understood properly about 2-spinors from lectures by the great quantum physicist Paul Dirac, earlier in the spring of 1958, in the academic year before Finkelstein's lecture. Dirac's course was basically on quantum field theory, but he appeared to have deviated from his normal course when he talked about 2-spinors. He was famous for discovering the dynamical equation for the electron, but this originally involved the introduction of 4-spinors. However, you can break each of them down into a pair of 2-spinors and Dirac had become well acquainted with this fact, having written an important paper [6] describing higher-spin fields in these terms. Yet, these techniques were not very familiar to most physicists at that time. I had myself heard of 2-spinors and had been intrigued by them, but I did not really understand them until Dirac's lectures had made them abundantly clear to me!

So, soon after Finkelstein's lecture, I began to wonder whether 2-spinor theory might be something I could apply to general relativity, which possibly might supply me with certain significant insights into general relativity that were, perhaps, unfamiliar to most theorists working in the field. But for this I needed to have a clear geometrical way of thinking about a 2-spinor. How was this to be achieved?

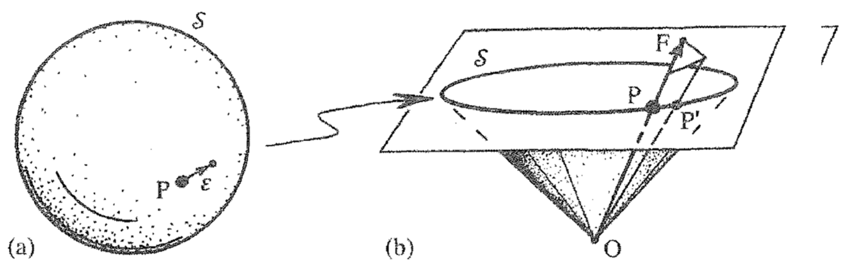

Figure 10.

In Fig. 10 we have a geometrical picture representing this remarkable concept. On the right-hand side of the picture (Fig. 10b) I have, in effect, depicted a 2-spinor in space-time terms, but this needs some clarification. On the left-hand side of Fig. 10 (10a), we see the celestial sphere *S* – although it is better to think of this as the *future* celestial sphere, rather than the past one (i.e. where light rays go to, rather than where they come from), and Fig. 10b depicts the different directions in which the 2-spinor's "flagpole" can point, this flagpole being a future-pointing null vector. We must bear in mind that there needs to be an additional spatial dimension in the picture of Fig. 10b, so that the bounding "ring" at the top is actually a *sphere S*, as in Fig. 10a. The 2-spinor also has a little flag attached to this null vector, whose plane touches this sphere S and determines the 2-spinor's *phase* (Payne [7], Penrose and Rindler [8]). We have to bear in mind that the upper ring in Fig. 10b is actually a sphere, so there is a freedom for the flag plane to rotate while still touching the sphere, because of the extra dimension, and it is better to refer back to Fig. 10a, where we note that the direction of the flag plane corresponds to a tangent vector ε on the spere *S* of Fig. 10a, corresponding to the direction towards *ṕ*, out from *p* in Fig 10b. The "spinorial" nature of this flag corresponds to the fact that a rotation of the flag through 360° changes the 2-spinor's sign, while a 720° rotation returns it to its original value.

The important thing about using the 2-spinor formalism in general relativity, for our purposes, is that it provides us with some insights about space-time curvature that are not so evident in the standard tensor formalism. Most particularly, there is a particular part of the space-time curvature, referred to as the *conformal curvature*, that is determined by a tensor quantity referred to as the *Weyl curvature tensor,* named after the highly esteemed mathematician Hermann Weyl. The Weyl tensor is defined from the full Riemann curvature tensor by a somewhat complicated-looking formula, but the 2-spinor formalism brings out an essential simplicity of the Weyl curvature that is not at all evident in the conventional tensor formalism.

It is not my purpose, here, to provide explicit expressions for all the tensor or spinor expressions for the quantities that are involved, these making use of the delicate interplay between tensor and spinor indices, often involving the metric tensor in its very elementary spinor form. The details of these algebraic manipulations are, of course, important, and are ultimately based on fundamentally simple rules [9], [8], but it is not necessary for us to go into this in detail here.

The Riemann curvature tensor, when written in its 2-spinor form, splits into three parts, one of which is the spinor form of the trace-free Ricci tensor, another being the trace of the Ricci tensor (namely the scalar curvature), and the third part being the spinor form of the Weyl tensor. Einstein's field equations, when expressed in 2-spinor form, effectively tell us that the spinor form of the trace-free Ricci part plays a role as a *source* of the gravitational field, the latter being described by the spinor form of the Weyl tensor, this "Weyl spinor" describing the free gravitational field. There is a strong analogy with electromagnetism revealed here, where the Ricci tensor (really in trace-reversed form) is analogous to the charge-current vector of electromagnetism, the Weyl spinor being analogous to a "Maxwell spinor", which is the spinor form of the Maxwell field tensor, which describes the free electromagnetic field. This analogy is made much more evident in the 2-spinor form than in the tensor form.

Indeed, in the 2-spinor formalism, we find a particularly simple expression for the Weyl curvature, which is not at all evident in the tensor formulation. We find that this "Weyl curvature spinor" is a 4-index 2-spinor which is *totally symmetric* in all its four 2-spinor indices, and it satisfies a very simple free-field equation in vacuum, where there are no gravitational sources (vanishing Ricci tensor). This is strikingly analogous to

Maxwell's electromagnetic theory in 2-spinor form, where the spinor form of the Maxwell field tensor is a 2-index 2-spinor which is *symmetric* in its two 2-spinor indices, and it satisfies the completely analogous free-field equations when the charge-current vector vanishes [9], [8]. All this fits in closely with a general study, made by Dirac in 1936, of field equations for fields of arbitrary spin, written in 2-spinor form. (Curiously, I once had a private discussion with Dirac, when I explained to him how Einstein's gravitational theory fits in with the other spin fields in Dirac's own 2-spinor analysis of higher-spin fields [6]. He had not been previously familiar with this aspect of Einstein's theory, and he found it interesting.)

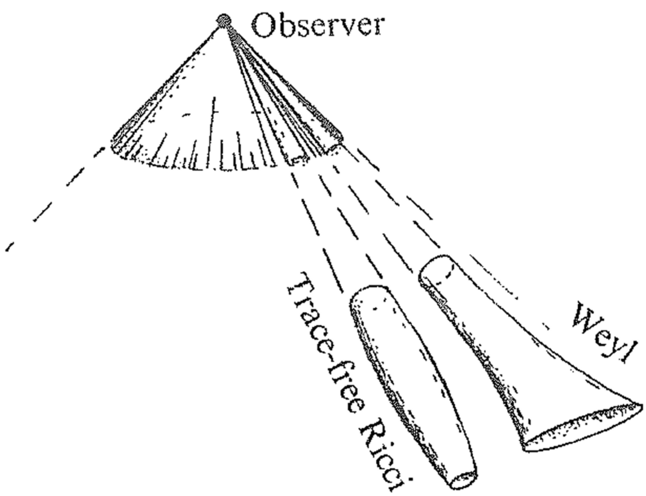

Figure 11.

In order to appreciate the geometrical effect of the Weyl and Ricci curvatures, we can have a look at Fig. 11. This depicts an observer, located at the top of the picture, looking back into the past and sees the inward focusing effect due to the trace-free Ricci curvature (behaving like an ordinary positive convex lens), and also sees the *distortion* due to the Weyl curvature W (like a purely astigmatic lens). In this way, we see the direct roles of the (trace-free) Ricci curvature and the Weyl curvature in their effects on light rays. The trace-free Ricci tensor acts as a lens which is positively focusing (like an ordinary magnifying lens) when the energy flux across the light ray is positive. (The trace part of the Ricci tensor, being proportional to the metric tensor, does not affect light rays.)

A few years after my time as a Cambridge Research Fellow, the quasars ("quasi-stellar objects") were being discovered, this being around 1962–1964. These objects were producing enormous quantities of energy and yet they seemed to be remarkably small, considering this output. They appeared to have an energy output of perhaps 100 to 1000 times an entire galaxy's emission, but yet they seemed to vary substantially in a few hours or days, which meant that they had to be very small compared with a galaxy, probably not larger than the solar system, but would have to be extremely massive to emit so much energy. How could all that mass-energy be squashed into that small volume? Astrophysicists, such as Fred Hoyle, started to speculate upon whether something gravitationally concentrated as in the O-S collapse might be relevant. But when you think just of a collapse that falls radially inwards, it doesn't give you any scope for signals coming out. If you were to have any involvement of gravitational waves, these would need to have at least a quadrupole structure. Moreover, the highly varying nature of the radio signals suggests that something very complicated must be involved. Certainly, the possibility of a very irregular gravitational collapse must be considered, which could be very different from the O-S picture of matter falling radially inwards. Perhaps material falling inwards in a complicated way might swirl round an come out again. Such possibilities had been suggested to me by John Wheeler and others, so I began to think again about such matters as a general gravitational collapse, in a serious way.

At around that time, in 1963, there had been a paper published by two Russians, Lifshitz and Khalatnikov [10], which seemed to have established that in general you would *not* get singularities, these occurring only in very special circumstances such as in the particular O-S collapse picture. Accordingly, in a physically realistic situation you might indeed consider that in a generic collapse the material would indeed just swish around and come swirling out again. As part of my worries about this problem, I had a good look at the Russians' paper. In fact, there was actually a serious mistake in the paper, but I didn't examine it carefully enough to notice that. What I did feel, however, was that the methods they were using were not altogether convincing to me, and that it was worth trying to think independently about whether or not you would get singularities in a generic gravitational collapse.

I remember walking in the woods near where I lived at the time, in Stanmore at the north of London, and trying to imagine that I was in the midst of such a gravitationally collapsing situation, I came to the conclusion that it could not just be a local curvature blow-up, but it had to involve a build-up of curvature due to some overall excessiveness of the material concentration. Some kind of non-local criterion would be needed to tell you that a "point of no return" had, in an appropriate sense, been passed.

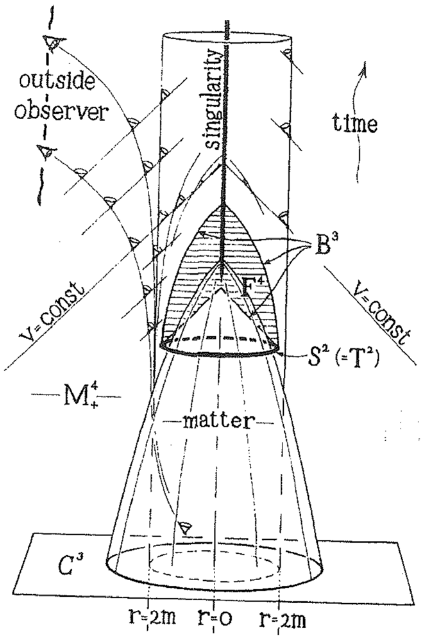

Figure 12.

A few weeks later, I hit upon the idea of a "*trapped surface*" which seemed to supply the kind of non-local criterion that I was looking for. The picture you see in Fig. 12 is a diagram that subsequently appeared in my 1965 paper [11], though a little earlier, in late 1964, I gave a talk at King's College London about it. The argument was that if you have a collapse which is generic, but in which you happen to have a trapped surface, then you will still have problems with singularities even though the infalling material would not be all aimed at a central point. In Fig. 12, you see what is basically the O-S collapse of Fig. 9, but you can also see that there is this little ring in the middle of the picture, marked S^2 (=T^2) surrounding the infalling matter. This is a trapped surface, and you have to realize that it is not actually a "ring" because I am only depicting two spatial dimensions, together with the time dimension and (as with Fig. 10b) the whole thing should actually be a 4-dimensional space-time, so that the "ring" is really a 2-dimensional surface, topologically like an ordinary sphere. In a general collapse, it might be distorted, not necessarily being a precise geometrical sphere.

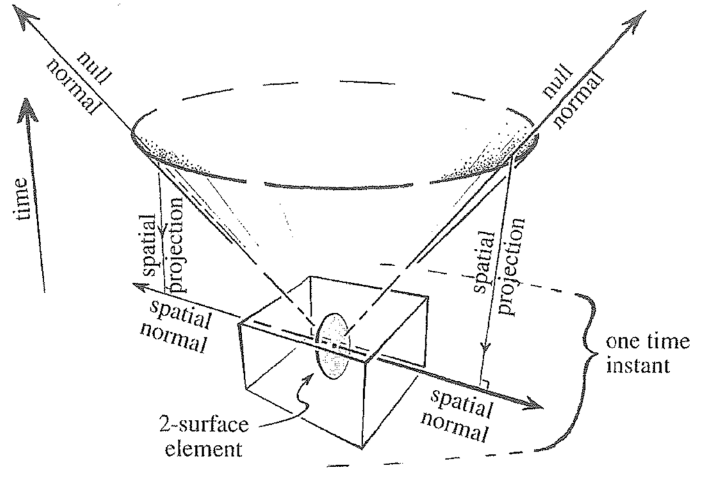

Figure 13.

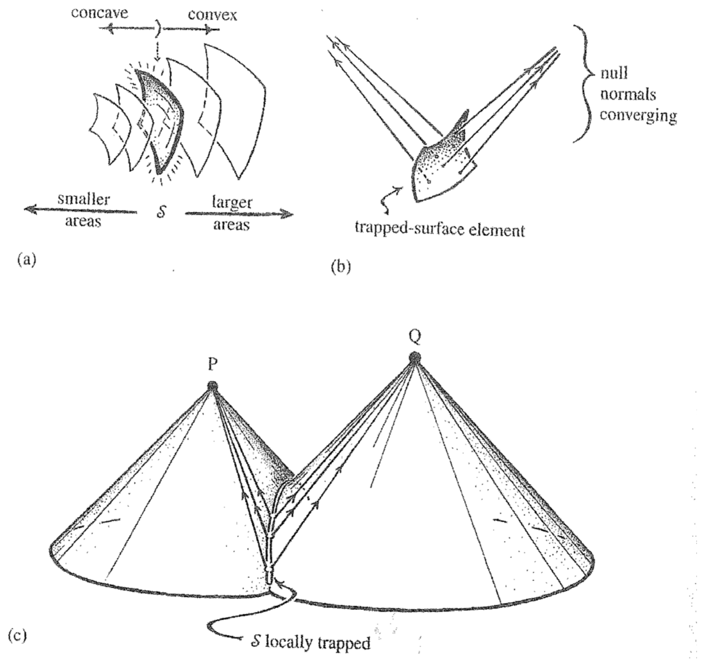

Figure 14.

But what exactly is a trapped surface? You have to imagine that there is a flash of light emitted simultaneously all over that surface and that this flash has a certain characteristic property. In the lower middle part of Fig. 13, you see a little 2-dimensional spatial surface element and you are to imagine that the surface emits a flash of light, which moves directly away from it in opposite directions, represented by light rays in the 4-space-time-dimensionally indicated top part of the picture. In Fig. 14, at the top left (Fig. 14a), we have a *curved* surface element which is concave on one side, where the light-flash rays from it are converging, and convex on the other side, with light-flash rays which diverge. That would be a normal thing for a curved surface element. Now, on the top right (Fig. 14b) you have something that's harder to imagine, namely that light-flash rays can converge on *both* sides. This is what we need for a trapped surface. In fact, in space-time terms, this is not really a problem for a curved surface element. At the bottom (Fig. 14c) you see a situation in ordinary flat (Minkowski) space-time, where the surface is the intersection of two past light cones (with vertices P and Q) and, indeed, the light flashes coming from *both* sides of such a surface element would indeed be converging. The definition of a *trapped surface* S^2, however, is that it is a *compact* spacelike 2-surface which has this *both-way* converging property occurs all over the surface S^2. The essential compactness condition means that it is closed up without any boundary (like an ordinary sphere, but which could be distorted in various ways without disturbing its topology). Thus, each point of this S^2 is *locally* like the intersection surface of Fig. 14c (i.e. like Fig. 14b) which, in itself, would present no problem, but what makes it the problematic "trapped" surface is that it is also compact, a feature that would certainly be the case for rotationally symmetrical 2-surfaces within the Schwarzschild horizon of Fig. 9.

A key aspect of this definition is that it is stable under small deformations. Because of the compactness of the surface S^2, if we were to vary the situation very slightly, the degree of the convergence of light rays moving out orthogonally away from S^2 must still be bounded away from zero, for a small-enough change in the geometry of the situation. Accordingly, a small-enough generic perturbation away from the spherically symmetrical case will not disturb the "trappedness" of the 2-surface S^2. We can certainly have many generic collapse situations in which a trapped surface indeed arises.

The next question for us is: why does the presence of a trapped surface cause us a problem? We need to consider that the trapped surface S^2 lies in a space-time region M^+ that is the future time-evolution away from a *non-compact* spacelike 3-surface C^3. (We are thinking of a reasonably local phenomenon, not a cosmological one for which we might, on the other hand, actually want C^3 to be some compact spacelike 3-surface.) We shall be concerned with the region F^4 (shown shaded in Fig. 12), within the future evolution of $C^{3,}$ which is the *chronological future* of S^2 [12], [13] That is to say, F^4 is the region swept out by (i.e. the union of) all the time-like curves with past end-points in S^2. What we are particularly concerned with is the *boundary* B^3 of F^4.

It follows from general results (see [12], [13]) that any point p of B^3, not on S^2, is the future end-point of a ray (i.e. null geodesic) lying on B^3 which is either past endless or else has a past end-point on the initial region S^2. The past-endless case is anomalous, because it implies that the region M^+ is not a standard time-evolution away from C^3, since it contains rays that wind endlessly into the past without ever reaching the initial surface C^3. It would not be unreasonable to regard such a departure from normal causality as "singular" behaviour, and no less anomalous than regions where the space-time curvature diverges to infinity. Accordingly, I shall here refer to such anomalous space-times as being "singular" whether or not their "incomplete" status actually arises from a curvature divergence.

We have just seen that the boundary B^3 of F^4 consists entirely of rays, and if M^+ is to be non-singular in the above sense, then all these rays have past end-points on S^2, but what do these rays do in their futures? For this, we need the "trapped" nature of S^2, which tells us that the divergence of these rays starts off as *negative* as they leave S^2. Now there is a result known as the "Raichaudhury effect" for rays [12], [13], that tells us that if we have a null hypersurface (such as B^3) for which the divergence of its generating rays is initially negative (which is here the trapped-surface condition) and for which the curvature along the rays is non-negative (i.e., according to Einstein's equations, that the energy flux across these ray s is non-negative, as indicated in Fig. 11), then after a finite affine distance along the ray its separation from some of its neighbouring rays becomes zero, which is what happens at a caustic point, as illustrated at the top right of Fig. 8. This tells us that – either at that caustic point, or, more usually earlier than that, at a crossing region of B^3 – the ray ceases to lie on the boundary B^3 of F^4 and enters into the interior of F^4. Either way, it is only a finite segment of the ray which lies on B^3, telling us that B^3 must, in fact, be *compact*, consisting entirely of a union of finite segments of rays, each of which has an initial end-point on the compact surface S^2.

It is this compactness of B^3 that actually leads us to a contradiction that drives us to the conclusion that there must be a *singularity* somewhere within F^4 (or perhaps at its boundary). However, in my original paper [11] I used a rather clumsy argument to demonstrate this, and afterwards Charles W. Misner pointed out to me that I could have used a much simpler argument, using the fact that there is a general theorem concerning Lorentzian manifolds (i.e. manifolds with the standard metric signature, with one time-like dimension and the rest space-like) that there exists a smooth time-like vector field all over the whole manifold. We can follow along such a vector field to map B^3 homeomorphically (i.e. preserving its topology) into C^3. Since B^3 is compact without boundary, the image of this map must also be compact without boundary. Since this 3-dimensional image indeed has no boundary, it must be the whole of C^3. But C^3 is non-compact, which immediately provides us with the required contradiction. I had been aware of the result that Misner was using in his subsequent contribution to the argument, but for some reason I had not thought of using it. In all my later accounts of this result I took advantage of Misner's simplification.

In the autumn of 1965 there was a conference at Imperial College, London about progress in general relativity, and Igor Novikov from the Russian school tried to present the aforementioned 1963 result, but Misner then pointed out the conflict with my result. Subsequently Belinskii joined with Lifschitz and Khalatnikov to provide a corrected paper with a conclusion opposite from what they had done before [14], and Misner provided his own version [15], these papers demonstrating how immensely complicated the singularities in a general gravitational collapse can be, with the Weyl tensor itself diverging in an extremely complicated way. I shall refer to these as *BKLM-type* singularities.

At this point, I should draw attention to the fact that what I showed was that the occurrence of *singularities* is a robust prediction of Einstein's general theory of relativity, not necessarily that *black holes* must necessarily be the consequence of a realistic gravitational collapse. Another possibility might be that "naked singularities" could arise, these being, in effect, space-time singularities that are not hidden behind event horizons, and so might actually be directly visible from a distance away. Unlike the case of a black hole, however, there is no reason to expect that actual naked singularities should exist in nature (although there are many exact solutions of the Einstein equations which do possess naked singularities, such as the Schwarzschild solution for a negative mass). The normal view – probably correct, in my opinion – is that naked singularities cannot occur in ordinary gravitational collapse situations but, as far as I am aware, no mathematical theorem has yet been provided to demonstrate what I have referred to as the "Cosmic censorship hypothesis", which might, in effect, prove that naked singularities are unstable, or something like that (though I did once write a paper exploring some mathematical implications of the production of naked singularities [16]).

After I gave my talk at King's College London, Dennis Sciama asked me to give a repeat in Cambridge, which I did in early 1965 and at that occasion Stephen Hawking was present (though he had not been present at my London talk). Immediately following my Cambridge lecture, I had a private session with Stephen, accompanied also by George Ellis, who had been collaborating with him on the use of certain ideas that they had hoped might be useful in addressing the necessity of the Big Bang singularity. I explained much more about the details of my arguments. Stephen Hawking very quickly picked up on the ideas and applied a version of my own result to cosmology in an original way. Subsequently, Hawking developed these techniques very considerably, often with the benefit of critical corrections from Brandon Carter. Hawking's developments were published in a series of three papers in the *Proceedings of the Royal Society* [17], [18], [19], [20]. Eventually I came back to collaborate with Hawking to provide a very general result which encompassed practically all that had gone before [21] (although with a slightly stronger energy condition than the one used in my original argument [11]).

Hawking's considerations were directly concerned with the problem of the Big Bang. The issue under consideration was whether the "singular" description of that event is necessary, and might it have been the case that there had been a previously collapsing phase of the universe which, through some extreme complication of the earlier collapse, the universe might have "bounced" into the expansion that we now perceive in our "Big Bang".

In Fig. 15 I have indicated our current picture of the overall history of the universe, from its Big Bang origin to the currently observed exponential expansion. The frilly part at the back is just to accommodate the currently popular view that the universe may well not be spatially closed, but may continue indefinitely in spatial directions. The exponential expansion is a conclusion arrived at by the 2011 Physics Nobel Prize laureates Saul Perlmutter, Brian Schmidt and Adam Riess, and it can be most directly explained by the presence of a positive value for the cosmological constant Λ in Einstein's modified gravitational field equations. Einstein introduced his Λ-term for the wrong reason in 1917. At that time, Einstein was hoping to be able to accommodate a spatially closed static universe into his equations. However, he later rejected that term in his equations after being convinced by Hubble and others that the universe is actually expanding. Nevertheless, a positive value for Λ is the most economical explanation for what is now rather misleadingly referred to as "dark energy", namely the cause of the exponential expansion that I have schematically depicted in Fig. 15, this having been convincingly observed at around the turn of the 21st century by Perlmutter, Schmidt and Riess.

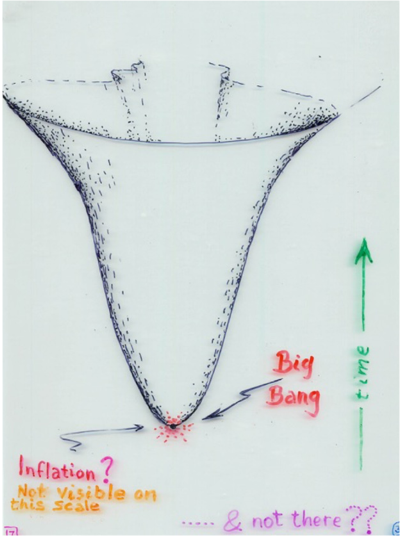

Figure 15.

The “singularity theorems” that Stephen Hawking and I and some others had developed, all have the character of being completely insensitive to the direction of time. I remember being puzzled by the fact that whereas one might well expect that some very complicated solutions of the Einstein equations might apply in the future, perhaps being realized through the interplay by systems of black holes, while, on the other hand, cosmologists seemed to restrict their attention to the very simplest of possibilities. I recall being very puzzled by why cosmologists did not study any of the many other kinds of possible singular origins for the universe. I remember an occasion when I was in Princeton and was about to go to one of the frequent conferences at Stevens Institute in Hoboken, New Jersey. We used to drive up in several cars, and I noticed in the back of one of the cars was James Peebles (later to become the 2019 Nobel Prize laureate in physics), so I asked him why serious cosmologists never seemed to consider any of these complicated possible alternative singularities that you might have for the description of a Big Bang, rather than just this simple highly symmetrical special case. Why, I asked, do cosmologists never consider any of these more complicated alternatives? He just looked at me and said, “because the universe is not like that”. So, I thought to myself: “my gosh, it isn’t like that is it – but why?”

I presumed that he was partly thinking about the uniformity of the *cosmic microwave background* radiation (CMB) which is indeed very uniform over the whole sky, having been discovered in 1963 by Arno Penzias and Robert Wilson, the 1978 Physics Nobel Prize laureates (shared also with

Pyotr Kapitsa), and this remarkable uniformity tells you that the universe really is indeed very spatially uniform. It struck me that there is something very strange about all these various singularities, namely that the big bang singularity is utterly different from the kinds that you might see in the future, namely in the collapses in black holes, with the distinct likelihood of things like the BKLM type of singularity arising! I was very puzzled by this, particularly since everybody seemed to think that the solution to the singularity problem would lie in combining general relativity with quantum mechanics. Accordingly, you need to find a *quantum gravity* theory to resolve the singularities – so everyone had supposed. But it seemed to me that it must be a very peculiar quantum gravity theory which is grossly asymmetrical in time, in order to give you a theory which makes the singularities quite different in structure in the past from in the future!

I held this view for quite a long time, but then I began to worry more about this problem in relation to the entropy in the universe. This is a key issue, which I shall come to shortly, but let us first consider what is currently a very popular view about what the very early universe was actually like. In Fig. 15, I have indeed sketched the history of our universe according to observation (and some theory), with time going up the picture. At the bottom is the big bang and at the top we see the beginnings of the exponential expansion which became an observationally established feature of current cosmology, through the work of Reiss, Schmidt and Perlmutter and which can be taken as an implication for a *positive* cosmological constant Λ in Einstein's 1917 version of his theory.

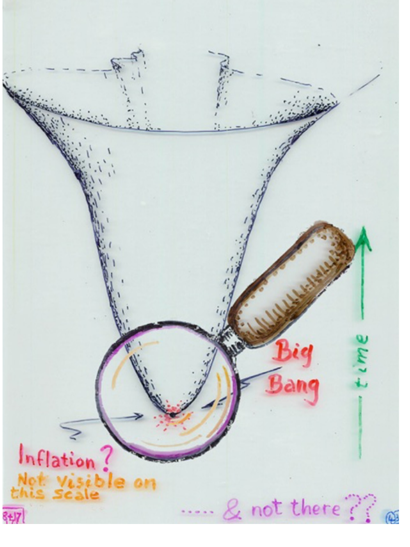

Figure 16.

Figure 17.

However, if you want to get a picture of what seems to be currently believed by most cosmologists concerning the extremely early universe, you need a very powerful magnifying glass (Fig. 16) in order have a good look at it in the picture. What you would see (Fig. 17), according to current "inflationary cosmology" would be a much earlier exponential expansion, supposedly all taking place within the absurdly tiny initial 10^{-32} seconds, or so, of the universe's existence, this being referred to as the *inflationary* phase of the universe's expansion. Inflation was initially introduced by Starobinski and Andrei Linde [22], and by Alan Guth [23], [24] in the early 1980s for various reasons, but a particularly important one in the present context was in order to explain an observed striking feature of the tiny variations in the temperature of the CMB radiation over the sky, namely that these variations are extremely closely *scale invariant,* which suggests some sort of exponentially expanding origin for these disturbances.

Another claim often made for the existence of this early inflationary phase was that it would iron out the very early universe immediately following the big bang so as to provide us the very uniform universe that James Peebles had pointed out to me.

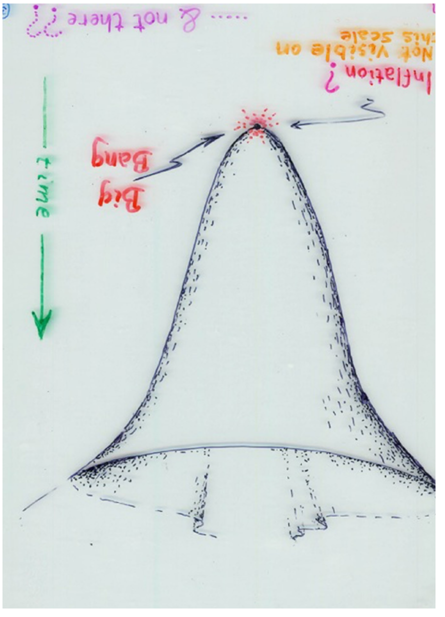

Figure 18.

However, I could never believe that it would actually do that, for the following reason. Imagine that the universe was contracting, rather than expanding, as shown in Fig. 18, this being simply Fig. 15 turned upside down, so the progression of time is still represented as going up the picture. We should take note of the fact that all the relevant dynamical equations – including those of the "inflation field", introduced solely for the purpose of making inflation work – are all unchanged if we reverse the direction of time. Accordingly, Fig. 18 represents a possible universe evolution. However, if we introduce some perturbations into the mass distribution, we must expect black holes to arise in this collapsing situation, which will start to merge with one another and finally produce a horrendous non-uniform mess of a "generic" singularity at the end, as indicated in Fig. 19, the presence of an "inflation field" making no essential difference at all to the picture. Now, we reverse the time-direction back again, as shown in Fig. 20, and we ask: why did the universe not have this "far more probable" type of big bang singularity, as opposed to the uniform one illustrated in indicated in Fig. 15? Inflation provides no answer to this fundamental problem.

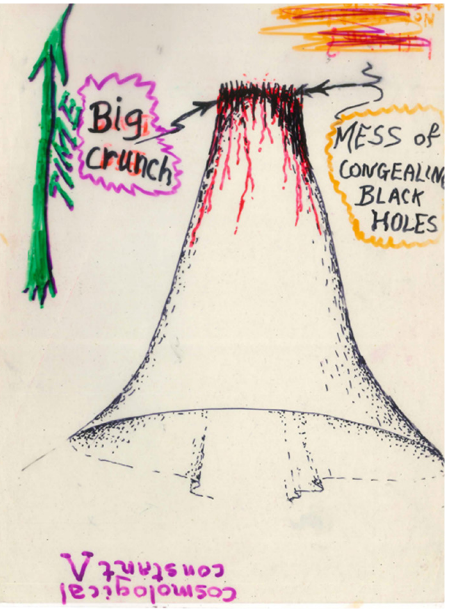

Figure 19.

In order to quantify this issue a little better, we must turn to thermodynamics, and most particularly to the Bekenstein-Hawking formula for the entropy of a black hole. We find, for the totality of the black holes currently within our observable universe, that the entropy is utterly dominated by that in the black holes. Moreover, in a collapse situation like that in Fig. 19, where we take into consideration only the amount of material (including dark matter) within our current observable universe, we find a value for the entropy in a collapse like that of Fig. 19 would be something like 10^{124}. Then, taking into consideration that entropies are really logarithms of probabilities, we come to the conclusion that if our big bang had come about as a singularity chosen somehow "by chance", then the odds against the uniform situation in the big bang that we actually appear to see (i.e. like Fig. 20, rather than the observed Fig. 15, now with time proceeding upwards in the picture), would be something like the utterly absurd figure of around $\exp(10^{124})$:1 (i.e. the probability of our observed universe having the uniformity that we see, this having been a chance occurrence – inflation or no inflation – would be the reciprokal of a number with around 10^{124} digits) which can hardly be the right answer! We need another explanation for the extraordinary specialness of our big bang.

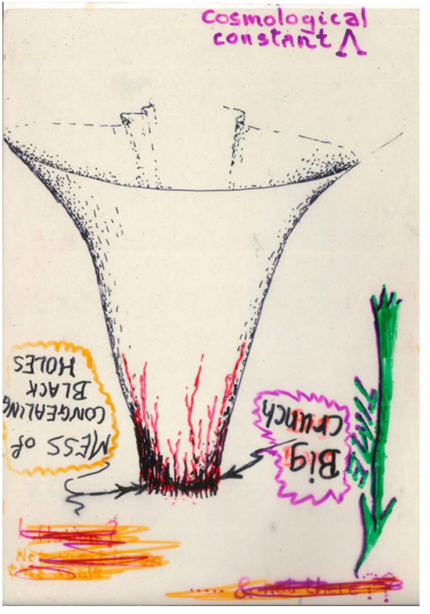

Figure 20.

In fact, the issue is even more curious than this. Let us consider Fig. 21. The top three pictures represent a gas in a box where in the top left-hand picture we see the gas initially constrained to be in a smaller box in its lower right-hand corner. Then you open this smaller box and the gas spreads out within the large box, so that as we move from the left-hand to the right-hand of the top three pictures, the gas gets more and more uniform. This illustrates the action of the 2nd law of thermodynamics, where the entropy (or the randomness) increases with time – time being represented as going from left to right in Fig. 21. Now let us consider the bottom three pictures in Fig. 21. These represent an imagined galactic-scale box containing a large number of stars, initially (bottom left picture) taken to be pretty uniformly distributed. Then, because of the universally attractive nature of gravity, the star distribution gets more and more clumpy, as we move from the lower left picture to the lower right one with, perhaps, the formation of a black hole, in the lower right-hand picture. Again, the time increases from left to right, and so also does the entropy.

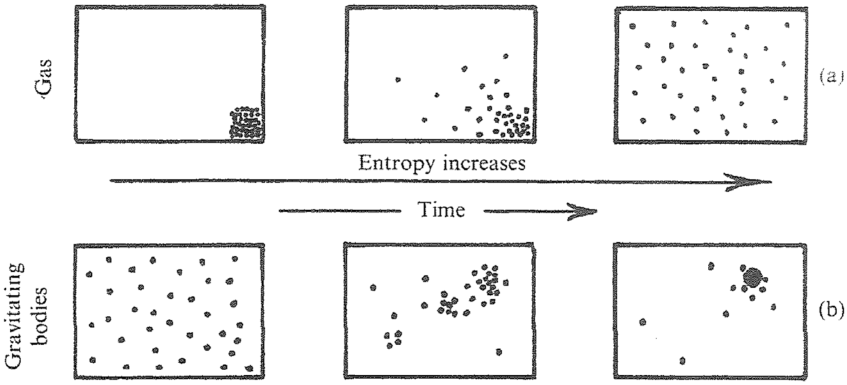

Figure 21.

Now, what is it that we see in the very early stages of our actual *universe*? It is uniformity which, as described in Fig. 21 by a combination of the upper right and lower left pictures. Moreover, one of the most striking features of the CMB observations is the extraordinarily precise Planck curve in temperature distribution for different frequencies, telling us that the matter – by which term I include both the electromagnetic field together with the actual material particles – is indeed in a very closely maximally entropic state. Where the initial very low entropy (a necessity for the 2^{nd} law) resided, was in *gravity* and *only* in the gravity! The puzzle about the origin of the 2^{nd} law lies in the very curious nature of the big-bang singularity, where the gravitational degrees of freedom appear to have been completely suppressed i.e. like Fig. 15 and not like Fig. 20.

Why is there such an extraordinary difference between the past-type and future-type singularities? As mentioned earlier, I used to think that there must be a very peculiar, blatantly time-asymmetric theory of quantum gravity governing this past/future distinction. Later, I simply *postulated* that past-type singularities must have vanishing Weyl curvature – i.e. vanishing gravitational degrees of freedom (what I called "the Weyl curvature hypothesis"). But this does not provide us with any kind of "physical reason" that, whereas the singularities of gravitational collapse must almost always have wildly diverging Weyl curvature, perhaps like the very exotic BKLM situation, the structure of the actual big bang appears to be quite the opposite.

As an approach to studying both the singularities and the asymptotic features of space-times, and also the massless fiends within them, I had come to realize the enormous value of looking at a space-time from the *conformal* point of view, rather than just its more restricting metric structure. The Weyl curvature is, after all, distinguished as describing the *conformal* curvature of a space-time. Moreover, massless fields, most notably Maxwell's electromagnetic field, all exhibit conformal invariance. Even more importantly, the conformal structure of any physically reasonable space-time is effectively the same as its *causal structure*, as we shall come to see very shortly.

Figure 22.

For the understanding of the asymptotic structure of a space-time, we need to understand what its "infinity" might be like. To this end, it is helpful to turn to Fig. 22, which shows (in his print "Circle limit I") how the Dutch artist M.C. Escher exquisitely illustrates how the infinity of the hyperbolic plane can be represented as a smooth circular boundary (Beltrami-Poincaré disc). Conformal maps preserve angles, rather than distances, so small shapes are accurately represented, but not necessarily their sizes. Note that, in this picture, the eyes of the fish-creatures remain exact circles, no matter how closely we approach the boundary.

Figure 23.

We can use similar representations for space-times, and to understand what is involved, it is useful to look first at Fig. 23 where, in addition to the null cone of Fig. 6, I have added some hill-shaped and bowl-shaped surfaces, marked in brown. These brown surfaces (in full 4-dimensional space-time being 3-dimensional surfaces) enable us to represent the full metric stricture at a point of space-time. They represent the ticks of identical clocks travelling at different velocities through the vertex point, as indicated in Fig. 24. At the bottom of this picture, I have written the two most famous formulae of 20th century physics, one of these being, of course, Einstein's well-known formula $E = mc^2$ (fundamental to relativity theory), and the other being Planck's earlier formula $E = h\nu$, (where ν is a frequency, fundamental to quantum mechanics). Einstein tells us that energy and mass are equivalent, and Planck tells us that energy and frequency are equivalent, so putting the two together we see that mass and frequency are equivalent (c and h being just conversion constants). This tells us that any stable massive particle is, in effect, a perfect clock! (The frequencies would be extremely high for individual fundamental particles, but appropriately scaled down, this gives us what is, in effect, the basis for atomic and nuclear clocks.) In fact, it is in the amount of "crowding" of these surfaces that the metric of space-time is defined.

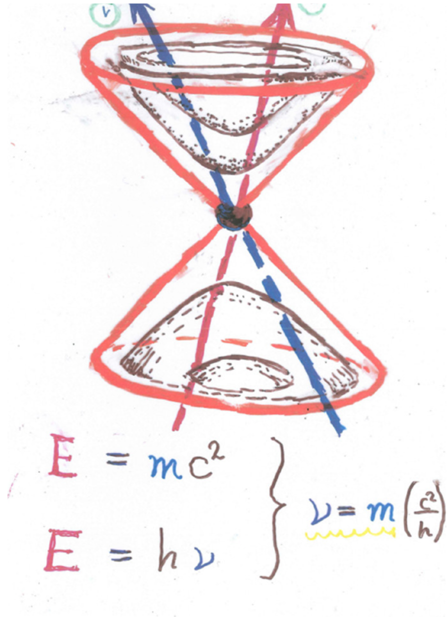

Figure 24.

Perhaps one usually thinks of a space-time's metric to be defining *distances* on an infinitesimal scale. However, it is much more physically direct to think of *times* as defining the metric structure. Spatial distances then arise as a secondary concept, determined in terms of times of transit, the time measures along world lines of particles being primary.

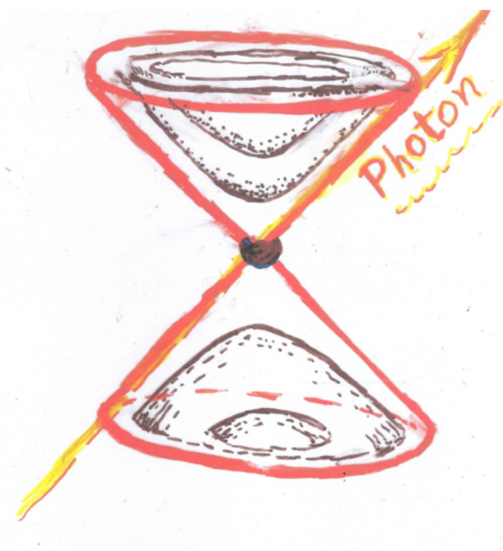

Figure 25.

But what about massless particles, where we now think of a photon in free space? We see from Fig. 25. that the photon does not even notice these scale-determining brown surfaces and does not "experience" the passage of time at all. To a photon, there is no elapse of time from one end of its trajectory to the other, so these scaling surfaces now play no role whatever, and we may as well remove them altogether, so we are left simply with the null cones themselves. Without the scaling, we simply have the metric up to proportionality, i.e. the space-time's *conformal* structure, i.e. that defined by the null cones themselves (Fig. 26, as in Fig. 6 and Fig. 7). It may be mentioned that, since causal signals are transmitted by effects on or within the null cones, the conformal stricture of space-time also defines its causal structure.

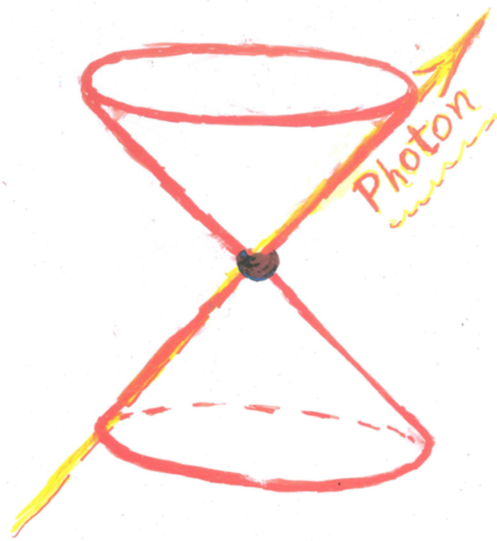

Figure 26.

Moreover, it is not just the *classical* picture of a photon that is concerned with only the conformal structure, rather than the full metric structure of space-time. James Clerk Maxwell's famous equations for the electromagnetic field can also be seen (when phrased appropriately) to be insensitive to the metric scaling, needing only the null-cone structure and not the full metric, i.e. Maxwell's equations are *conformally* invariant. Moreover, the Schrödinger equation for a photon's wave-function is, in effect, just Maxwell's equations and is therefore also conformally invariant. This conformal invariance would extend also to other massless particles and, in an appropriate sense, to gravitational wave propagation.

Now, as with Escher's representation of the infinity of the hyperbolic plane shown in Fig. 22, we can also use conformal re-scalings of the metric to get a good picture of the future infinity of cosmological models.

When there is a positive value for the cosmological constant Λ, we find that this future conformal infinity is space-like [25], so it represents this infinity as a temporal "moment", albeit a moment that would be at time "infinity" according to the normal space-time metric. This conformal "squashing down" of temporal infinity is represented at the top part of Fig. 27. This procedure is very general for space-times with positive Λ, as has been demonstrated by Helmut Friedrich [26], who showed that such a conformal future boundary is generic for space-times with $\Lambda > 0$ and massless field sources.

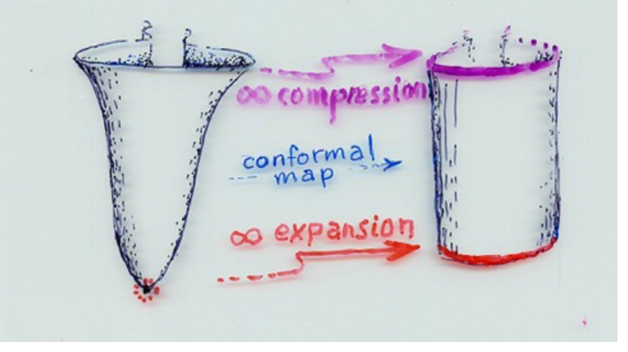

Figure 27.

This conformal squashing of the future can be accompanied by a conformal stretching out of the big bang to obtain a smooth spacelike past boundary, as illustrated at the bottom part of Fig. 27. This is a standard procedure that can be applied to most conventional cosmological models [25], [27]. However, these models assume isotropy and homogeneity, and therefore do not address the issue, raised earlier, that singularities of the type illustrated in Fig. 20 would have been vastly more probable, and we need some form of "Weyl curvature hypothesis", asserting that "past-type" space-time singularities must have highly restricted (perhaps zero) Weyl curvature, while no such restriction would be appropriate for future-type space-time singularities.

Yet, we definitely need such a hypothesis to restrict the possibilities for the *past* if we are to describe space-time models which have any real hope of describing the actual universe, in which there is a 2^{nd} Law of Thermodynamics in accordance with what is observed, for which the low entropy in the early universe arises from the initial suppression of gravitational degrees of freedom. As was remarked earlier, the presence of a very early phase of the universe in which there was an inflationary expansion does not in itself resolve this issue (and it is my own opinion that there was actually no such inflationary phase – an issue that we shall need to return to shortly). In any case, inflation or no inflation a huge constraint on the Big Bang is required, which is indeed of the nature of some kind of "Weyl curvature hypothesis", as appears to be a feature of the actual universe in which we find ourselves. Accordingly, the ontological status of adopting each of the two conformal boundaries, depicted in Fig. 27, could hardly be more different, the one in the future being a generic procedure, very broadly applicable, and imposing no significant constraint on the applicability of the procedure being presented, whereas the one in the past provides an enormous restriction on the type of universe model under serious consideration.

Nevertheless, we can regard such a restriction as indeed formulating a version of Weyl curvature hypothesis, where we may regard such a hypothesis as an essential feature of any universe model having a chance of representing the actual world that we see around us. In fact, it was Paul Tod (of the university of Oxford, and a former graduate student of mine) who first formally proposed, and then studied in detail, this form of Weyl curvature hypothesis [28] – namely that for our actual universe, the stretching procedure, illustrated at the bottom of Fig. 27 should result in a smooth initial spacelike hypersurface boundary, this being a far more attractive and mathematically tractable procedure than my original rather vague form of this hypothesis. Tod's procedure allows detailed calculations to be performed, and this enables the implications of the hypothesis to be studied in some considerable detail.

With regard to the future conformal infinity, it had been a useful "mathematical trick" in the study of gravitational radiation, etc., to imagine that the future infinity could be conformally extended smoothly to a fictional space-time continuation beyond this future infinity [29], [30]. But now we can imagine that a kind of "time-reverse" of this trick is applied to the Big Bang, where we contemplate a fictional pre-Big-Bang extension of our universe. Thus, not only can we imagine a fictional extension of the universe's future to some kind of world beyond, but we can also contemplate a fictional world that extends our universe in a conformally smooth way to some world prior to our Big Bang.

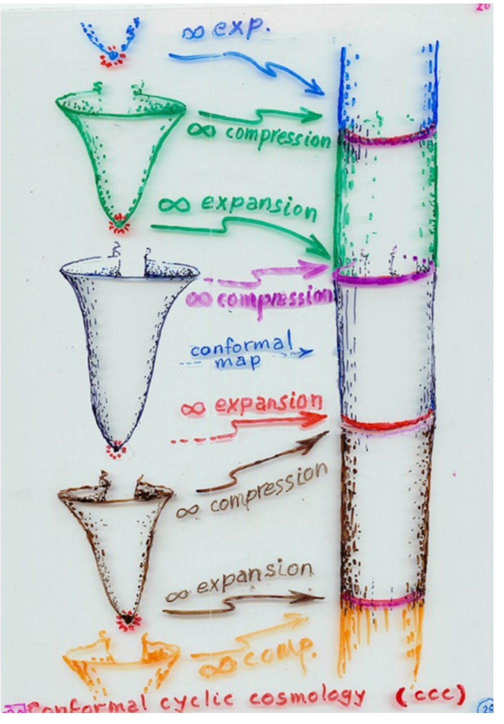

Figure 28.

This may not be quite the usual way that cosmologists have thought to picture the universe, but there is nothing outrageous about it. What may actually be regarded as outrageous, on the other hand, is the picture presented in Fig. 28, where the view is taken that neither of these conformal extensions is taken to be "fictional", but *both* are regarded as being actually *real*. The new feature about these extensions is that we do not continue the Big Bang singularity to another future-like singularity or perhaps a remote future boundary to a remote past-type boundary, but we preserve the time directions, to continue our non-singular remote future boundary to a succeeding big bang singularity, and precede our Big Bang singularity by a non-singular remote future boundary, thereby automatically *forcing* our Big Bang to satisfy a Weyl curvature hypothesis, as appears to be required.

Such a picture makes geometrical sense, but we must ask whether it can possibly make *physical* sense. On first consideration one might well take the view that this is unreasonable because the remote future is extremely cold, and the density very low, whereas at the Big Bang, things were very much the opposite, with an extraordinarily high temperature and density. However, when you conformally rescale things, the conjugate variables go the opposite way. Time scales oppositely to energy and space scales oppositely to momentum. Thus, the very cold and rarefied remote future re-scales to the very hot and dense next big bang, this being consistent with the very large measures of space and time in the remote future rescaling to very tiny measures of space and time in the next big bang, all this being consistent with the model that I have been proposing [31].

There is also an issue about the cosmological constant Λ. This needs to be positive for the scheme to work. The stretched-out Big Bang appears to be space-like in all serious models, but for the scheme to work we are restricted to those for which the conformal infinity is also space-like, this corresponding to $\Lambda > 0$ which, fortunately for the scheme, appears to be the case [25]!

I refer to this scheme as *conformal cyclic cosmology* or CCC for short. The portion of the sequence from a big bang moment to its following remote future I call an aeon. I adopt the conceptually simplest version of CCC that there is an infinite succession of aeons, infinite in both directions, though other possibilities might also be considered. I also adopt the view that the aeons are qualitatively similar to one another so that, the constants of nature do not vary from aeon to aeon, but other possibilities are certainly open to consideration. I tend to use the capitalised “Big Bang”, when this refers to the specific moment that initiated our current aeon, and “big bang” otherwise.

A comment needs to be made in relation to inflation. Certainly, CCC is not compatible with the version of inflation that is currently favoured by many cosmologists, because this would provide a causal gap between aeons that would ruin the observational issues that will be described below, though a very small inflationary phase could be considered. A more satisfactory CCC picture would be to eliminate inflation altogether, the hope being that the things that inflation is useful for in cosmology can be taken over by the final exponential behaviour of the previous aeon, which plays any role that in conventional cosmology an exponentially expanding phase seems to be required. In short, CCC does provide an “inflationary phase” in effect, but it occurred prior to the Big Bang rather than following it!

Our final issue has to do with possible observational tests of CCC. In fact, there are several of these, especially if one takes the view that the various cosmic aeons are necessarily qualitatively similar to each other. Then there are certainly issues about physical parameters having to match on two sides of the crossover 3-surface joining one aeon to the next. Another possibility might be signals, such as electromagnetic ones getting through from one aeon to the next, as indicated in Fig. 29. These might be important magnetic fields getting across, for example, but that has not yet been looked at.

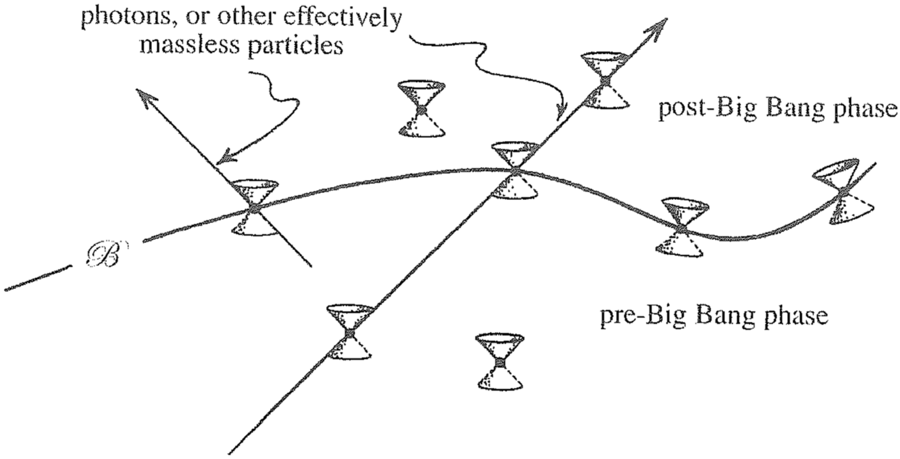

Figure 29.

An alternative, which has indeed been examined is the possibility of gravitational wave signals getting across. These could certainly get through, in the CCC picture. We can imagine that encounters between supermassive black holes within a previous-aeon galactic cluster should encounter one another from time to time, and in doing so, should emit enormous amounts of energy in the form of gravitational waves. Such waves ought certainly to be able to get through into our aeon after smoothly propagating through the crossover between the two. When in our aeon, such waves would transfer some of their energy into electromagnetic form and slightly affect the temperature of our microwave background, each wave often making a large ring of such slightly increased temperature. Such a ring can be understood as being the intersection of the sphere W^2 here the gravitational wave encounters our last scattering 3-surface L^3 win the sphere M^2 which is where our past light cone meets L^3, so that M^2 is, is fact, our own microwave background sky. The circle in the microwave sky that we are looking for is the intersection W2∩M2 of these two spheres. Rather remarkably, two groups appear to have independently observed rings consistently with such expectations. One was a Polish group, [32], [33], who regard their observations to support our CCC expectations, with around 99.5% confidence level.

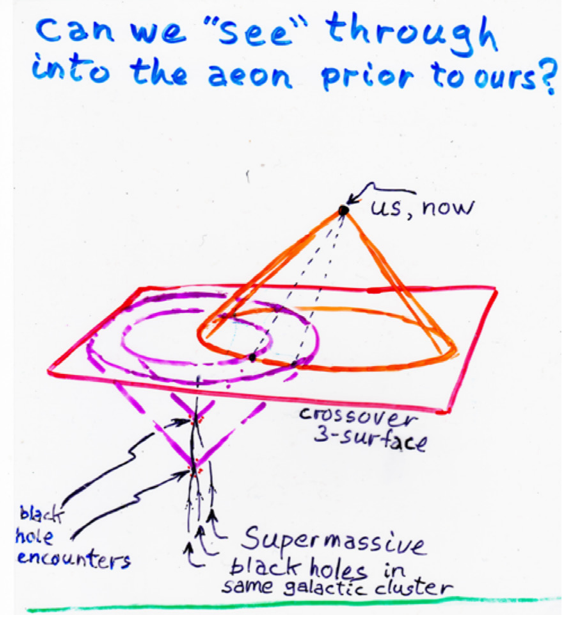

Figure 30.

The other is a collaboration between myself and my Armenian colleague Vahe Gurzadyan. In order to get a significant enough signal, we took advantage of the fact that within any particular large-enough galactic cluster in the previous aeon there ought normally to be several such encounters between supermassive black holes. Each such black-hole encounter should provide one of these rings in our CMB, but since there should be several of these events within the same cluster, these should provide rings with the same centre (see Fig. 30). In our papers [34], [35] we considered that the temperature disturbance in the CMB should be able to be seen as a reduced temperature variance around the ring, and to get a strong enough signal, we looked for occurrences where there are at least 3 different low-variance rings with the same centre, being from the same cluster. This would be plotted as a single point in the map of our CMB sky. The results are illustrated in Fig. 31, and we find what we regard as a rather remarkable effect, First of all, we notice that the points, indicating the ring centres and therefore, according to the theory, the locations of galactic clusters in the previous aeon, are extremely clustered, and certainly not uniformly distributed across the sky, as would have been expected according to conventional ideas about the uniformity of the universe, that should reveal itself on a large-enough scale. Moreover, the points are also distinctively clustered with regard to colour, this colour coding referring to the temperature assigned to the individual ring whose centre is being marked in the picture. We must also recall that the criterion for selecting the points was not the overall energy (temperature) of the ring, but the fact that the *variance* in the temperature was low. The actual temperature is an independent parameter, and (according to the theory) should be a signal of the distance of the source of the signal from us. The theory says that the red points in the picture are extremely distant and the blue ones somewhat less distant. Accordingly, according to the theory, the galactic clusters that we seem to be seeing are very non-uniformly distributed, not only in their angular separation across the sky, but also in the distance away from us!

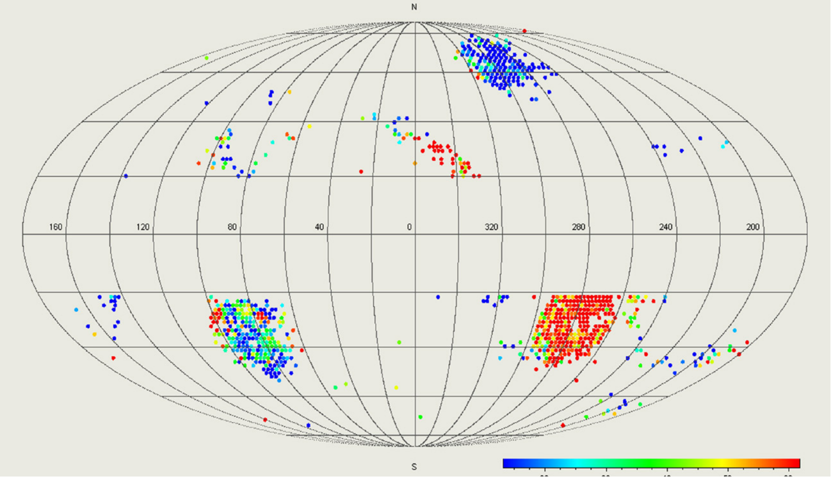

Figure 31.

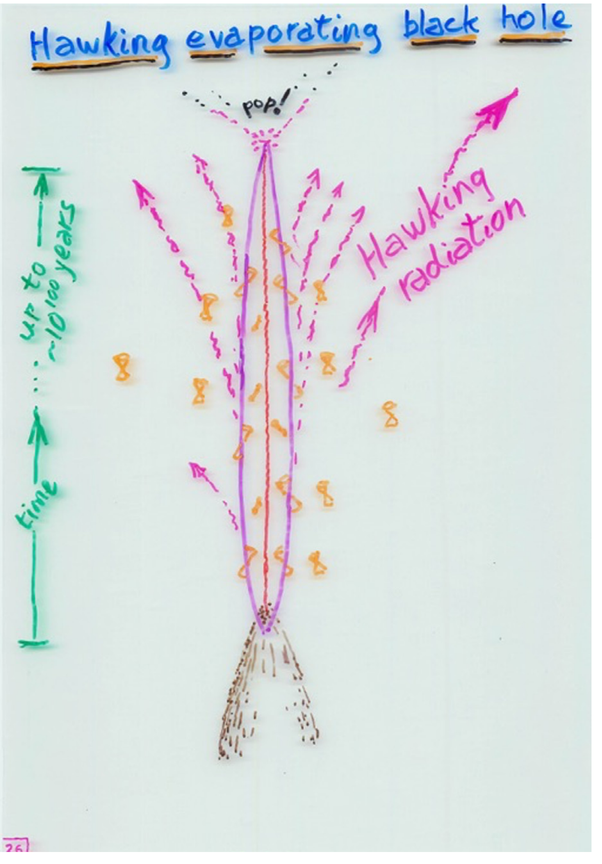

Figure 32.

Finally, we should consider the completely different effect illustrated in Fig. 32. This concerns the ultimate phenomenon that we presume to have taken place in the previous aeon. According to the effect predicted by Hawking, a black hole ought to have a very tiny temperature, referred to as the Hawking temperature, which for a very large black hole would be exceptionally tiny. Nevertheless, as the universe expands, the temperature of the universe gets lower and lower until it becomes smaller even than the Hawking temperature of the supermassive black hole, at which point the hole itself begins to evaporate away, all of its enormous mass being eventually radiated away into this Hawking radiation. However, because this occurs so extremely late in that aeon's existence (perhaps at least 10^{100} years), this entire activity occupies what is effectively a single point on the crossover surface, as exhibited in Fig. 33. However, the mass-energy cannot be lost, and it bursts through into the subsequent aeon (our aeon) at a single point that we refer to as a *Hawing point*. The energy bursting through at that point would disperse itself through that early material reaching a certain diameter until revealing itself as a heated spot of a certain diameter on the last scattering surface. See Fig. 34. It turns out that we actually see such spots in the CMB sky with an angular diameter of about 4° radians (about 8 times the diameter of the full moon), which is close to what one should expect on theoretical grounds, with a confidence level of about 99.98%. [36]

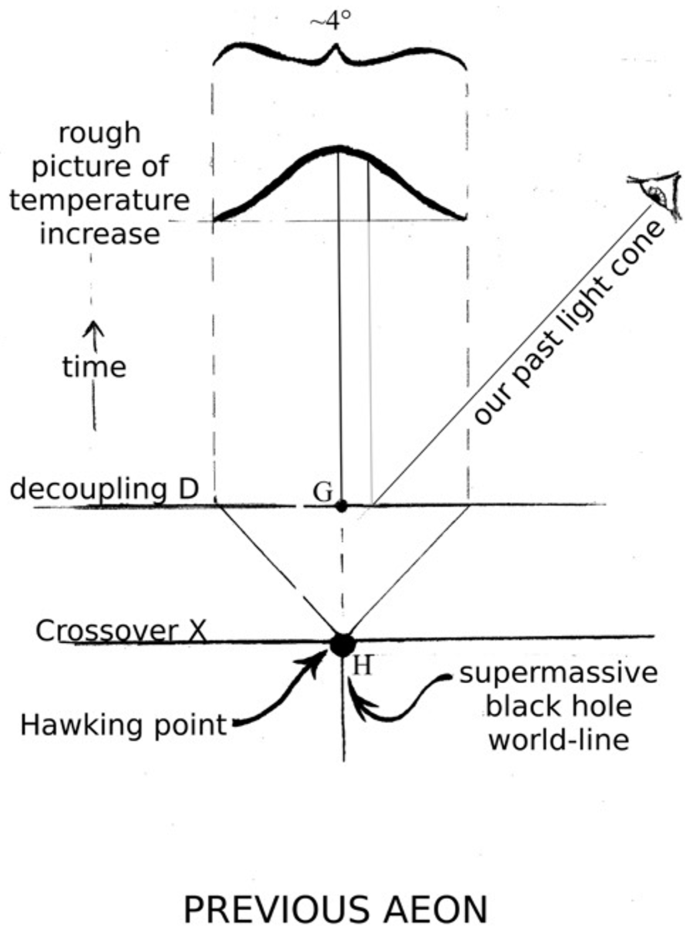

Figure 33.

We have been provided with two different kinds of signal that seem to exhibit rather strong evidence for a cosmic aeon prior tours, in accordance with CCC. A natural question to ask is whether there might, in addition, be some interconnection between the two. This ought to be the case, but with the data currently available the connection remains a somewhat weak one so far. The gravitational waves accessible to us from a galactic cluster that ends up as a Hawking point that we can directly see would have to b "edge-on"' to our line of sight and would therefore be marked as "green" (middle-range temperature) in Fig. 31. Of the 5 points that are clearly seen in both the Planck and WMAP data at exactly the same places in each [36], it is noteworthy that 3 of then are remarkably close to such green points in [35]. Various factors could contribute to any slight discrepancies, such as imprecision in locating the exact centres of the rings in question or even proper motions building up over the vast intervening timescales involved.

Warm thanks go to Dennis Lehmkuhl for several historical clarifications.

REFERENCES

It should be noted that the "phi" coordinates in [35] and [36] differ by a right/left reflection.

1. Oppenheimer, J.R. and Snyder, H. (1939) On continued gravitational contraction. *Phys. Rev*. **56**, 455–9.
2. Chandrasekhar, S. (1931) The maximum mass of ideal white dwarfs. *Astrophys. J.,* **74**, 81–2.
3. Finkelstein, D. (1958) Past-future asymmetry of the gravitational field of a point particle, *Phys. Rev.* **110**, 965–7.
4. Painleve, P (1921) La mécanique Classique et la théorie de la relativitité. *R. Acad. Sci*. (Paris) **173**, 677–80
5. Lemaître, G. (1933) L'universe en expansion *Ann. Soc. Sci. Bruxelles* I **A53**, 51–85 (cf. p. 82).
6. Dirac, P.A.M. (1936) Relativistic wave equations. *Proc. Roy. Soc*. (Lond.) **A155**, 447–59.
7. Payne, W.T. (1952) Elementary spinor theory *Am. J. Phys.* **20**, 253–62.
8. Penrose, R. and Rindler, W. (1984) *Spinors and Space-Time, Vol. 1: Two-Spinor Calculus and Relativistic Fields* (Cambridge University Press, Cambridge).
9. Bade, W.L. and Jehle, H. (1953). An introduction to spinors. *Rev. Mod. Phys.* **25**, 714–28.
10. Lifshitz, E.M. and Khalatnikov, I.M. (1963) Investigations in relativistic cosmology, in *Advances in Physics* **12**,185–249.
11. Penrose, R. (1965) Gravitational collapse and space–time singularities, *Phys. Rev. Lett*. **14**, 57–59.
12. Penrose, R. (1972) Techniques of Differential Topology in Relativity, *CBMS Regional Conf. Ser. in Appl. Math.*, No. 7 (S.I.A.M., Philadelphia).
13. Hawking, S.W. and Ellis, G.F.R. (1973) *The Large-Scale Structure of Space-Time* (Cambridge University Press, Cambridge).
14. Belinskii, V.A., Khalatnikov, I.M. and Lifshitz, E.M. (1970) Oscilliatory

approach to a singular point in the relativistic cosmology, *Usp. Fiz. Nauk* 102, 463–500. Engl. transl. in *Adv. in Phys.* **19**, 525–73.

15. Misner, C.W. (1969) Mixmaster universe. *Phys. Rev. Lett.,* **22**, 1071–4.
16. Penrose, R. (1973) Naked singularities, *Ann. N.Y. Acad. Sci.*, **224**, 125–34.
17. Hawking, S.W. (1965) Occurrence of singularities in open universes. *Phys. Rev. Lett.* **15**, 689–90.
18. Hawking, S.W. (1966) The occurrence of singularities in cosmology. *Proc. Roy. Soc.* (London) **A294**, 511–21.
19. Hawking, S.W. (1966) The occurrence of singularities in cosmology II. …. . *Proc. Roy. Soc.* (London) **A295**, 490–3. [Adams Prize Essay: Singularities in the geometry of space-time.]
20. Hawking, S.W. (1967) The occurrence of singularities in cosmology III. Causality and singularities. *Proc. Roy. Soc.* (London) **A300**, 187–201.
21. Hawking, S.W. and Penrose, R. (1970) The singularities of gravitational collapse and cosmology *Proc. Roy. Soc.* (London) **A314** (1519) 529–48.
22. Linde, A.D. (1982) A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B* **108**:389–93
23. Guth, A.H. (1981) Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev.* **D23**, 347–56.
24. Guth, A.H. (1997) *The Inflationary Universe* (Jonathan Cape, London).
25. Penrose, R. (1964) Conformal approach to infinity, in Relativity, Groups and Topology: *The 1963 Les Houches Lectures,* eds. B.S. DeWitt and C.M. DeWitt (Gordon and Breach, New York).
26. Friedrich, H. (1998) Einstein's equation and conformal structure. In T*he Geometric Universe; Science, Geometry, and the Work of Roger Penrose*, eds. S.A. Huggett, L.J. Mason, K.P. Tod, S.T. Tsou, and N.M.J. Woodhouse (Oxford Univ. Press, Oxford).
27. Penrose, R. (2004) *The Road to Reality: A Complete Guide to the Laws of the Universe* (Jonathan Cape, London). Vintage IBN: 9780-679-77631-4.
28. Tod, K.P. (2003) Isotropic cosmological singularities: other matter models. *Class. Quantum Grav*. **20**, 521–534 doi:10.1088/0264-9381/20/3/309.
29. Penrose, R. (1963) Asymptotic properties of fields and space-times, *Phys. Rev. Lett.* **10**, 66–8.
30. Penrose, R. (1965) Zero rest-mass fields including gravitation: asymptotic behaviour, *Proc. Roy. Soc. London*, **A284**, 159–203.
31. Penrose, R. (2010) *Cycles of Time: An Extraordinary New View of the Universe.* (Bodley Head, London). ISBN 978-0-224-08036-1; 9780099505945.
32. Meissner, K. A., Nurowski, P. and Ruszczycki, B. (2013) Structures in the microwave background radiation, *Proc. R. Soc*. A469:2155, 20130116, arXiv:1207.2498[astro-ph.CO].
33. D. An, Meissner, K. A. Nurowski, P, *Mon. Not. Roy. Astron. Soc.* **473** (2018) no. 3, 3251.
34. Gurzadyan, V.G. and Penrose R. (2013) On CCC-predicted concentric low-variance circles in the CMB sky. *Eur. Phys. J. Plus* 128: 22 DOI 10.1140/epjp/i2013-13022-4.
35. Gurzadyan, V.G. and Penrose R. (2016) CCC and the Fermi paradox *Eur Phys. J. Plus* (2016) 131:11.
36. An, D, Meissner, A, Nurowski,P, Penrose, R, (2020) Apparent evidence for Hawking Points, *Mon. Not. Roy. Astron. Soc*. **495** no. 3 3403.