12. Singularities and time-asymmetry

R. PENROSE

12.1 Introduction

It has been a source of worry to many people that the general theory of relativity – that supremely beautiful description of the geometry of the world – should lead to a picture of spacetime in which singularities are apparently inevitable. Einstein himself had fought against the inevitability of such seeming blemishes to his theory, suggesting different possible ways out, of considerable ingenuity (e.g. the Einstein–Rosen bridge,[1] the attempt at a black-hole-avoiding stable relativistic star cluster,[2] the idea that a non-singular 'bounce' of the universe might be achieved through irregularities,[3] even his attempts at modifying general relativity to obtain a singularity-free unified field theory[4,5]). Yet, the researches of theorists in more recent years have driven us more and more in the direction of having to accept, and face up to, the existence of such singularities as true features of the geometry of the universe.

This is not to say that some mathematically precise concept of 'singularity' should now form part of our description of physical geometry – though much elegant work has been done in this direction in recent years. Rather, it seems to be that the very notion of spacetime geometry, and consequently the physical laws as we presently understand them, are limited in their scope. Indeed, these laws are even *self*-limiting, as the singularity theorems[6,7,8] seem to show. But I do not feel that this is a cause for pessimism. There is a need for new laws in any case; while, in my opinion, the presence and the apparent structure of spacetime singularities contain the key to the solution to one of the long-standing mysteries of physics: the origin of the *arrow of time.*

The point of view I am going to present does not stem from any radical view of things. I shall adopt a basically conventional attitude on most issues – or so it would have seemed, were it not for the fact that my ultimate conclusions appear to differ in their basic essentials from those most commonly expressed on this subject! My arguments do not depend on detailed calculations, but on what seem to me to be certain 'obvious'

facts, whose very obviousness may contribute to their being frequently overlooked.

It was Einstein who was, after all, the supreme master at deriving profound physical insights from 'obvious' facts. I hope I may be forgiven for trying to emulate him on his hundredth birthday, for I feel sure that he would have cared not one fig for the supposed significance of so arbitrary an anniversary!

12.2 Statement of the problem

The basic issue is a familiar one.[9,10,11] The local physical laws we know and understand are all symmetrical in time. Yet on a macroscopic level, time-asymmetry is manifest. In fact, a number of apparently different such macroscopic arrows of time may be perceived. To these may be added the only observed time-asymmetry of particle physics – which features in the decay of the K^0-meson. And there is one further related issue, namely the interpretation of quantum mechanics, which I feel should not be banished prematurely from our minds in this connection. The conventional wisdom has it that, despite an initial appearance to the contrary, the framework of quantum mechanics contains no arrow.[9,12–16] I do not dispute this wisdom, but nevertheless believe, for reasons that I shall indicate, that the question must be kept alive.

Let me list, therefore, seven apparently independent arrows (or possible arrows) that have been discussed in the literature:[9] section 12.2.1, K^0-meson; section 12.2.2, quantum-mechanical observations; section 12.2.3, general entropy increase; section 12.2.4, retardation of radiation; section 12.2.5, psychological time; section 12.2.6, expansion of the universe; and section 12.2.7, black holes versus white holes. I shall discuss each of these in turn.

12.2.1 Decay of the K^0-meson

Can the asymmetry that is present in the decay rate of the K^0-meson have any remote connection with the other arrows? The effect, after all, is utterly minute. The T-violating component in the decay is perhaps only about one part in 10^9 of the T-conserving component[17–21] – and, in any case, the presence of this T-violation has to be inferred, rather than directly measured, from the presence of a minute CP-violation ($\sim 10^{-9}$) together with the observation that CPT-violation, if it exists, must be even smaller in this interaction ($\ll 10^{-9}$). A very weak interaction indeed

(or very weak component of a weak interaction) seems to be involved, and it plays no significant role in any of the important processes that govern the behaviour of matter as we know it. Gravitation, of course, is even weaker, and does, in fact, dominate the motion of matter on a large scale. But the differential equations of general relativity are completely time-symmetric as are Maxwell's equations and, apparently, the laws of strong interactions and ordinary weak interactions.

Yet the tiny effect of an almost completely hidden time-asymmetry seems genuinely to be present in the K^0-decay. It is hard to believe that Nature is not, so to speak, 'trying to tell us something' through the results of this delicate and beautiful experiment, which has been confirmed several times.[20] One of the suggestions that was put forward early on was that the T-violating effect arose via some cosmological long-range interaction, whereby matter-tô-antimatter imbalance provided the required asymmetry.[50] But subsequent analysis[19,20] has rendered this viewpoint implausible. It seems that the asymmetry is really present in the *local* dynamical laws. This is a matter that I shall return to later. I believe that it is a feature of key significance.

12.2.2 Quantum-mechanical observations

In standard quantum mechanics, the dynamical evolution of a state takes place according to Schrödinger's equation. Under time-reversal this equation is transformed to itself provided i is replaced by $-$i. But Schrödinger's equation must be supplemented by the procedure ('collapse of the wavefunction') whereby the current state vector is discarded whenever an 'observation' is made on a system, and is replaced by an eigenstate ψ_Q of the Hermitian operator Q which represents the observable being measured. As it stands, this procedure is time-asymmetric since the state of the system is an eigenstate of Q just after the observation is made, but not (normally) just before (figure 12.1(a)). However, this asymmetry of description is easily remedied:[14] in the time-reversed description, one simply regards this same eigenstate ψ_Q as referring, instead, to the state just before the observation and Schrödinger's equation is used to propagate backwards until the previous observation (with operator P) is reached. Thereupon this (backwards-evolved) state vector is discarded (according to a time-reversed version of the 'collapse of the wavefunction') and an eigenstate ψ_P of P (corresponding to the actual result of the observation P) is substituted (figure 12.1(b)).

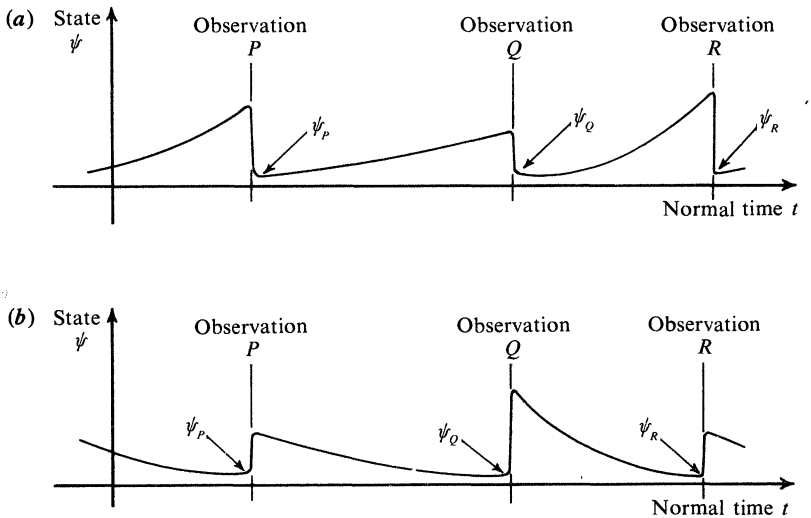

Figure 12.1. (*a*) Conventional (Schrödinger) picture of development of a wavefunction. (*b*) Essentially equivalent 'time-reverse' development of a wavefunction.

The relative probability of the observed Q-value, given the observed P-value, or of the observed P-value, given the observed Q-value, is the same in each mode of description, being

$$|\langle\psi_Q|U\psi_P\rangle|^2 = |\langle\psi_P|U^{-1}\psi_Q\rangle|^2,$$

where U is the unitary operator representing the evolution (according to Schrödinger's equation) of a state from the time of P to the time of Q. Since these probabilities are the only observational manifestation of the wavefunction ψ in any case, we see that the two modes of description are equivalent, and that the framework of quantum mechanics is time-symmetric.

It is often stated[16,22] that the actual value of the wavefunction at any time is not properly a description of physical reality. This is strikingly illustrated by figure 12.1, if time-symmetry is to be a feature of microscopic physics. But in any case there is a well-known difficulty, even with special relativity, concerning taking too strong a view that the wavefunction describes physical reality. For an 'observation' would seem to collapse the wavefunction into one of its eigenstates simultaneously with that observation – where 'simultaneous' presumably refers to the rest-frame of the one making the observation. This can lead to conceptual problems, when two spacelike-separated observations are carried out,

concerning the question of the *ordering* in which two collapses of the wavefunction take place. The difficulty is pinpointed particularly well in the famous thought experiment of Einstein, Podolsky and Rosen.[22]

There are, however, *other* situations in which it seems hard to maintain the view that the wavefunction (or state vector) does *not* give a proper description of physical reality in accordance with figure 12.1(*a*). Consider an isolated system on which observation P has just been made, giving a *conventional* description of a state ψ_P (eigenstate of P) which evolves forwards in time according to Schrödinger's equation to give a state $U\psi_P$ at some later time. Now (according to the conventional framework of quantum-mechanics, and assuming that no additional principles are incorporated) the operator UPU^{-1} is just as 'good' an observable as P. Furthermore, provided that the eigenvalue λ corresponding to ψ_P is simple, $U\psi_P$ has the property that it (or a nonzero multiple of it) is the *unique* state for which the probability is unity of giving the value λ for the observation UPU^{-1}. The isolated system cannot (in the ordinary way of looking at things) 'know' that the observation UPU^{-1} may be about to be performed upon it, but it must be prepared for that eventuality! So it seems that the information of $U\psi_P$ (up to phase) must be stored in the system, i.e. that the wavefunction *does* describe physical reality.

Of course UPU^{-1} may correspond to an utterly outlandish and totally impracticable experiment, as, for example, if in the case of the Schrödinger 'cat paradox'[24] we were to attempt to verify a resulting state $U\psi_P$: 'complex linear combination of dead cat and alive cat'. The very outlandishness of such an experiment suggests that $U\psi_P$ (and UPU^{-1}) may not, after all, 'really' refer to reality! But that is my whole point. There is something missing or something inappropriate about the laws of quantum mechanics, when applied to such situations. There is also something absurd about the whole idea of a collapsing wavefunction (or of any of the other essentially equivalent alternative ways of describing the same phenomenon, such as the conscious observer threading his way through an Everett-type[26–29] many-sheeted universe) as a description of physical reality. Yet what is a physical theory for if not to describe reality? In this I feel that I must align myself with many of the originators and main developers† of quantum mechanics – and, not least, with Einstein himself[23] – in believing that the resolution of the question of 'observations' is not to be found within the formalism of quantum mechanics itself. Some new (presumably nonlinear) theory seems to be required in

† E.g., in their different ways, Bohr,[30] Schrödinger,[24] Dirac,[31] Wigner.[25]

which quantum mechanics and classical mechanics each emerges as a separate limiting case.

The issue has importance here for two reasons. In the first place, the making of an observation seems to be associated with an *irreversible* process, depending upon an essential entropy increase. It is not obvious that what is missing (or wrong) about quantum mechanics is not some fundamentally *time-asymmetric* law. So the demonstration of time-symmetry in the formalism of quantum mechanics does not really settle the question of time-(a)symmetry in quantum-mechanical observations.

The second reason why the issue is important here has to do with the question of the role of quantum gravity. One must bear in mind the possibility that it might be the presence of a (measurable) *gravitational field* that takes the description of a physical system out of the realm of pure quantum physics.[32,33] And if a *new law* is needed, especially to cover situations in which both quantum and classical physics are being stretched in the extreme, then quantizers of gravity, beware!

I shall return to these questions in section 12.4. But, for the present, I propose to be wholly conventional in my attitude to quantum mechanics: it contains no manifest arrow, and the solution to the problem of macroscopic time-asymmetry must be sought elsewhere.

12.2.3 General entropy increase

The statistical notion of entropy is, of course, crucial for the discussion of time-asymmetry. And if the (important) local laws are all time-symmetric, then the place to look for the origin of statistical asymmetries is in the boundary conditions. This assumes that the local laws are of the form that, like Newtonian theory, standard Maxwell–Lorentz theory, Hamiltonian theory, Schrödinger's equation, etc., they determine the evolution of the system once boundary conditions are specified, it being sufficient to give such boundary conditions *either* in the past *or* in the future. (The boundary values are normally specified on a spacelike hypersurface.) Then the statistical arrow of time can arise via the fact that, for some reason, the *initial* boundary conditions have an overwhelmingly *lower* entropy than do the *final* boundary conditions.

There are several issues that must be raised here before we proceed further. In the first place, there is something rather unreasonable about determining the behaviour of a system by specifying boundary conditions at all, whether in the past *or* future. The 'unreasonableness' is particularly apparent in the case of future boundary conditions. Suppose I throw my

watch on a stone floor so that it shatters irreparably, and then wait for 10 minutes. The future boundary condition is a mess of cogs and springs, but with minutely organized velocities of such incredible accuracy that when reversed in direction (i.e. with the clock run backwards) they suddenly reassemble my watch after a 10-minute period of apparent motionlessness. Though the models of physics that we are using (e.g. Newtonian theory) allow, in principle, such accuracy to be defined, we do not know (nor do we demand) that our models of physical laws correspond with such precision to reality. The problem is there also even for past boundary conditions, as stressed by Born.[34] And Feynman[35] has pointed out that, in Newtonian theory, if all the positions and velocities of a complex system are known to a certain accuracy, then all the accuracy is lost in less time than it takes to state that accuracy in words! Born (and Feynman) invoke this argument to demonstrate that classical mechanics is, in a sense, no more deterministic than quantum mechanics. Of course quantum mechanics has the additional problem of what happens when an 'observation' is made – and 'observations' seem to be necessary, in the normal view, to keep the wavefunctions from spreading throughout space.

I mention such things mainly to point out difficulties. But I shall ignore them henceforth and follow the conventional path that boundary conditions work! There is, however, a somewhat related question that needs further comment. Consider, again, my shattered watch as a future boundary condition. It will be seen that though this state has a higher entropy than that before my watch was shattered – and, therefore, in the normal way of looking at things, is a less 'unusual' state than the earlier one – the later state is nevertheless a very strange one indeed when one comes to examine it in minute detail, in view of the very precise correlations between the particle motions. But again I shall adopt a conventional 'macroscopic' view here. This strangeness is not of the kind that is described as 'low entropy'. And the 'reason' that I had a watch earlier is not that these precise correlations exist in the *future* boundary conditions, but that something in the *past* (say a watch factory) had a lower entropy than it might otherwise have had. Likewise, the 'reason' for the precise correlations in the particle motions of the shattered watch can be traced back to the factory, not the other way around.

I do not feel that this is begging the question of time-asymmetry. It could perfectly well have been the case, in a suitably designed universe, that some processes behave like my watch, while others (using the time-sense defined by my watch while it was still working!) indulge in apparently miraculous assembly procedures which suggest that special

(low-entropy) *future* boundary conditions should be invoked to provide the 'reason' for *their* behaviour. But our universe seems not to work that way.

An important related question that I have glossed over so far is that of coarse-graining.[36] What does entropy mean anyway? Is it a definite physical attribute of a system which, like energy–momentum, seems to be independent of the way that we look at it? In practice, entropy can normally be treated that way (for example, in physical chemistry). But for the most general definitions of entropy that might be expected to apply to a complicated system such as a watch, we need some apparently rather arbitrary (i.e. non-objective) way of collecting together physical states into larger classes (coarse-graining) where the members of each class are considered to be indistinguishable from one another. The entropy concept then refers to the classes of states and not to the individual states. Then, the (Boltzmann) entropy of a class containing N distinct (quantum) states is

$$S = k \log N$$

(where k is Boltzmann's constant). In fact, several conceptually different definitions of entropy are available.[36–38] But the whole question is clearly fraught with difficulties.† (The phenomenon of 'spin-echo' is one striking example that emphasizes these difficulties.[39] I am not even convinced that 'entropy increase' is at all an appropriate concept for describing the shattering of my watch. Probably taking a bath increases the entropy enormously more – while, in the case of my watch, the proportional increase in entropy must be quite insignificant.) The question of the objectivity of entropy will be returned to in section 12.4. But for the moment I hastily take refuge once more in conventionality: entropy is a concept that may be bandied about in a totally cavalier fashion!

12.2.4 Retardation of radiation

The question of boundary conditions is also intimately involved in the next of our arrows, that of retarded radiation. We may separate this phenemenon into two quite distinct aspects: the entropy question again, and the question of source-free or sink-free radiation. Retardation is not just a feature of electromagnetic radiation, of course, though it is usually

† I am leaving aside also such important questions as the 'H-theorem'[36–38] and 'branch systems'.[10,11] They do not *explain* the time-asymmetric origin of the total entropy imbalance.

discussed in that context. Imagine a stone thrown into a pond. We expect to see ripples expanding outwards from the point of entry to have their energy gradually dissipated, especially when they hit the bank. We do not expect to see, before the stone reaches the water, ripples being produced at the bank with such precise organization that they converge upon the point of entry at the exact moment that the stone enters the water. Still less do we expect to see such ripples converging inwards from the bank to some point in the middle of the pond, at which they entail the sudden ejection of a stone into the air! Such extraordinary behaviour is perfectly in agreement with the local physical laws. But its occurrence would require the sort of precise correlations in particle motions that could only be explained by some low-entropy future boundary condition.

I should emphasize, once more, a point made in the last section, since I feel that it is a key issue: correlations are present in the detailed particle motions in the future *because* the entropy was low in the past. Likewise, similar (but time-reversed) correlations are absent in the past *because* the entropy is high in the future. The latter statement is an unusual form of words, but I am trying to be unbiased with respect to time-ordering.† My point of view is that the correlations are not to be viewed as the 'reason' for anything; but low entropy (itself to be explained from some other cause) *can* provide the 'reason' for the correlations. (This way around we avoid the problem of over-extreme precision in physical laws.) And once more I stress that the 'specialness' of a state, due to its possessing intricate particle correlations of this kind, is *not* the type of 'specialness' that *at that time* corresponds to a low entropy. That is an essential point of coarse-graining.

So we see that the normal retarded behaviour of the ripples correspond to low entropy in the past and correlations in the future, while the two situations I have described, which seem to involve *advanced* behaviour of the ripples, involve some very precise correlations in the past of the kind leading to a reduction in entropy. Furthermore, an alternative hypothetical *retarded* situation, in which a stone is suddenly ejected from the pond accompanied by ripples propagating outwards towards the bank,‡ also involves such precise correlations (this time in the motions of the

† This leads to a logical reversal of the viewpoint expressed by O. Penrose and Percival[40] in their 'law of conditional independence', according to which the absence of initial correlations is *postulated*. The world-view of section 12.3.3 provides a certain justification for this law.

‡ The reader may notice the close relation between this situation and the 'zag' motion described by Gold.[9,41] Likewise, the converging waves meeting the falling stone are 'zaglike', while the other two are 'ziglike'.

particles near the stone at the bottom of the pond). Thus, in these situations we have no need to invoke any extra hypothesis to explain why the ripples are retarded. The entropy hypothesis is already sufficient to rule out the two advanced situations as overwhelmingly improbable – and it also rules out the above unreasonable retarded situation with the ejected stone (assuming the absence of any other agency responsible for ejecting the stone, such as an underwater swimmer, etc.).

The situation with electromagnetic radiation is, for the most part, similar to that of the ripples. There is a minor difference here, however, in the case of stars shining in a largely empty universe: it might well be that some of this radiation is never absorbed by any matter, but continues on indefinitely as the universe expands, or else terminates its existence on a spacetime singularity; likewise, there might be source-free radiation present in the universe, which had been produced directly in the big bang (or in a white hole) or possibly had come in from infinity from a previously collapsing phase of the universe.

I do not believe that these possibilities really make any essential difference to the discussion. I mention them mainly because a great deal has been written on the subject of 'the absorber theory of radiation'. In this view,[42,43] the contribution to the electromagnetic field due to each charged particle is taken to be half advanced and half retarded, while any additional source-free or sink-free radiation is regarded as 'undesirable'. By postulating the absence of such additional radiation, a link between the expansion of the universe and the retardation of radiation is obtained – though, in my opinion, not very convincingly. (And I have to confess to being rather out of sympathy with the whole programme, which strikes me as being unfairly biased against the poor photon, not allowing it the degrees of freedom admitted to all massive particles!)

In any case, the relevance of the entropy argument to the question of retardation seems to me to be quite independent of this.[44] The presence of free radiation coming in from infinity (or from the big bang singularity, say) which converges on a searchlight the moment it is switched on – or some other such absurdity – corresponds just as much to entropy-decreasing-type correlations in the initial state as would radiation coming in from sources. The only difference is that the correlations are just put directly into the photons themselves rather than into the particles producing the photons. So such correlations would be expected to be absent if the entropy in the future is to be high. Correspondingly, there is no objection to such correlations (in time-reversed form) being present in the *future* boundary conditions, because the entropy was low in the

past – and this is, of course, necessary in order that the stars should shine.

The reader may be concerned about how one actually specifies boundary conditions at infinity, or on a spacetime singularity, in order to discuss such correlations in any detail. Of course serious technical problems can arise, particularly in the case of singularities. But the details of these problems should not substantially affect the foregoing discussion – at least if cosmic censorship holds true. I prefer to postpone these questions until section 12.3.2 except just to mention that under certain circumstances (e.g. in a big bang model in which the total charge within some observer's past light-cone is nonzero) there is necessarily a certain amount of source-free radiation present (and in other circumstances, a certain amount of sink-free radiation).[45] There is no reason to believe that this radiation should be correlated in any way which is incompatible with the entropy arguments. The stars still shine 'outwards', rather than 'inwards', whether or not there is some additional radiation permeating space – provided that this radiation has less than the intensity of a star and that it is not specially correlated.

12.2.5 Psychological time

The arrow most difficult to comprehend is, ironically, that which is most immediate to our experiences, namely the feeling of relentless forward temporal progression, according to which potentialities seem to be transformed into actualities. But since the advent of special relativity it has become clear that at least in *some* respects this feeling is illusory. One has the instinctive (or perhaps learnt) impression that one's own concept of temporal progression is universal, so that the transformation of potentialities into actualities that each one of us feels to be taking place ought to occur simultaneously for all of us. Special relativity tells us that this view of the world is false (and it is this lesson that had probably represented the major obstacle to the understanding and acceptance of the theory). Two people amble past one another in the street. What potential events are then becoming actual events on some planet in the Andromeda galaxy? According to the two people, there could be a discrepancy of several days!† (And adopting a view that events are becoming actual on, say, one's past or future light-cone – rather than

† To make the question seem more relevant, imagine that 'at that very moment' a committee is sitting on the planet, deciding the future of humanity!

using Einstein's definition of simultaneity – only makes the subjectivity of these occurrences even worse!)

So relativity *seems* to lead to a picture in which 'potentialities becoming actualities' is either highly subjective or meaningless. Nevertheless the feeling remains very strong within us that there is a very fundamental difference between the past and the future, namely that the past is 'actual' and unchangeable, whereas the future can yet be influenced and is somehow not really fixed. The usual view of the world according to relativity denies this, of course, presenting a rigid four-dimensional determinate picture and telling us that our instinctive feelings concerning the changeability of the future are illusory.

But I do not think that we should just dismiss such feelings out of hand. It is possible to envisage model universes in which the future is yet indeterminate, while the past is fixed. Imagine a continually branching universe, like that depicted in figure 12.2. One is to depict oneself located

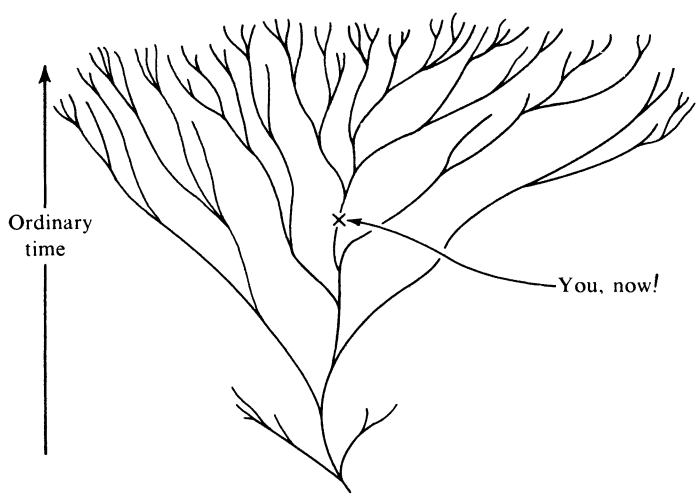

Figure 12.2. A model universe, branching into the future.

at some universe-point in the midst of it all, 'moving' up the picture as one's psychological time unfolds. In this model, the branching takes place only into the future. The path into the past from that point (i.e. the past history of the universe) is absolutely unique, whereas there are many alternative branches into the future (i.e. many alternative possible future histories for the universe, given the present state).

There are, in fact, (at least) two ways to make such a model relativistic. In the first (figure 12.3(a)), the branching takes place along the future

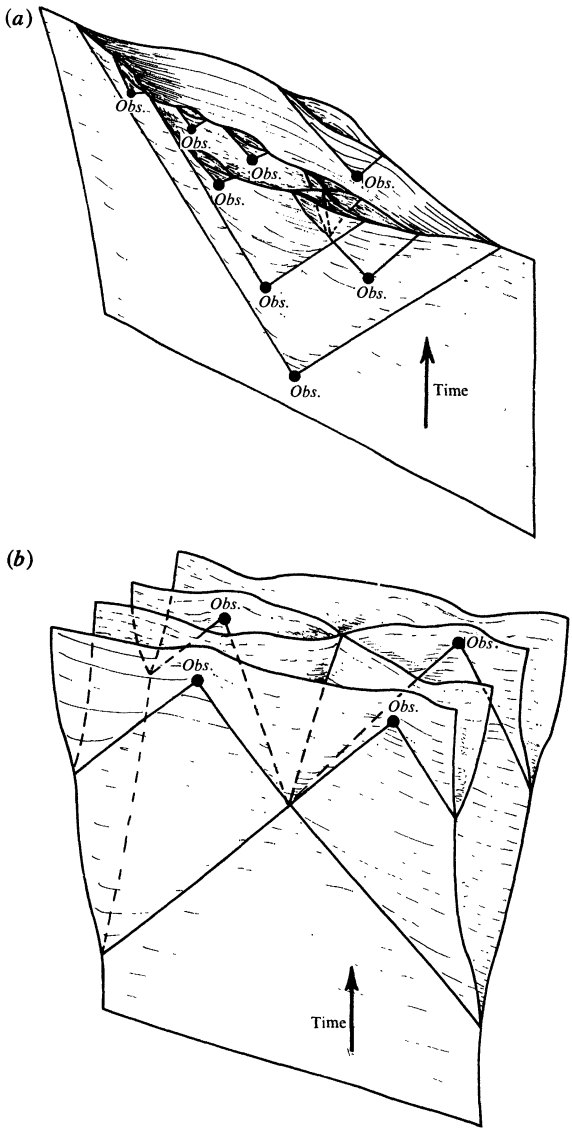

Figure 12.3. Two possible versions of a relativistic future-branching universe: *Obs.* means event at which an 'observation' is made.

light-cones of the points at which 'observations' (presumably quantum-mechanical) are made, and in the second (figure 12.3(*b*)), along the past light-cones of such points. This second case would seem to have considerably less plausibility than the first, since the universe has to 'know' that it has to branch in advance of the observation – which suggests

the presence of some sort of precise correlations in the past, like those leading to advanced-type radiation. The first model, however, is not altogether implausible. It is possible to envisage, for example, that the branching accompanies a kind of retarded collapse of the wavefunction, where on each branch the wavefunction starts out as a different eigenvector of the operator representing the observation.

Such a model may be referred to as an 'Everett-type' universe,[26–29] although it is by no means clear to me that this is really the kind of picture that the Everett formalism of quantum mechanics leads to. For, in the first place, if the Everett picture is to be essentially a reformulation of standard quantum mechanics, it ought to be time-reversible, i.e. there ought to be as much branching into the past as into the future. In the second place, in the Everett picture, one envisages a single wavefunction for the whole universe, which never itself 'collapses'. Instead, it becomes naturally represented as a linear combination of states, in each of which is a measuring apparatus and a physical system appearing to be in a separate eigenstate of the apparatus. Possibly, and with further assumptions as to, say, the nature of the Hamiltonian involved, etc., an Everett picture could, to a certain degree of approximation, be shown to lead to a picture somewhat resembling figures 12.2 and 12.3. But no-one appears to have done this. In particular, without some grossly time-asymmetrical additional assumption there would be nothing to rule out the time-reverses of figures 12.2 and 12.3, or innumerable possibilities in which branching occurs both to the past and to the future, with many branches temporarily separating and then coming together again.

Such considerations are clearly wildly speculative. But one is, in any case, groping at matters that are barely understood at all from the point of view of physics – particularly where the question of human (or non-human) 'consciousness' is implicitly involved, this being bound up with the whole question of psychological time, though not necessarily with the Everett picture. But spacetime models resembling figure 12.3(*a*), say, might well be worthy of study. They could be legitimate mathematical objects, e.g. four-dimensional Lorentzian manifolds subject to Einstein's field equations (say), but where the Hausdorff condition is dropped.[46] It could be argued that such a model is more in accordance with one's intuitive feelings of a determinate past and an indeterminate future than is our normal picture of a Hausdorff space-time. And one could claim that our feeling that 'time moves forwards into the future', rather than 'into the past', is natural in view of the fact that in the 'forward' direction potentialities become actualities, rather than the other way about! Such a

model could be viewed as an 'objective' description of a world containing some strongly 'subjective' elements. One could envisage different conscious observers threading different routes through the myriads of branches (either by chance, say, or perhaps even by the exercise of some 'free will'). And each such observer would have a different 'subjective' view of the world.

This is not to say that I have any strong inclinations to believe in such a picture. I feel particularly uncomfortable about my friends having all (presumably) disappeared down different branches of the universe, leaving me with nothing but unconscious zombies to talk to! I have, in any case, strayed far too long from my avowed conventionality in this discussion, and no new insights as to the *origin* of time-asymmetry have, in any case, been obtained. I must therefore return firmly to sanity by repeating to myself three times: 'spacetime is a *Hausdorff* differentiable manifold; spacetime is a *Hausdorff*...'!

I have as yet made no real attempt to relate the *direction* of psychological time to the key question of entropy. It is clear that any attempted answer must necessarily be very incomplete, in view of our very rudimentary understanding of what constitutes psychological time. And I should emphasize that it is not just a question of the past being (apparently) more certainly *knowable* than the future. Indeed, it is not always the case that the past is easier to ascertain than the future, i.e. that 'retrodiction' is more certain than 'prediction'. At a meeting in Cornell in 1963, Gold[41] pointed out that it was far easier to predict the future behaviour of a recently launched Soviet satellite than to ascertain the time and place of launching! Furthermore, it is not very clear why the phenomenon of increasing entropy should, in any case, be compatible with ease in retrodiction and difficulty in prediction. If earlier states are more 'exceptional' and later ones more 'usual', then it would seem that we should be on safer ground predicting later ones than retrodicting earlier ones! Since, in practice, retrodiction is normally easier (or, at least, more accurate) than prediction (memory being more reliable than soothsaying), the precise relation between this and the entropy question is obscure.

But the issue is really rather different from this. It is not the ease in inferring the past that is relevant here, but the feeling that the past is unchangeable. Likewise, it is not the difficulty that we might have in guessing (or trying to calculate) the future that concerns us, but the feeling that we can affect the future. Thus, despite Gold's observation that the Soviet satellite's future was accurately predictable, one might have the

lingering unease that (had the technical expertise been available at that time) someone might have tried to intercept it. On the other hand, it seems totally inconceivable that any action could be applied *after* the satellite was in orbit to change the launching date! But if the future is to be, in principle, not essentially different from the past, one must entertain the awesome possibility[47] that even the 'fixed' past might conceivably, under suitable circumstances, be changeable! Is it just a question of 'money' that prevents this? (The science-fiction possibility suggests itself of a powerful government going one stage further than falsifying history: namely, actually *changing* the past!) I prefer to leave this question well alone.

There is a rather more helpful way of looking at the intuitive past–future distinction, however. One tends to view events in the past as providing 'causes' or 'reasons' for events in the future, not the other way around. This, at least, *is* compatible with the views I have been presenting in sections 12.2.3 and 12.2.4. My attitude has been that a low entropy at one time may be regarded as providing a 'reason' for precise correlations in particle motions at another, but that the presence or absence of such correlations should not be regarded as providing 'reasons' for anything else. We observe low entropy in the past and infer precise correlations in the future. Thus it is the presence of low entropy in the past that implies that the state of the past provides 'reasons' for the state of the future, not the other way around. This seems to me to be wholly sensible. If it had been the case in our world that collections of broken cogs and springs would sometimes spontaneously assemble themselves into working watches, then people would surely not be averse to attributing the 'causes' of such occurrences to events in the future. Such occurrences might co-exist with others of the more familiar kind, whose 'causes' could be attributed to events in the past. But our universe happens not to be quite like that! The 'causes' of things in *both* types of universe would be traced to situations of low entropy. And in *our* universe these low-entropy states turn out to be in the past.

So at least in this case our psychological feeling of a distinction between past and future can be directly linked to the entropy question. Perhaps the other aspects can too.

12.2.6 Expansion of the universe

I have implicitly indicated in section 12.2.4 that the expansion of the universe should not be regarded as directly responsible for the retar-

dation of radiation, the latter phenomenon being simply one of the many consequences of an assumption that the initial state of the universe was of a far lower entropy than its final state will be (and, correspondingly, that entropy-decreasing correlations were absent in the initial state). I now wish to argue that the expansion of the universe cannot, in itself, be responsible for this entropy inbalance either.

For let us suppose that the contrary is the case and that, for some reason, increasing entropy is a necessary concomitant of an expanding universe. By time-reversal symmetry, this view would entail, correspondingly, that in a contracting universe the entropy should decrease.[48] There are two main situations to consider. First, it might be that the expansion of our actual universe will some day reverse and become a contraction, in which case, according to this view, the entropy would start decreasing again to attain a final low value. The second possibility is that the expansion will continue indefinitely and that the entropy will likewise continue increasing for ever – until a maximum entropy state is reached (ignoring the question of Poincaré cycles, etc.).

It seems that there are very serious objections to the idea that the trend of increasing entropy will reverse itself when the universe reaches maximum expansion. It is hard to see how such reversal could take place without some sort of thermal equilibrium state having been reached in the middle. Otherwise one would have to envisage, it seems to me, a middle state in which phenomena of the normal sort (e.g. retarded radiation and shattering watches) would co-exist with phenomena of the 'time-reversed' sort (e.g. advanced radiation and self-assembling watches). While it is possible to contemplate such situations 'for the purposes of argument', it is a different matter altogether for us to take them seriously for our *actual* universe. Furthermore, the moment of time-symmetry would be reached, it would seem, whilst normal retarded light from very distant galaxies is still coming in (those distant galaxies appearing still to be receding) and, at the same time, there would be advanced light behaving in the time-reversed way (i.e. specially correlated and converging on approaching galaxies). There would seem to be some serious self-consistency problems here[9] (though I am not claiming that they are totally insurmountable). I cannot find it in myself to take such a picture seriously – though some others have apparently not found their intuitions to be so constrained![48]

We might suppose, on the other hand, that the timescale for the reversal of the expansion is so enormously long that an effective thermal equilibrium can be achieved at maximum expansion. But such times that

one must contemplate for this are of a completely different order from the normal cosmological scales. *In effect*, then, such a universe does not recontract at all, and the situation can be considered alongside that of the indefinitely expanding universe-models.

One might think that these models would avoid the problems just considered, but this is not so. Let us envisage an astronaut in such a universe who falls into a black hole. For definiteness, suppose that it is a hole of $10^{10} M_{\odot}$ so that our astronaut will have something like a day inside, for most of which time he will encounter no appreciable tidal forces and during which he could conduct experiments in a leisurely way. In figure 12.4 the situation is depicted in a standard conformal diagram

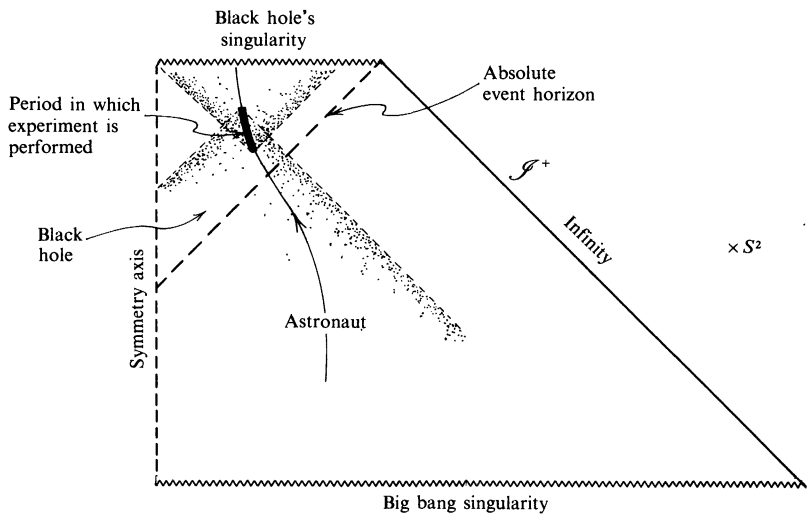

Figure 12.4. Conformal diagram of astronaut falling into black hole in Friedmann $k = -1$ universe.

(light-cones drawn at 45° and spherical symmetry assumed). Notice that the entire history of the astronaut beyond his point of crossing the absolute event horizon lies within the past domain of dependence of the black hole's singularity – and also within the future domain of dependence of the big bang singularity. Suppose that experiments are performed by the astronaut for a period while he is inside the hole. The behaviour of his apparatus (indeed, of the metabolic processes within his own body) is entirely determined by the conditions at the black hole's singularity (assuming that behaviour is governed by the usual hyperbolic-type differential equations) – as, equally, it is entirely determined by the

conditions at the big bang. The situation inside the black hole differs in no essential respect from that at the late stages of a recollapsing universe. If one's viewpoint is to link the local direction of time's arrow directly to the expansion of the universe, then one must surely be driven to expect that our astronaut's experiments will behave in an entropy-*decreasing* way (with respect to 'normal' time). Indeed, one should presumably be driven to expect that the astronaut would believe himself to be coming *out* of the hole rather than falling in (assuming his metabolic processes could operate consistently through such a drastic reversal of the normal progression of entropy).

I have to say that I cannot help regarding this possibility as even *more* of an absurdity than the entropy reversal at maximum expansion of a recollapsing universe! One would presumably not expect the entropy to reverse suddenly as the astronaut crosses the horizon (and thereafter, the collapsing aspect of his boundary conditions *enormously* outweighs the expanding one). So, strange behaviour in the entropy would have to occur well before the astronaut actually crosses the horizon – whereupon the astronaut could change his mind and accelerate outwards, so avoiding capture by the hole and being in a position to report his strange findings to the outside world! Indeed, the black hole argument can also be applied in the recollapsing universe. We do not need to wait for the whole universe to recollapse in order for the absurdities of this viewpoint to manifest themselves. (It may be that some holders of this viewpoint have a disinclination to accept the reality of black holes in any case. I have no desire to enter into the arguments – in my view very compelling – in favour of black holes here, but refer the reader to the literature.[7,51])

An argument could be put forward that the spacetime depicted in figure 12.4 is based on a too strong and unrealistic assumption of spherical symmetry. In fact this is not really the case; the dropping of spherical symmetry should make no essential qualitative difference to the picture, provided only that a (suitably strong) assumption of cosmic censorship is made. I prefer to postpone a more detailed discussion of this point until section 12.3.2. For the moment it is sufficient that I may fall back on conventionality again to conclude that the expansion of the universe is *not, in itself*, somehow responsible for the fact that the entropy of our universe is increasing.

This is not to say that I regard the correspondence between these two awesome facts as entirely fortuitous. Far from it. For I shall argue later on that both are consequences of the very special nature of the big bang – a special nature that is *not* to be expected in the singularities of recollapse.

12.2.7 Black holes versus white holes

General relativity is a time-symmetric theory. So, to any solution of its equations (with time-symmetric equations of state) that is asymmetric in time, there must correspond another solution for which the time-ordering is reversed.† One of the most familiar solutions is that representing (spherically symmetric) collapse of a star (described using, say, the T_{ab} of 'dust') to become a black hole.[54,55,7] In time-reversed form this represents what is referred to as a 'white hole' finally exploding into a cloud of matter. Spacetime diagrams for the two situations are given in figure 12.5.

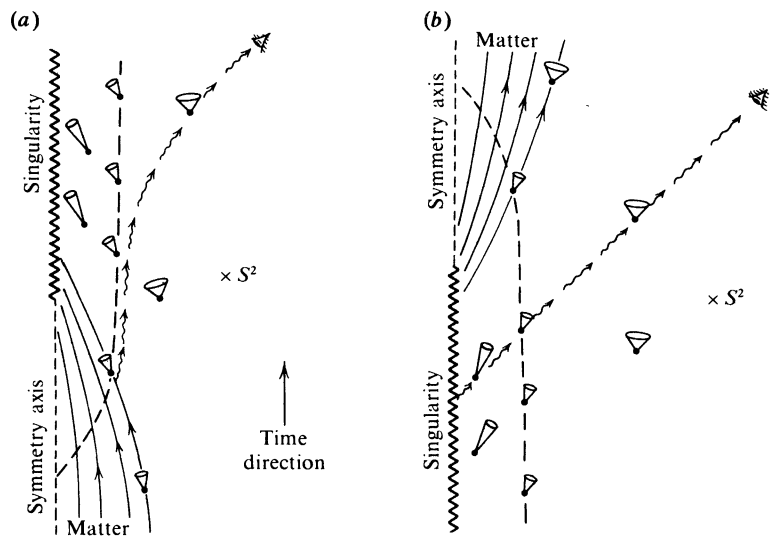

Figure 12.5. Black and white holes (Finkelstein-type picture): (*a*) collapse to a black hole, (*b*) explosion from a white hole. An observer's eye is at the top right.

Various authors[52,53] have attempted to invoke white holes as explanations for quasars or other violent astronomical phenomena (sometimes using the name of 'lagging cores'). Recall that in a (classical) collapse to a black hole, the situation starts out with a perfectly normal distribution of matter which follows deterministic evolutionary equations. At a certain stage a trapped surface may form, leading to the presence of an *absolute event horizon* into which particles can fall, but out of which none can escape. After all the available matter has been swallowed, the hole settles

† For this purpose, a 'solution of the Einstein equations' would be a Lorentzian 4-manifold with a *time-orientation* (and perhaps a space-orientation). 'Reversing the time-ordering' amounts to selecting the opposite time-orientation.

down and remains unchanging until the end of time (possibly at recollapse of the universe). (This ignores the quantum-mechanical effects of the Hawking process,[56] which I shall discuss in a moment.) A white hole, therefore, is created at the beginning of time (i.e. in the big bang) and remains in an essentially unchanging state for an indefinite period. Then it disappears by exploding into a cloud of ordinary matter. During the long quiescent period, the boundary of the white hole is a stationary horizon – the *absolute particle horizon* – into which no particle can fall, but out through which particles may, in principle, be ejected.

There is something that seems rather 'thermodynamically unsatisfactory' (or physically improbable) about this supposed behaviour of a white hole, though it is difficult to pin down what seems wrong in a definitive way. The normal picture of collapse to a black hole seems to be 'satisfactory' as regards one's conventional ideas of classical determinism. Assuming that (strong) cosmic censorship[57–59] holds true, the entire spacetime is determined to the *future* of some 'reasonable' Cauchy hypersurface, on which curvatures are small. But in the case of the white hole, there is no way of specifying such boundary conditions in the past because an initial Cauchy hypersurface has to encounter (or get very close to) the singularity. Put another way, the future behaviour of such a white hole does not, in any *sensible* way, seem to be determined by its past. In particular, the precise moment at which the white hole explodes into ordinary matter seems to be entirely of its own 'choosing', being unpredictable by the use of normal physical laws. Of course one could use future boundary conditions to retrodict the white hole's behaviour, but this (in our entropy-increasing universe) is the thermodynamically unnatural way around. (And, in any case, one can resort to memory as a more effective means of retrodiction!)

Related to this is the fact that while an external observer (using normal retarded light) cannot directly see the singularity in the case of a black hole, he can do so in the case of a white hole (figure 12.5). Since a spacetime singularity is supposed to be a place where the known physical laws break down it is perhaps not surprising, then, that this implies a strong element of indeterminism for the white hole. Causal effects of the singularity can, in this case, influence the outside world.

The presently accepted picture of the physical effects that are expected to accompany a spacetime singularity, is that *particle creation* should take place.[62,63,66] This is the general conclusion of various different investigations into curved-space quantum field theory. However, owing to the incomplete state of this theory, these investigations do not always agree

on the details of their conclusions. As applied to a white hole, two particular schools of thought have arisen. According to Zel'dovich[64] the white hole ought to be completely unstable to this process, evaporating away instantaneously, while Hawking[67] has put forward the ingenious viewpoint that the white hole evaporation ought to be much lower, and indistinguishable in nature from that produced, according to the Hawking process, by a black hole of the same mass – indeed, that a white hole ought itself to be indistinguishable from a black hole! This Hawking viewpoint is, in a number of respects, a very radical one which carries with it some serious difficulties. I shall consider these in a moment. The other viewpoint has the implication that white holes should not physically exist (though, owing to the tentative nature of the particle-creation calculations, this may not carry a great deal of weight; however, cf. also Eardley[65]).

There is another reason for thinking of white holes as antithermodynamic objects (though this reason, too, must be modified if one attempts to adopt the above-mentioned more radical of Hawking's viewpoints). According to the Bekenstein–Hawking formula,[56,69] the surface area A of a *black* hole's absolute event horizon is proportional to the intrinsic entropy S of the hole:

$$S = kAc^3/4\hbar G$$

(k being Boltzmann's constant and G being Newton's gravitational constant). The area principle of classical general relativity[7,70,71] tells us that A is non-decreasing with time in classical processes, and this is compatible with the thermodynamic time's arrow that entropy should be non-decreasing. Now, if a white hole is likewise to be attributed an intrinsic entropy, it is hard to see how the value of this entropy can be other than that given by the Bekenstein–Hawking formula again, but where A now refers to the absolute *particle* horizon. The time-reverse of the area principle then tells us that A is *non-increasing* in classical processes, which is the opposite of the normal thermodynamic time's arrow for entropy. In particular, the value of A will substantially *decrease* whenever the white hole ejects a substantial amount of matter, such as in the final explosion shown in figure 12.5. Thus, this is a strongly anti-thermodynamic behaviour.

It seems that there are two main possibilities to be considered concerning the physics of white holes. One of these is that there is a general principle that rules out their existence (or, at least, that renders them overwhelmingly improbable). The other possibility is contained in the

aforementioned line of argument due to Hawking, which suggests that because of quantum-mechanical effects, black and white holes are to be regarded as physically indistinguishable.[67] I shall discuss, first, Hawking's remarkable idea. Then I shall attempt to indicate why I nevertheless believe that this cannot be the true explanation, and that it is necessary that white holes do *not* physically exist.

Recall, first, the Hawking radiation that is calculated to accompany any black hole. The temperature of the radiation is inversely proportional to the mass of the hole, being of the general order of 10^{-7} K in the case of a black hole of 1 $M_{\odot}$. Of course this temperature is utterly insignificant for stellar-mass holes, but it could be relevant observationally for very tiny holes, if such exist. In an otherwise empty universe, the Hawking radiation would cause the black hole to lose mass, become hotter, radiate more, lose more mass, etc., the whole process accelerating until the hole disappears (presumably) in a final explosion. But for black hole of solar mass (or more) the process would take $>10^{53}$ Hubble times! And, so that the process could even begin, a wait of 10^{7}, or so, Hubble times would be needed to enable the expansion of the universe to reduce the present background radiation to below that of the hole – assuming an indefinitely expanding universe-model!

The absurdity of such figures notwithstanding, it is of some considerable theoretical interest to contemplate, as Hawking has done,[67,72] the state of thermal equilibrium that would be achieved by a black hole in a large container with perfectly reflecting walls. If the container is sufficiently large for a given total mass–energy content (case (*a*)), the black hole will radiate itself away completely (presumably) – after having swallowed whatever other stray matter there had been in the container – to leave, finally, nothing but thermal radiation (with perhaps a few thermalized particles). This final state will be the 'thermal equilibrium' state of maximum entropy (see figure 12.6(*a*)).

If the container is substantially smaller (case (*c*)) – or, alternatively, if the mass–energy content is substantially larger (though still not large enough to collapse the whole container) – the maximum-entropy state will be achieved by a single spherical black hole in thermal equilibrium with its surrounding radiation. Stability is here achieved because, if by a fluctuation the hole radiates a bit too much and consequently heats up, its surroundings heat up even more and cause it to absorb more than it emits and thus to return to its original size; if by a fluctuation it radiates less than it absorbs, its surroundings cool by more than it does and again it returns to equilibrium (figure 12.6(*c*)).

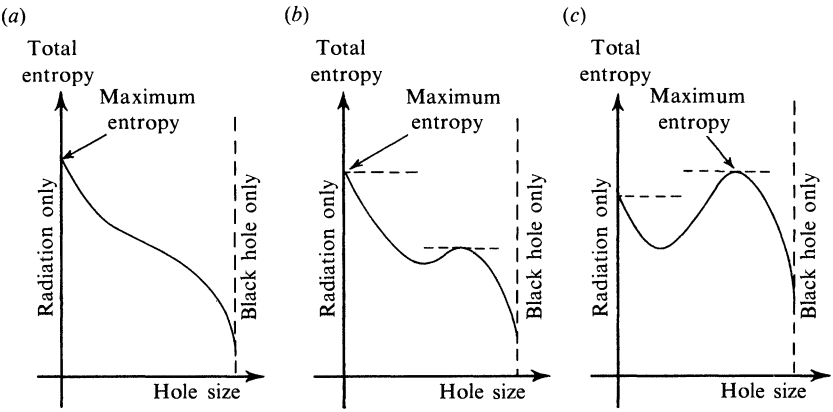

Figure 12.6. Hawking's black hole in a perfectly reflecting container: (*a*) large container, (*b*) intermediate container, (*c*) smallish container.

There is also a situation (case (*b*)) in which the container lies in an intermediate size range, for the given mass–energy content, and where the black hole is still stable, but only represents a *local* entropy maximum, the absolute maximum being a state in which there is only thermal radiation (and perhaps thermalized particles) but no black hole. In this case the black hole can remain in equilibrium with its surrounding radiation for a very long period of time. A large fluctuation would be needed in which a considerable amount of radiation is emitted by the hole, sufficient to get across the low entropy barrier between the two local maxima (figure 12.6(*b*)). With such a large mass loss from the hole, it is able to heat up by an amount greater than its surroundings are able to do; it then loses more mass, heats up more, and, as in case (*a*), radiates itself away completely, to give the required state of thermal radiation (plus occasional thermalized particles).

It should be pointed out that whereas we are dealing with processes that have absurdly long timescales,† these situations have a rather fundamental significance for physics. We are concerned, in fact, with the states of maximum entropy for *all* physical processes. In cases (*a*) and (*b*), the maximum-entropy state is the familiar 'heat death of the universe', but in case (*c*) we have something new: a black hole in thermal equilibrium with radiation. There are, of course, many detailed theoretical difficulties with this setup (e.g. the Brownian motion of the black hole would occasionally send it up against a wall of the container, whereupon

† If we allow (virtual) black holes of down to 10^{-33} cm (i.e. $\sim 10^{-20}$ of an elementary particle radius) then we obtain a picture[73] for which these timescales can be very short.

the container should be destroyed). Such problems will be ignored as irrelevant to the main issues! But also, since the relaxation times are so much greater than the present age of the universe, the interpretations of one's conclusions do need some care. Nevertheless, I feel that there are important insights to be gained here.

To proceed further with Hawking's argument, consider case (*c*). For most of the time the situation remains close to maximum entropy: a black hole with radiation. But occasionally, via an initial large fluctuation in which a considerable energy is emitted by the hole, a sequence like that just considered for case (*b*) will occur, where the black hole evaporates away to give thermal radiation. But then, after a further long wait, enough radiation (again by a fluctuation) collects together in a sufficiently small region for a black hole to form. Provided this hole is large enough, the system settles back into the maximum-entropy state again, where it remains for a very long while.

Cycles like this can also occur in case (*b*) (and even in case (*a*)), but with the difference that most of the time is spent in a state where there is no black hole. In case (*c*), a black hole is present most of the time. Hawking now argues that since the essential physical theories involved are time-symmetric (general relativity, Maxwell theory, neutrinos, possibly electrons, pions, etc., and the general framework of quantum mechanics), the equilibrium states ought to be time-symmetric also. But reversing the time-sense leads to white holes, not to black holes. Thus, Hawking proposes, white holes ought to be physically indistinguishable from black holes!

This identification is not so absurd as one might think at first. The Hawking radiation from the black hole becomes reinterpreted as particle creation near the singularity of the white hole (and hence Hawking proposes a rather slow rate of particle production at the white hole singularity). The swallowing of radiation by the black hole becomes time-reversed Hawking radiation from the white hole. One can, of course, envisage a black hole swallowing a complicated object such as a television set. How can this be thought of as time-reversed Hawking radiation? The argument is that Hawking radiation, being thermal,[67,74] produces all possible configurations with equal probability. It is *possible* to produce a television set as part of the Hawking radiation of a black hole, but such an occurrence is overwhelmingly improbable and would correspond to a large reduction in entropy. A black hole swallowing a television set only seems more 'natural' because we are used to situations in which the entropy is low in the initial state. We can equally well

envisage initial boundary conditions of low entropy for the time-reversed Hawking radiation – and this would be the case for a television set being annihilated as time-reversed Hawking radiation of a white hole.

So far, this all seems quite plausible, and there is even a certain unexpected elegance and economy in the whole scheme. But unfortunately is suffers from two (or perhaps three) very severe drawbacks which, in my opinion, rule it out as a serious possibility.

In the first place, whereas the geometry of the spacetime outside a stationary black hole's horizon is identical to that outside a stationary white hole's horizon, it is definitely *not* so that the exterior geometry of a black hole that forms by standard gravitational collapse and then finally disappears according to the Hawking process is time-symmetric. This time-asymmetry is made particularly apparent by use of conformal diagrams as shown in figure 12.7. A precise distinction between the

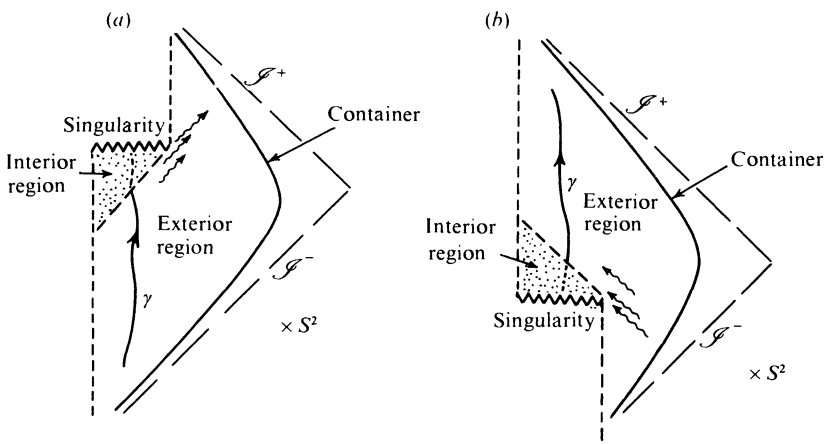

Figure 12.7. Conformal diagrams illustrating time-asymmetry of a transient black hole. (*a*) Classical collapse to a black hole followed by complete Hawking evaporation. (*b*) Hawking condensation to a white hole followed by its classical disappearance.

transient black hole and white hole external geometries can be made in terms of their TIP and TIF structures (cf. section 12.3.2). But intuitively, the distinction should be clear from the presence of timelike curves γ which, in the case of the black hole, 'leave' the exterior geometry to 'enter' the hole; and the other way around in the case of the white hole. The reason for this distinction is simply that the process of *classical* collapse is not the time-reverse of the *quantum* Hawking process. This should not really surprise us since each relies on quite different physical

theories (classical general relativity as opposed to quantum field theory on a fixed curved-space background).

A point of view adopted by Hawking which might avoid this difficulty is to regard the spacetime geometry as being somewhat observer-dependent. Thus, as soon as quantum mechanics and curved-space geometry have become essentially intertwined, so this viewpoint would maintain, one cannot consistently talk about a classically objective spacetime manifold. An observer who falls into a white hole to be evaporated away as its time-reversed Hawking radiation would, accordingly, believe the geometry to be, instead, that of a black hole whose horizon he crosses and inside which he awaits his 'classical' fate of final destruction by excessive tidal forces.

I have to say that I find this picture almost as hard to accept as those according to which the entropy starts decreasing when the observer's universe collapses about him. If one is considering black or white holes whose radius is of the order of the Planck length ($\sim 10^{-33}$ cm) – or even, possibly, of the order of an elementary particle size ($\sim 10^{-13}$ cm) – then such indeterminancy in the geometry might be acceptable. But for a black hole of solar mass (or more) this would entail a very radical change in our views about geometry, a change which would drastically affect almost any other application of general relativity to astrophysical phenomena. It is true that in section 12.2.5 I have briefly entertained the possibility of a world-view which allows for an element of 'observer-dependence' in the geometry. But I have as yet seen no way to relate such a view to the kind of indeterminacy in classical geometry that the physical identification of black holes with white holes seems to lead to.

But there are also other objections to attempting to regard classical gravitational collapse as being effectively the time-reverse of a quantum-mechanical particle creation process. One of these refers not so much to the attempted identification of the Hawking process with the time-reverse of the classical swallowing of matter by a black hole, but with the attempt to identify either of these processes with the phenomenon of particle production at regions of large spacetime curvature. Such a further identification *seems* to be an integral part of the time-symmetric view that I have been discussing, though it may well not really be what is intended. I am referring to the picture of Hawking radiation by a black hole as being alternatively regarded as a process of particle production near the singularity of a white hole. If, indeed, it can be so regarded, then this is not the 'normal' process of particle production at regions of large curvature that has many times been discussed in the literature.[62] For in

that process, particles are always produced in *pairs*: baryon with antibaryon, lepton with antilepton; positively charged particle with negatively charged particle. But the Hawking process is explicitly not of this form, as its thermal nature (for particles escaping to infinity) implies.[67,74,75]

The contrast is even more blatant if we try to relate this pairwise particle production near white hole singularities to the time-reverse of the destruction of matter at a black hole's singularity. For there are no constraints whatever on the type of matter that a black hole can classically absorb. And if strong cosmic censorship is accepted in classical processes, it seems that even individual charged particles have then to be separately destroyed at the singularity. (This point will be amplified in section 12.3.2.) There is no suggestion that the particles must somehow contrive to sort themselves into particle–antiparticle pairs before they encounter the singularity. The difficulty here is, perhaps, not so much directly to do with Hawking's identification of black holes with white holes, but with the whole idea of hoping to deal with the matter destruction–creation process in terms of known physics. Thus I maintain that, whereas it *may* be that matter creation at the big bang can be treated in terms of known (or at least partially understood) particle-creation processes, this seems *not* to be true of the destruction processes at black hole singularities – nor, if Hawking is right, of the creation processes at white hole singularities. Thus, the Hawking view would seem to lead to direct conflict with the often expressed hope that particle creation at the big bang can be understood in terms of processes of particle production by spacetime curvature. This relates to the question of whether time-symmetric physics can be maintained at spacetime singularities. This is a key issue that I shall discuss in more detail in section 12.3.

More directly related to Hawking's proposal is a difficulty which arises if we examine in detail the cycles whereby a solar-mass black hole (say), in stable equilibrium with its surroundings, in our perfectly reflecting container, may disappear and reform, owing to fluctuations, in cases (*b*) and (*c*) just discussed. What, in fact, is the most probable way for the black hole to evaporate completely? It might, of course, simply throw out its entire mass in one gigantic fluctuation. But the random Hawking process would achieve this only absurdly infrequently. Overwhelmingly *less* absurdly infrequent would be the emission, in one huge fluctuation, of that *fraction* of the hole's mass needed to raise its Bekenstein–Hawking temperature above that to which its surroundings would be consequently raised. From there on, the evaporation would proceed

'normally' needing no further improbable occurrence. But consider the final explosion according to which the hole 'normally' disappears. Electrons and positrons have been appearing, followed by pions; then, at the last moment, a whole host of unstable particles is produced which undergo complicated decays. Finally one expects many protons and antiprotons separately escaping the annihilation point. Only much later, by chance encounters as they move randomly about the container, would one expect the protons and antiprotons gradually to annihilate one another (or possibly occasionally to decay, themselves, by a Pati–Salam-type process[76] into, say, positrons and into electrons which would then mostly annihilate one another).

What, now, must we regard as the most probable way in which a black hole forms again in the container, to reach another point of stable equilibrium with its surroundings? Surely it is *not* the time-reverse of the above, according to which the protons and antiprotons must first (with long preparation) contrive to form themselves (quite unnecessarily) out of the background radiation before aiming themselves with immense accuracy at a tiny point, only to indulge in some (again unnecessary) highly contrived particle physics whereby they meet up with carefully aimed γ-rays, etc., to form various unstable particles, etc., etc.; and then (also with choreography of the utmost precision) other particles (pions, then electrons and positrons) must aim themselves inwards, having first formed themselves out of the background at the right moment and in the correct proportions. Only later does the background radiation itself fall inwards to form the bulk of the mass needed to form the hole.

My point is not that this curious beginning is necessarily the most improbable part of the process. I can imagine that it may well not be. But it is unnecessary. The essential part of the hole formation of a *black* hole would occur when the radiation itself collects together in a sufficiently small region to undergo what is, in effect, a standard gravitational collapse. In fact, it would seem that the tiny core that has been formed with such elaborate preparations ought somewhat to *inhibit* the subsequent inward collapse, owing to its excessive temperature!

So what has gone wrong? What time-asymmetric physics has been smuggled into the description of the 'most probable' mode of disappearance of the hole, that it should so disagree with the time-reverse of its 'most probable' mode of reappearance? Possibly none, if white holes, in principle, exist, and are simply *different* objects from black holes. While the above elaborate preparations are not necessary for the formation of a black hole, they could be for a white hole. After all, one has to contrive

some way of producing the white hole's initial singularity which, as is evident from figure 12.7, has a quite different structure from that of a black hole. It might be that the production of such a singularity is an extraordinarily delicate process, requiring particles of just the right kind and in an essentially right order to be aimed with high energy and with extraordinary accuracy. There is an additional seeming difficulty, however, in that one also has to conjure up a region of spacetime (namely that inside the horizon) which does not lie in the domain of dependence of some initial Cauchy hypersurface drawn before the white hole appears (figure 12.7(*b*)). Of course, the time-reverse of this problem also occurs for the (Hawking) disappearance of a black hole, but one is not in the habit of trying to retrodict from Cauchy hypersurfaces, so this seems less worrisome! An additional difficulty, if white holes are allowed, is that we now encounter the problem, mentioned earlier, of trying to predict what the white hole is going to emit, and when. As I have indicated, *if* time-symmetric physics holds at singularities, then 'normal' ideas of particle creation due to curvature will not do. Hawking's concept of 'randomicity'[67] might be nearer the mark, but it is unfortunately too vague to enable any calculations to be made – now that the essential guiding idea of an identification between black holes and white holes has been removed.

I find the picture of an equilibrium involving the occasional production of such white holes a very unpleasant one. And what other monstrous zebroid combinations of black and white might also have to be contemplated? I feel that such things have nothing really to do with physics (at least on the macroscopic scale). The only reason that we have had to consider white holes *at all* is in order to save time-symmetry! The consequent unpleasantness and unpredictability seems a high price to pay for something that is *not even true* of our universe on a large scale.

One of the consequences of the hypothesis that I shall set forth in the next section is that it rules out the white hole's singularity as an unacceptable boundary condition. The hypothesis is time-asymmetric, but this is necessary in order to explain the other arrows of time. When we add this hypothesis to the discussion of equilibrium within the perfectly reflecting container, we see explicitly what time-asymmetric physics has been 'smuggled' in. For the hypothesis is designed not to constrain the behaviour of black holes in any way, but it forbids white holes and therefore renders irrelevant the extraordinary scenario that we seem to need in order to produce one!

I hope the reader will forgive me for having discussed white holes at such length only to end by claiming that they do not exist! But hypothetical situations can often lead to important understandings, especially when they border on the paradoxical, as seems to be the case here.

12.3 Singularities: the key?

What is the upshot of the discussion so far? According to sections 12.2.3, 12.2.4 and 12.2.5, the arrows of entropy and retarded radiation, and *possibly* of psychological time, can all be explained if a reason is found for the initial state of the universe (big bang singularity) to be of comparatively low entropy and for the final state to be of high entropy. According to section 12.2.6, some low-entropy assumption *does* need to be *imposed* on the big bang; that is, the mere fact that the universe expands away from a singularity is in no way sufficient. And according to section 12.2.7, we need some assumption on initial singularities that rules out those which would lie at the centres of white holes. On the other hand, the discussions in sections 12.2.1 and 12.2.2 were inconclusive, and I shall need to return to them briefly at the end.

But what is it in the nature of the big bang that is of 'low entropy'? At first sight, it would seem that the knowledge we have of the big bang points in the opposite direction. The matter (including radiation) in the early stages appears to have been completely thermalized (at least so far as this is possible, compatibly with the expansion). If this had not been so, one would not get correct answers for the helium abundance, etc.[77,78] And it is often remarked upon that the 'entropy per baryon' (i.e. the ratio of photons to baryons) in the universe has the 'high' value of $\sim 10^9$. Ignoring the contribution to the entropy due to black holes, this value has remained roughly constant since the very early stages, and then represents easily the major contribution to the entropy of the universe – despite all the 'interesting' processes going on in the world, so important to our life here on Earth, that depend upon 'small' further taking up of entropy by stars like our Sun. The answer to this apparent paradox – that the big bang thus *seems* to represent a state of *high* entropy – lies in the unusual nature of gravitational entropy. This I next discuss, and then show how this relates to the structure of singularities.

12.3.1 Gravitational entropy

It has been pointed out by many authors[79] that gravity behaves in a somewhat anomalous way with regard to entropy. This is true just as

much for Newtonian theory as for general relativity. (In fact, the situation is rather worse for Newtonian theory.) Thus, in many circumstances in which gravity is involved, a system may behave as though it has a negative specific heat. This is directly true in the case of a black hole emitting Hawking radiation, since the more it emits, the hotter it gets. But even in such familiar situations as a satellite in orbit about the Earth, we observe a phenomenon of this kind. For dissipation (in the form of frictional effects in the atmosphere) will cause the satellite to speed up, rather than slow down, i.e. cause the kinetic energy to increase.

This is essentially an effect of the universally attractive nature of the gravitational interaction. As a gravitating system 'relaxes' more and more, velocities increase and the sources clump together – instead of uniformly spreading throughout space in a more familiar high-entropy arrangement. With other types of force, their attractive aspects tend to saturate (such as with a system bound electromagnetically), but this is not the case with gravity. Only non-gravitational forces can prevent parts of a gravitationally bound system from collapsing further inwards as the system relaxes. Kinetic energy itself can halt collapse only temporarily. In the absence of significant non-gravitational forces, when dissipative effects come further into play, clumping becomes more and more marked as the entropy increases. Finally, maximum entropy is achieved with collapse to a black hole – and this leads us back into the discussion of section 12.2.7.

Consider a universe that expands from a 'big bang' singularity and then recollapses to an all-embracing final singularity. As was argued in section 12.2.6, the entropy in the late stages ought to be much higher than the entropy in the early stages. How does this increase in entropy manifest itself? In what way does the high entropy of the final singularity distinguish it from the big bang, with its comparatively low entropy? We may suppose that, as is apparently the case with the actual universe, the entropy in the initial *matter* is high. The kinetic energy of the big bang, also, is easily sufficient (at least on average) to overcome the attraction due to gravity, and the universe expands. But then, relentlessly, gravity begins to win out. The precise moment at which it does so, locally, depends upon the degree of irregularity already present, and probably on various other unknown factors. Then clumping occurs, resulting in clusters of galaxies, galaxies themselves, globular clusters, ordinary stars, planets, white dwarfs, neutron stars, black holes, etc. The elaborate and interesting structures that we are familiar with all owe their existence to

this clumping, whereby the gravitational potential energy begins to be taken up and the entropy can consequently begin to rise above the *apparently* very high value that the system had initially. This clumping must be expected to increase; more black holes are formed; smallish black holes swallow material and congeal with each other to form bigger ones. This process accelerates in the final stages of recollapse when the average density becomes very large again, and one must expect a very irregular and clumpy final state.

There is a slight technical difficulty in that the concept of a black hole is normally only defined for asymptotically flat (or otherwise open) spacetimes. This difficulty could affect the discussion of the final stages of collapse when black holes begin to congeal with one another, and with the final all-embracing singularity of recollapse. But I am not really concerned with the location of the black holes' event horizons, and it is only in precisely defining these that the aforementioned difficulty arises. A black hole that is formed early in the universe's history has a singularity that is reached at early proper times for observers who encounter it;[57] for holes that are formed later, they can be reached at later proper times. On the basis of strong cosmic censorship (cf. section 12.3.2), one expects all these singularities eventually to link up with the final singularity of recollapse.[57] I do not require that the singularities of black holes be, in any clear-cut way, distinguishable from each other or from the final singularity of recollapse. The point is merely that the gravitational clumping which is characteristic of a state of high gravitational entropy should manifest itself in a very complicated structure for the final singularity (or singularities).

The picture is not altogether dissimilar for a universe that continues to expand indefinitely away from its big bang. We still expect local clumping, and (provided that the initial density is not altogether too low or too uniform for galaxies to form at all) a certain number of black holes should arise. For the regions inside these black holes, the situation is not essentially different from that inside a collapsing universe (as was remarked upon in section 12.2.6), so we expect to find, inside each hole, a very complicated singularity corresponding to a very high gravitational entropy. For those regions not inside black holes there will still be certain localized portions, such as rocks, planets, black dwarfs, or neutron stars, which represent a certain ultimate raising of the entropy level owing to gravitational clumping, but the gain in gravitational entropy will be relatively modest, though sufficient, apparently, for all that we need for life here on Earth.

I have been emphasizing a qualitative relation between gravitational clumping and an entropy increase due to the taking up of gravitational potential energy. In terms of spacetime curvature, the absence of clumping corresponds, very roughly, to the absence of Weyl conformal curvature (since absence of clumping implies spatial-isotropy, and hence no gravitational principal null-directions).[45] When clumping takes place, each clump is surrounded by a region of nonzero Weyl curvature. As the clumping gets more pronounced owing to gravitational contraction, new regions of empty space appear with Weyl curvature of greatly increased magnitude. Finally, when gravitational collapse takes place and a black hole forms, the Weyl curvature in the interior region is larger still and diverges to infinity at the singularity.

At least, that is the picture presented in spherically symmetric collapse, the magnitude of the Weyl curvature diverging as the inverse cube of the distance from the centre. But there are various reasons for believing that in generic collapse, also, the Weyl curvature should diverge to infinity at the singularity, and (at most places near the singularity) should dominate completely over the Ricci curvature.

This can be seen explicitly in the details of the Belinskii–Khalatnikov–Lifshitz analysis.[80] Moreover, one can also infer it on crude qualitative grounds. In the exact Friedmann models, it is true, the Ricci tensor dominates, the Weyl tensor being zero throughout. In these cases, as a matter world-line is followed into the singularity, it is approached isotropically by the neighbouring matter world-lines, so we have simultaneous convergence in three mutually perpendicular directions orthogonal to the world-line. In the case of spherically symmetrical collapse to a black hole, on the other hand, if we envisage some further matter falling symmetrically into the central singularity, it will normally converge in towards a given matter world-line only in *two* mutually perpendicular directions orthogonal to the world-line (and diverge in the third). This is the situation of the Kantowski–Sachs[81] cosmological model, giving a so-called 'cigar'-type singularity.[7] If r is the usual Schwarzschild coordinate, the volume gets reduced like $r^{3/2}$ near the singularity, so the densities are $\sim r^{-3/2}$. Thus, for a typical Ricci tensor component, $\Phi \sim r^{-3/2}$. However, in general, for a typical Weyl tensor component, $\Psi \sim r^{-3}$, showing that the Weyl tensor dominates near the singularity in these situations. Also, in the 'pancake' type of singularity, where there is convergence in only *one* direction orthogonal to a matter world-line, we again expect the Weyl tensor to dominate with $\Phi \sim r^{-1}$ and $\Psi \sim r^{-2}$ in this case.

Now the Friedmann type of situation, with simultaneous convergence of all matter from all directions at once, would seem to be a very special setup. If there is somewhat less convergence in one direction than in the other two, then a cigar-type configuration seems more probable very close to the singularity, while a pancake-type appears to result when the main convergence is only in one direction. Moreover, with a generic setup, a considerable amount of oscillation seems probable.[80] An oscillating Weyl curvature of frequency ν and complex amplitude Ψ, supplies an *effective* additional 'gravitational-energy' contribution to the Ricci tensor[61] of magnitude $\sim|\Psi|^2\nu^{-2}$. If ν becomes very large so that many oscillations occur before the singularity is reached, then[49] $\nu^2 \gg \Phi^{-1}$, where Φ is a typical Ricci tensor component. Thus if, as seems reasonable in general, the 'energy content' of Ψ is to be comparable with Φ as the singularity is approached, we have $|\Psi|^2\nu^{-2} \sim \Phi$, so $|\Psi| \gg \Phi$. These considerations are very rough, it is true, but they seem to concur with more detailed analysis[80] which indicates that in generic behaviour near singularities the contributions due to matter can be ignored to a first approximation and the solution treated as though it were a vacuum, i.e. that the Weyl part of the curvature dominates over the Ricci part.

The indications are, then, that a high-entropy singularity should involve a very large Weyl curvature, unlike the situation of the singularity in the Friedmann dust-filled universe or any other models of the Robertson–Walker class. At the time of writing, however, no clear-cut integral formula (say) which could be regarded as giving mathematical expression to this suggested relation between Weyl curvature and gravitational entropy has come to light. Some clues as to the nature of such a formula (if such exists at all) may be obtained, firstly, from the Bekenstein–Hawking formula for the entropy of a black hole and, secondly, from the expression for the particle number operator for a linear spin-2 massless quantized free field – since an estimate of the 'number of gravitons' in a gravitational field could be taken as a measure of its entropy.† Thus this entropy measures the number of *quantum* states that contribute to a given classical geometry.

There is one final point that should be mentioned in connection with the question of the entropy in the gravitational field. It was pointed out some time ago by Tolman[84] that a model universe containing matter that

† This point of view does not seem to agree with that of Gibbons and Hawking,[82] who apparently regard the gravitational entropy as being zero when black holes are absent. But an estimate of 'photon number' in a classical electromagnetic field gives a measure of its entropy[83] (without black holes). Gravity is presumably similar.

appeared to be in thermal equilibrium in its early stages can lead to a situation in which the matter gets out of equilibrium as the universe expands (a specific example of matter illustrating such behaviour being a diatomic gas which is capable of dissociating into its elements and recombining). Then, if such a model represents an expanding and recollapsing universe, the state of the matter during recollapse would differ from the corresponding state during expansion, where we make the correspondence at equal values of the universe radius R (or comoving radius R). In fact, the matter, during recollapse, would have acquired some energy out of the global geometry of the universe, the resulting difference in geometry showing up in the fact that $\dot{R}^2$ is greater, for given R, at recollapse than it is during the expansion. So the entropy of the system as a *whole* increases with time even though the matter *itself* is in thermal equilibrium during an initial stage of the expansion. There is, in fact, a contribution to the entropy from R (and $\dot{R}$), which must be regarded as a dynamical variable in the model. (This arises because of the phenomenon of *bulk viscosity*.[85])

One can view what is involved here as basically a transfer of potential energy from the global structure of the universe (gravitational potential energy) into the local energy of the matter, though there are well-known difficulties about defining energy in a precise way for models of this kind. But these difficulties should not concern us unduly here, since it is actually the entropy rather than the energy that is really relevant, and entropy has much more to do with probabilities and coarse-graining than it has to do with any particular definition of energy. In the example given by Tolman there is no state of maximum entropy, either achieved in any one specific model or throughout all models of this type. By choosing the value of R at maximum expansion to be sufficiently large (for fixed matter content), the total entropy can be made as large as we please. Tolman envisaged successive cycles of an 'oscillating' universe with gradually increasing maximum values of R. However, it is hard for us to maintain such a world-view now, because the singularity theorems[7,8] tell us that the universe cannot achieve an effective 'bounce' at minimum radius without violating the known† laws of physics.

From my own point of view, the situation envisaged by Tolman may be regarded as one aspect of the question of how the structure of the universe as a whole contributes to the entropy. It apparently concerns a somewhat different aspect of this question than does gravitational

† I am counting quantum gravity as 'unknown' whether or not it helps with the singularity problem!

clumping, since the Weyl tensor is everywhere zero in Tolman's models. It is clear that this has also to be understood in detail if we are to perceive, fully, the role of gravitational entropy. Nevertheless, it appears that the entropy available in Tolman's type of situation is relatively insignificant[78] compared with that which can be obtained – and, indeed, *is* obtained – by gravitational clumping (cf. section 12.3.3).

The key question must ultimately concern the structure of the singularities. These singularities, in any case, provide the boundary conditions for the various cycles in Tolman's 'oscillating' universe. Moreover, as we shall see in a moment, if strong cosmic censorship holds true, the presence of irregularities should not alter the all-embracing nature of these singularities in the case of an expanding and recollapsing universe.

12.3.2 Cosmic censorship†

Though it is by no means essential, for the viewpoint that I am proposing, to suppose that naked singularities cannot occur, such an assumption of 'cosmic censorship' does nevertheless greatly simplify and clarify the discussion. It has been my stated intention to adopt basically conventional attitudes on most issues, so I should not be out of line here, also, were I simply to align myself with what appears to be a majority view and (at least for purposes of argument) adopt a suitable assumption of cosmic censorship. But in the following pages I shall also give some independent justification of this view.

A preliminary remark is required before considering the details of this, however. In the Hawking process of black hole evaporation, the (supposed) final disappearance of the hole produces, momentarily, a naked singularity. This is not normally considered to be a violation of cosmic censorship because the Hawking process is a quantum-mechanical process, whereas cosmic censorship is taken normally to refer only to classical general relativity. (In the words of Hawking,[68] cosmic censorship is 'transcended' rather than violated!) Nevertheless, the presence of actual naked singularities of this kind in the geometry of the world would make a certain difference to the discussion. But the difference seems unlikely to be of any great relevance to the problem at hand. The black holes that we have any clear reason to believe actually exist in the universe are all of the order of a solar mass or more – and we have seen

† Parts of this section are considerably more technical than the others, and it can be omitted without seriously affecting the thread of argument.

that their lifetimes are greater than 10^{53} Hubble times, so their final naked singularities (!) can be safely ignored. Moreover, an implication of the viewpoint I shall set out in section 12.3.3 seems to be that mini-holes are unlikely to exist, these being the only black holes (say of mass 10^{20} g or less) whose final naked singularity could occur early enough to be of any remote relevance to the discussion. But since, in any case, such holes would be of no greater than atomic dimension, they could be 'smoothed over' and need not be considered as constituting a significant part of the classical geometry.

It seems, then, that a discussion of cosmic censorship entirely within classical general relativity should be perfectly adequate for our purposes. So what form of statement should we adopt? The usual formulation involves some assertion such as:

A system which evolves, according to classical general relativity with reasonable equations of state, from generic non-singular initial data on a suitable Cauchy hypersurface, does not develop any spacetime singularity which is visible from infinity.

Something of this kind, forbidding naked singularities, appears to be required in order that the usual general discussion of black holes can be carried through (e.g. the area-increase principle, the merging of two black holes necessarily forming a third, the general macrostability of black holes – indeed the very physical existence of black holes at all, rather than something worse, in a generic collapse[51]). The statement is vague in several respects, and a considerable increase in precision would be required in order to obtain something capable of clear mathematical proof, or disproof.

But it is probably not too helpful simply to make the various conditions more precise in some way, without having a deeper idea of what is likely to be involved. For example, it seems to me to be quite unreasonable to suppose that the physics in a comparatively local region of spacetime should really 'care' whether a light ray setting out from a singularity should ultimately escape to 'infinity' or not. To put things another way, some observer (timelike world-line) might intercept the light ray and see the singularity as 'naked', though he be not actually situated at infinity (and no actual observer would be so situated in any case). The observer might be close by the singularity and possibly himself trapped, e.g. inside the usual black hole of figure 12.5(*a*). The unpredictability entailed by the presence of naked singularities which is so abhorrent to many people

would be present just as much for this local observer – observing a 'locally naked' singularity – as for an observer at infinity.

It seems to me to be comparatively unimportant whether the observer himself can escape to infinity. Classical general relativity is a scale-independent theory, so if locally naked singularities occur on a very tiny scale, they should also, in principle, occur on a very large scale in which a 'trapped' observer could have days or even years to ponder upon the implications of the uncertainties introduced by his observation of such a singularity (compare the analogous discussion in section 12.2.6 of the astronaut inside a large black hole). Indeed, for inhabitants of recollapsing closed universes (as possibly we ourselves are) there is no 'infinity', so the question of being locally 'trapped' is one of degree rather than principle.

It would seem, therefore, that if cosmic censorship is a principle of Nature, it should be formulated in such a way as to preclude such *locally* naked singularities.[57,58,86] This viewpoint gains some support from, first of all, the standard picture of spherically symmetrical collapse inside a black hole as shown in figure 12.5(a). An observer who falls inside the hole cannot, in fact, see the singularity at all until he encounters it. This is perhaps clearer if we use a standard conformal diagram (with null-cones sloping at 45°) to depict the situation, as in figure 12.8, since then the spacelike nature of the singularity is brought out clearly and shows that it is *not* locally naked in the sense described above.

Secondly, there are some reasons for believing that generic perturbations away from spherical symmetry will not change the spacelike nature of the singularity (whose continued existence, in the perturbed case, is guaranteed by the singularity theorems[6–8]). The situation is slightly complicated, however, because the Schwarzschild–Kruskal singularity of figure 12.8 is, in fact, unstable. The introduction of a minute amount of angular momentum into the black hole to give a Kerr solution (with $a \ll m$) will actually change the singularity structure completely, and it ceases to be spacelike. Only when a further perturbation of a generic nature is made, may a spacelike structure of the singularity be expected to be restored.

It is somewhat easier to examine this behaviour if we consider adding charge rather than angular momentum, so that we get the Reissner–Nordstrom solution instead of the Kerr solution. The corresponding conformal diagram is shown in figure 12.9(a), and, indeed, it is evident that the singularity *is* now locally naked in the sense described above, since the observer whose world-line is γ can see the singularity. With a

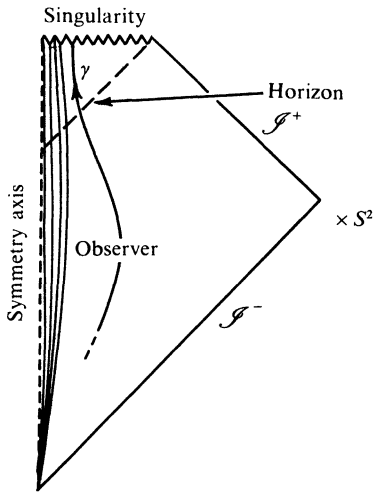

Figure 12.8. Conformal diagram of spherically symmetric collapse (with asymptotic flatness) illustrating the spacelike nature of the singularity.

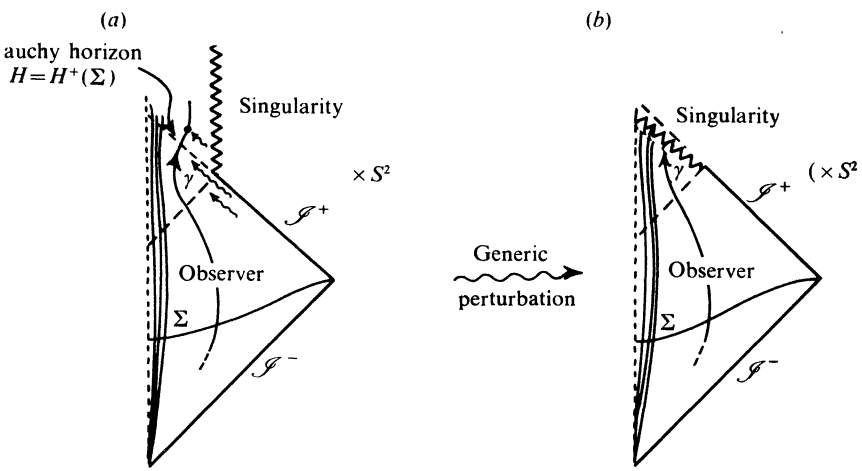

Figure 12.9. Conformal diagrams illustrating collapse to a black hole with a small charge: (*a*) spherically symmetric, (*b*) generically perturbed.

further perturbation of a generic nature one expects the situation more to resemble that of figure 12.9(*b*) in which the singularity is spacelike again (or perhaps null). The reason is that in the situation of figure 12.9(*a*) there is a null-hypersurface, H, in the spacetime, which is the *Cauchy horizon* of an (edgeless) spacelike hypersurface Σ extending all the way out to spatial infinity. (For terminology and notation, see reference 49.) An

observer γ who crosses $H = H^+(\Sigma)$ will see, as he does so, the entire future history of the outside world flash by in an instant. If the data on Σ are perturbed in some mild way out near infinity, then this will lead to a drastic alteration of the geometry in the neighbourhood of H. This is because signals from the infinite regions of Σ will be blueshifted at H by an infinite amount. Indeed, the analysis of weak-field perturbations (or test fields)[87,88] explicitly indicates that these perturbations diverge along H. With the full nonlinear coupling, this would presumably give a curvature singularity in place of H. Furthermore, it *could* well be that nonlinear effects actually result in a spacelike rather than just a null-singularity, because the effects of large curvature might proliferate and reinforce one another further and further up along the singularity (cf. figure 12.9(*b*)).

In order to make precise the notion of a spacelike (or null) spacetime singularity it will be convenient to recall the concept of an *ideal point*[89] of a spacetime $\mathcal{M}$, defined in terms of the TIPs or TIFs (terminal indecomposable past-sets or future-sets) of $\mathcal{M}$. The ideal points may be thought of as some 'extra points' adjoined to the manifold $\mathcal{M}$, either 'singularities' or 'points at infinity', for which the timelike curves in $\mathcal{M}$ that are future-endless in $\mathcal{M}$ acquire future ideal endpoints (via TIPs), and those that are past-endless in $\mathcal{M}$ acquire past ideal endpoints (via TIFs).

For simplicity, assume that $\mathcal{M}$ is strongly causal. Let γ and γ' be two future-endless timelike curves in $\mathcal{M}$. Then γ and γ' have the same future ideal endpoint if and only if they have the same pasts, which, in standard notation, is written $I^-[\gamma] = I^-[\gamma']$. The TIPs of $\mathcal{M}$ are, in fact, the sets of the form $I^-[\gamma]$, with γ future-endless and timelike, and may be 'identified' with the future ideal points. Similarly, past-endless timelike curves η and η' have the same ideal past endpoints whenever their futures are the same: $I^+[\eta] = I^+[\eta']$, these sets being the TIFs of $\mathcal{M}$. (See figure 12.10). In each case, the TIP or TIF is said to be *generated* by the timelike curve in question. A simple criterion[57,58] that may be used to distinguish those TIPs representing points at infinity from those representing singular points is to define a TIP as an ∞-*TIP* if it is generated by some timelike curve of infinite proper length into the future, and as a *singular TIP* if it is generated by no such curve. The ∞-*TIFs* and *singular TIFs* are similarly defined. (One may also choose to call some of the ∞-TIPs and ∞-TIFs singular, in some appropriate sense, but I shall not bother with this in detail here.)

Next, a locally *naked* singularity can be defined as either a singular TIP contained in the past $I^-(q)$ of some point q of $\mathcal{M}$, or as a singular TIF

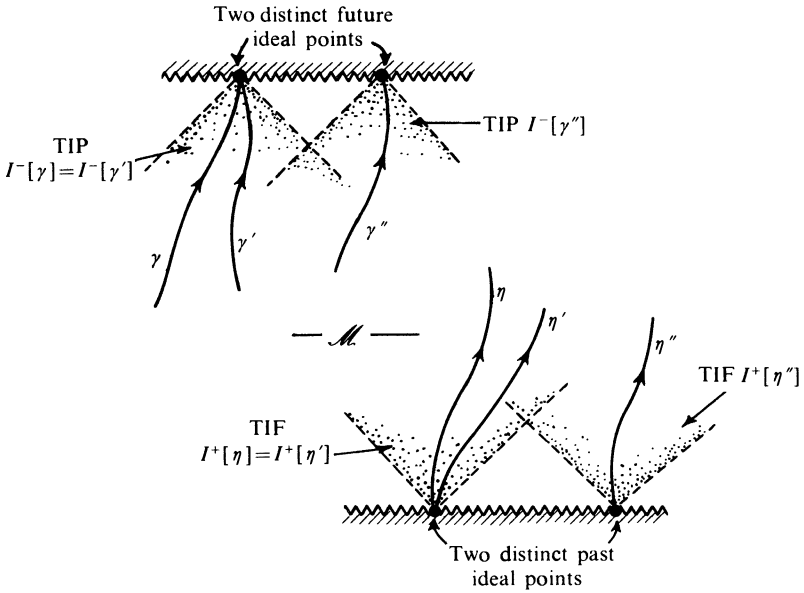

Figure 12.10. TIPs and TIFs defining ideal points of $\mathscr{M}$.

contained in the future $I^+(p)$ of some point p of $\mathscr{M}$ (figure 12.11). The upshot of such a definition is that in each case there is a timelike curve ζ (observer's world-line) from a point p to a point q, where the singularity lies to the future of p and to the past of q. (Take p in the TIP, in the first case, and q in the TIF, in the second.) The significance of this is that not

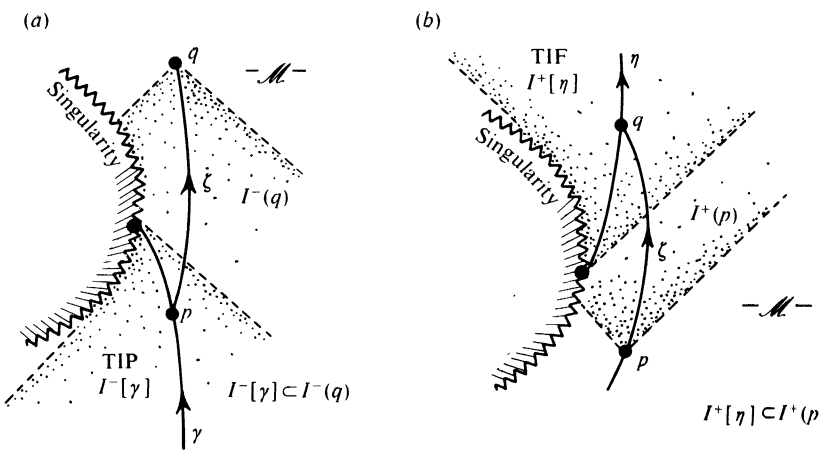

Figure 12.11. A locally naked singularity, lying to the future of p and to the past of q: (*a*) TIP definition, (*b*) TIF definition.

only can some observer (namely ζ) see the singularity (namely from q) but that at some earlier stage of his existence (namely at p) the singularity has yet to be produced. Thus, the usual all-embracing big bang does *not* qualify as locally naked, since no observer existed before it was produced.

One may say that $\mathcal{M}$ accords with a form of cosmic censorship if such locally naked singularities (as illustrated in figures 12.11(a) and (b)) do not occur.[57] However, it is convenient to go somewhat further than this and to exclude 'naked points at infinity' also, by dropping the condition that the TIP or TIF involved in figure 12.11 is necessarily a *singular* TIP or TIF. Indeed, it could be argued that such a naked point at infinity introduces as much indeterminacy into the future behaviour of a universe-model as does a naked singularity. But, in fact, naked points at infinity seem unlikely to arise unless they are also, in some appropriate sense, *singular* points (though still defined by ∞-TIPs or ∞-TIFs). The reason is that for a smooth conformal infinity $\mathscr{I}$, naked points at infinity can occur only if $\mathscr{I}$ is (in some places at least) timelike. And, with reasonable matter density, $\mathscr{I}$ can be timelike only if $\lambda < 0$, where λ is the cosmological constant.[90] However, the usual (Friedmann-type) $\lambda < 0$ models expand from, and recontract to, all-embracing singularities,[91] and hence possess no ∞-TIPs or ∞-TIFs, so (non-singular) naked points at infinity are not to be expected even in these cases.

Whether the TIPs or TIFs refer to singular points or to points at infinity, the situation of figure 12.11 is what characterizes ideal points as belonging to a boundary to $\mathcal{M}$ which is, in a sense, *timelike*. For if we extend ζ indefinitely into the future, in figure 12.11(a), to becomes a future-endless timelike curve ζ' we have

$$I^-[\gamma] \subset I^-(q), \quad q \in I^-[\zeta']$$

which is the condition[89] for the TIP $I^-[\gamma]$ to lie to the chronological (i.e. timelike) past of the TIP $I^-[\zeta']$. (The weaker condition $I^-[\gamma] \subset I^-[\zeta']$ obtains if $I^-[\gamma]$ lies to the *causal* past[89] of $I^-[\zeta']$.) Similarly, figure 12.11(b) gives us

$$I^+[\eta] \subset I^+(p), \quad p \in I^+[\zeta'']$$

with ζ'' past-endless, characterizing the TIF $I^+[\eta]$ as lying to the chronological future of the TIF $I^+[\zeta'']$. Thus, to rule out figure 12.11(a) configurations is to say that the future ideal points constitute an *achronal*[49] (i.e. roughly, spacelike or null) future boundary to $\mathcal{M}$, while to rule out those of figure 12.11(b) is to say that the past ideal points constitute an *achronal* past boundary to $\mathcal{M}$.

What is more, ruling out *either one* of these configurations throughout $\mathcal{M}$ is equivalent to ruling out the other, since each condition turns out to be equivalent[57] to the time-symmetric condition that $\mathcal{M}$ be *globally hyperbolic.*[7,49]

The proof of this fact is quite simple and is worth outlining here since it has not been spelled out in the standard literature. First, a globally hyperbolic spacetime is one with the property that for any two of its points p, q, the space of causal curves from p to q is compact. (A causal curve is a curve, not necessarily smooth, that everywhere proceeds locally within, or on, the future light-cone. Thus, causal curves are timelike curves or curves that are everywhere locally the C^0 *limits* of timelike curves.) Strong causality being assumed, global hyperbolicity is in fact equivalent to the statement that each $I^+(q) \cap I^-(p)$ has compact closure.[7,49] I also remark here that for a causal curve ζ, the set $I^-[\zeta]$ is a TIP if and only if ζ is future-endless,[89] with a similar result holding for TIFs.

Now suppose that $\mathcal{M}$ contains a point q and a future-endless timelike curve γ such that $I^-[\gamma] \subset I^-(q)$. It follows that $\mathcal{M}$ cannot be globally hyperbolic. For if p is a fixed point on γ, and $r_1, r_2, r_3, \ldots$ is a sequence of points proceeding indefinitely far up γ, we obtain a sequence $\zeta_1, \zeta_2, \zeta_3, \ldots$ of causal curves from p to q where ζ_i consists of that segment of γ from p to r_i together with a timelike curve from r_i to q (which exists because $r_i \in I^-[\gamma]$ and $I^-[\gamma] \subset I^-(q)$). If $\zeta_1, \zeta_2, \ldots$ had a limit causal curve ζ, from p to q, then $I^-[\zeta'] = I^-[\gamma]$, where $\zeta' = \zeta \cap I^-[\gamma]$ (as is easily seen), whereas ζ' cannot be future-endless, being continued as ζ to the future endpoint q. This would contradict the last statement of the preceding paragraph, showing that ζ does not exist and that $\mathcal{M}$ is consequently not globally hyperbolic.

The converse argument is slightly more technical. I use the notation and proposition numbers of reference 49. Suppose $\mathcal{M}$ is not globally hyperbolic. Then points p and q exist for which $I^+(p) \cap I^-(q)$ does not have compact closure. Hence (propositions 3.9, 5.20) $p \notin \text{int } D^-(\partial I^-(q))$, whereas $p \in I^-(q)$. Consequently (proposition 5.5h), $p' \notin D^-(\partial I^-(q))$, where $p' \in I^-(p)$, so there is a future-endless timelike curve γ from p' which does not meet $\partial I^-(q)$. But $p' \in I^-(q)$, whence $\gamma \subset I^-(q)$ and the TIP $I^-[\gamma]$ lies entirely within $I^-(q)$, as is required to prove.

My proposal[57,59] for a *strong cosmic censorship principle* is, therefore, that a physically reasonable classical spacetime $\mathcal{M}$ ought to have the property which can be stated in any one of the following equivalent forms: no TIP lies entirely to the past of any point in $\mathcal{M}$; no TIF lies entirely to the future of any point in $\mathcal{M}$; the TIPs form an achronal set; the

TIFs form an achronal set; $\mathscr{M}$ is globally hyperbolic; there exists a Cauchy hypersurface for $\mathscr{M}$. (This last equivalence is a well-known result due to Geroch.[93]) The plausibility of this rests, of course, on what we expect of a 'physically reasonable' spacetime. Indeed, global hyperbolicity has tended to be regarded by many people as an over-strong restriction. Nevertheless, I believe that one *can* put forward plausibility arguments to support strong cosmic censorship. I next give an indication of this.

To begin with (except in the case of the big bang) it seems not unreasonable to restrict attention, in the first instance, to the case of vacuum only. The reason for this was indicated in section 12.3.1, namely that near 'generic' singularities we expect the Weyl curvature to dominate over the Ricci tensor. This is not totally satisfactory, however, because there are 'cumulative' effects due to the Ricci tensor (namely focussing) which the Weyl tensor can achieve only indirectly through nonlinearities. Nevertheless, the behaviour of vacuum solutions would seem to give a good first approximation near a generic singularity, which avoids the problems raised, for example, by the presence of apparently inessential 'shell-crossing' naked singularities[92] in idealized matter such as 'dust'. As a second approximation one could consider the Einstein–Maxwell theory, for example, which likewise avoids these problems. However, the big bang is a special situation (which relates to its low entropy) and the criterion of 'genericity' need not apply.† But strong cosmic censorship still seems to hold – for different reasons (cf. section 12.3.3). In the singularities of collapse, however, a high-entropy assumption of 'genericity' seems physically reasonable.

The sequence depicted in figures 12.8, 12.9(a) and 12.9(b) would seem to provide a plausible pattern for the general situation. In figure 12.8 (extended Schwarzschild solution), global hyperbolicity holds, but apparently fortuitously. Each singular TIP intersects a Cauchy hypersurface Σ in a compact region. The data in that region *alone* are all that are needed to imply the existence and nature of the singularity that the TIP represents. But perturb the solution slightly, so that it becomes that of Kerr and extend it maximally in the usual way (as in figure 12.9(a)); then the singularity disappears – in the sense that no singular TIP near to the original one now exists, but is replaced by the past $I^-(x)$ of a non-singular interior point x. Thus the original singularity apparently owes its existence to some very special aspect (e.g. the exact spherical symmetry) of the original initial data. Global hyperbolicity is violated in the slightly

† Indeed, it is perfectly in order for the big bang to be what, in the reverse direction in time, would be an *unstable* singularity (cf. section 12.3.3).

perturbed solution, but *because* of this violation there is now a Cauchy horizon $H = H^+(\Sigma)$. The past $I^-(y)$ of a point y of H intersects Σ in a region with *non-compact* closure (extending to infinity). We may take it that the structure of the spacetime (e.g. the curvature) at y is the result of some form of integral of the initial data over this region. If the data are perturbed generically, we are liable to get a *divergence* (owing to the non-compactness and resulting infinite blueshift) so that the non-singular point y changes, in effect, to a singular ideal point with (presumably) diverging curvature.

Suppose, now, that, instead of the asymptotically flat situation we have been examining, we consider a spacelike initial hypersurface Σ that is *compact*. It may still be that, as in figure 12.9(a), certain sets of data on Σ lead to a maximally extended vacuum spacetime which violates global hyperbolicity (e.g. Taub–NUT space $\mathcal{M}$), and a Cauchy horizon $H = H^+(\Sigma)$ is produced. Take $y \in H$, as before, and consider $I^-(y) \cap \Sigma$. This must now have compact closure (since Σ is compact) but it seems that, in a sense, it is liable to be *effectively non-compact* owing to an infinite wrapping round and round Σ by $I^-(y)$ – at least this is what appears to be indicated by an examination of the Bianchi IX models.[80,94] This effective non-compactness would seem to lead to a situation similar to the one just discussed, in which H gets converted to a curvature singularity upon generic perturbation, and strong cosmic censorship holds, because the integrals which define the perturbed curvature on H involve the same data on Σ over and over again, infinitely many times.

One is tempted to conjecture, therefore, that the singularities, like that of figure 12.8, which result from data on an effectively compact region are a 'measure zero' special case, and that while some perturbations give rise to Cauchy horizons, the points on these horizons are liable to be dependent upon effectively non-compact data regions, with infinite blueshift, so that the horizons should be unstable (like that of figure 12.9(a)). In this sense, strong cosmic censorship looks to be very plausibly true† – but the above argument is yet a long way from a proof.

We are thus presented with the picture of a globally hyperbolic universe, which starts from an achronal set $\partial\mathcal{M}$ of initial ideal points (the big bang), then remains topologically unchanging (an implication of global hyperbolicity[93]) – despite the presence of black holes – until the

† From many different viewpoints it is *mathematically* very desirable to be able to restrict attention to globally hyperbolic spacetimes. For example, most of the technical difficulties[89] concerning the topology and identifications for TIP and TIF structure now disappear.

achronal set $\partial\hat{\mathscr{M}}$ of final ideal points is reached. (Strong cosmic censorship implies, in fact, that $\partial\check{\mathscr{M}}$ and $\partial\hat{\mathscr{M}}$ must be regarded as totally disjoint sets.) The initial ideal points are normally taken to be all singular points, but the final ideal points may be either points at infinity or singular points. Points at infinity would normally be thought to arise only in the case of an ever-expanding universe-model, in which case one would also expect singular final ideal points in black holes. But it is also conceivable (though in my view rather unlikely, for reasons similar to those just outlined above) that a universe which recollapses as a whole may nevertheless have certain limited portions that 'escape' to infinity in a (non-singular) ∞-TIP.

In figure 12.12, an indefinitely expanding universe-model is illustrated, showing how $\mathscr{M}$ can remain topologically unchanging – in the sense that $\mathscr{M} \cong \mathbb{R} \times \Sigma$, with each copy of $\mathbb{R}$ a timelike curve and each copy of Σ a spacelike Cauchy hypersurface[93] – even though there may be several

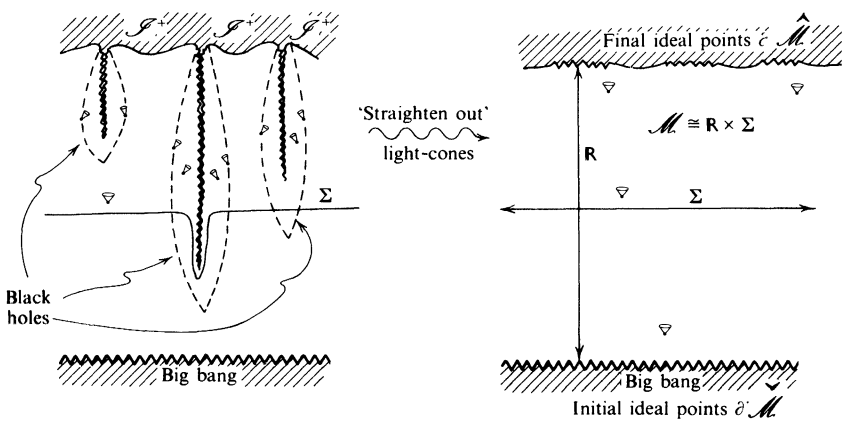

Figure 12.12. A universe-model subject to strong cosmic censorship. The indefinitely expanding case is illustrated. (The different $\mathscr{I}^+$ regions are actually all connected.)

black holes present.[57] The situation for a recollapsing universe-model is similar. The sets $\partial\check{\mathscr{M}}$ and $\partial\hat{\mathscr{M}}$ may or may not have the same topology as Σ, however. (For example, in the Einstein static universe, each of $\partial\check{\mathscr{M}}$ and $\partial\hat{\mathscr{M}}$ is a point, whereas Σ is an S^3.) According to the viewpoint in section 12.3.3, the big bang $\partial\check{\mathscr{M}}$ should, indeed, have the topology of Σ. But it is not at all clear that this should be so of $\partial\hat{\mathscr{M}}$. Another way of phrasing Misner's original hope that the generic Bianchi type IX 'mixmaster' empty models should be free of particle horizons[94] is that $\partial\check{\mathscr{M}}$ should be a

single point.† In time-reversed form, one might have correspondingly expected $\partial\hat{\mathcal{M}}$ to be a single point. But later analysis[95–98] has shown this kind of behaviour to be a very unlikely possibility. Nevertheless, one might anticipate that under certain circumstances $\partial\hat{\mathcal{M}}$ has, in an appropriate sense, fewer than three dimensions (as is the case for the two-dimensional $\partial\hat{\mathcal{M}}$ of a pancake singularity and for Taub space[7]).

If the big bang $\partial\check{\mathcal{M}}$ is, indeed, a smooth spacelike hypersurface with the topology of Σ (cf. section 12.3.3) then it constitutes a very satisfactory initial Cauchy hypersurface. But irrespective of the smoothness of three-dimensionality of $\partial\check{\mathcal{M}}$ and $\partial\hat{\mathcal{M}}$, each can be regarded as providing a home for the ultimate Cauchy data for $\mathcal{M}$ – the ultimate initial data in the case of $\partial\check{\mathcal{M}}$, and the ultimate final data in the case of $\partial\hat{\mathcal{M}}$. Each is, indeed, all-embracing in the sense of being intersected by every endless causal curve in $\mathcal{M}$, and is also achronal. Of course, the exact form that such Cauchy data would take is not yet clear in the general case, but the potentiality is certainly there in principle.

As a final consideration of this section, let us examine the possibility of an initial creation or the final destruction of a charged particle (cf. also section 12.2.7). If the initial creation, at $\partial\check{\mathcal{M}}$, can be regarded as the result of a more-or-less understood process[62,63,66] whereby the curvature itself creates particles, then we would expect not just a single charged particle, but a *pair* of oppositely charged particles. However, it might be that some totally unknown process is involved in particle production at the big bang, whereby charged particles could be produced singly.[67] In this situation, the particle's Coulomb field must also be conjured up at $\partial\check{\mathcal{M}}$, leading, as was remarked in section 12.2.4, to the existence of some effectively source-free incoming radiation.[45] Indeed, such a picture (at first sight, at least) would be presented as the time-reverse of particle destruction at $\partial\hat{\mathcal{M}}$. Imagine a single charge being swallowed by a large spherically symmetrical uncharged black hole. According to the geometry depicted in figure 12.8, this particle must be annihilated alone, on a singular part of $\partial\hat{\mathcal{M}}$. Is it conceivable that this one particle should so disturb the geometry of $\partial\hat{\mathcal{M}}$ that its destruction is held off until finally a particle of the opposite charge is swallowed by the hole and guided in to be annihilated by the first charge, before the energy of both can be absorbed by the singularity? It seems hardly credible. Yet the structure of generic singularities appears to be so delicate and elaborate that even this should not be dismissed out of hand. But the issue that I am attempting to raise here is whether or not

† This is also, and more explicitly, what is demanded for a Sachs–Budic 'deterministic' spacetime.[99]

the laws that dominate physical behaviour at a spacetime singularity can, after all, be time-symmetric. It seems to me that those who believe that they can be must face additional serious problems of principle!

12.3.3 A hypothesis and a world-view

Almost all the arguments that I have been presenting seem to focus themselves on one issue: what special geometric structure did the big bang possess that distinguishes it from the time-reverse of the generic singularities of collapse – and why? At this point I should mention the viewpoint of *chaotic cosmology*,[94] which has been in vogue for a number of years. According to this view, the big bang was not initially a very uniform singularity, but came to appear so because dissipative effects (e.g. neutrino viscosity, hadron collisions or particle creation[100–102]) served to iron out all its irregularities. One intended effect of this dissipation is to produce the observed 'high' entropy-per-baryon figure of $\sim 10^9$. Another is to produce the presently observed isotropy. And in order not to impose apparently arbitrary restrictions on the big bang, it was originally conjectured that the chaos of the initial singularity was, in some appropriate sense, *maximal*. In this way a supposedly canonical type of initial state was suggested, whose detailed implications could be worked out and compared with observation.

I have to say, however, that I regard the programme of chaotic cosmology – at least in its 'pure' form of maximal initial chaos – to be basically misconceived. For to assert that the initial chaos is maximal is presumably to assert that the initial *entropy* is maximal. And if this were the case, with time-symmetric physical laws, there would be no time's arrow and therefore no dissipation. (It is no good appealing to the expansion of the universe here – for reasons that were amply discussed in section 12.2.6!)

We might, however, entertain some milder version of chaotic cosmology in which the initial geometric chaos was not maximal but was constrained in some suitable way – the constraints being so adjusted that not only could the observed entropy per baryon be correctly obtained from early dissipation, but that density irregularities adequate for galaxy formation might also be obtainable as a result of a non-uniform initial geometry.[103,105] In my opinion, this version of chaotic cosmology is also quite unsatisfactory. The essential misconception seems to be to regard an entropy per baryon of $\sim 10^9$ as a *high* figure in a gravitational context. Consider a closed recollapsing universe containing, say, 10^{80} baryons.

When an $M_\odot$ black hole forms, it achieves, by the Bekenstein–Hawking formula, an entropy per baryon of 10^{18}, while a $10^{10}\,M_\odot$ galaxy with a $10^6\,M_\odot$ black hole core has an entropy per baryon of 10^{20}. When finally the bulk of this galaxy collapses into the hole, the figure is 10^{28}. But as the collapse of the universe proceeds, these black holes unite into bigger and bigger ones, yielding an immense final entropy per baryon of $\sim 10^{40}$. The entropy resides in the irregularities of the geometry of the final singularity. Reversing the time-sense, we now see what a stupendous entropy would really have been available, had the Creator chosen to make use of it by an initial chaotic geometry. The supposedly 'high' figure of $\sim 10^9$ is small fry indeed! If we are to adopt 'mild' chaotic cosmology as an explanation for the figure 10^9, then we must ask why only the absurdly small fraction of at most $\sim 10^{-31}$ of the 'available chaos' was actually used! (Indeed, the figure would be $\sim 10^{-32}$ if the closed-universe value $\sim 10^{-8}$ for the entropy per baryon were used.)

It would seem that, with such an enormous discrepancy, we should not look for a gravitational explanation† for the figure 10^9. A more hopeful place to look might be in particle physics. I shall return to this question in section 12.4. Likewise it would seem that a *purely* gravitational explanation of the irregularities needed for galaxy formation will not be forthcoming. Again (though with considerably less confidence) I appeal to particle physics – in the very early stages of the expansion.

I propose,[57–60] then, that there should be a complete lack of chaos in the initial *geometry*. We need, in any case, some kind of low-entropy constraint on the initial state. But thermal equilibrium apparently held (at least very closely so) for the *matter* (including radiation) in the early stages. So the 'lowness' of the initial entropy was not a result of some special matter distribution, but, instead, of some very special initial spacetime geometry. The indications of sections 12.2.6 and 12.3.1, in particular, are that this restriction on the early geometry should be something like: *the Weyl curvature C_{abcd} vanishes at any initial*‡ *singularity.*[58–60]

This hypothesis is still a little vague, and is open to a number of different interpretations. We could require, for example, that the Weyl curvature tend to zero as the initial singularity is approached, or that it should do so at some preassigned faster rate, or, perhaps, that it should

† This is borne out by a recent detailed analysis by Barrow and Matzner[104] who come to the same essential conclusion.

‡ That is, with the notation of section 12.3.2, at $\partial\mathscr{M}$. Note that this includes the final singular TIF of a Hawking black hole explosion – at which the Weyl curvature indeed vanishes!

only remain bounded (or merely dominated by the Ricci tensor near the singularity – so that the curvature tensor becomes, in the limit, *proportional* to one whose Weyl part vanishes). I have not examined the differences that might be entailed by adopting different versions of this hypothesis. My inclination is to try, first, the simplest of these, namely that $C_{abcd} \to 0$ (in, say, any parallelly propagated frame), as initial singularities are approached. I shall indicate, shortly, what the rough implications of this would be, though complete details have yet to be worked out.

Let me first explain my view of the role of this hypothesis, as it affects the 'selection' that the Creator might make of *one* particular universe out of the apparently infinite choice available, consistent with given physical laws. Imagine some vast manifold $\mathcal{U}_0$ (where I use the word 'manifold' in a rather loose sense) representing the different possible initial data for the universe, compatible with the physical laws. To select one universe, the Creator simply places a 'pin' somewhere in $\mathcal{U}_0$. But a viewpoint of this article is that we should not be biased in choosing initial rather than final data. So, equally well, the selection of the universe could be envisaged as the Creator's pin being placed in the manifold $\mathcal{U}_\infty$, representing all possible sets of final data compatible with the physical laws. Indeed, one could use any intermediate $\mathcal{U}_t$ representing the possible data 'at time t'.† All are equivalent, the equations of motion (which I am crudely assuming, for the purposes of the present discussion, to be of a classical determinate type) effecting canonical isomorphisms between $\mathcal{U}_0$, $\mathcal{U}_\infty$ and each $\mathcal{U}_t$. Thus we may envisage a *single* isomorphic abstract manifold $\mathcal{U}$ which represents any (or every) one of these, being the set of all possible universe-histories compatible with the physical laws.

We would like to be able to envisage that the Creator's pin is simply placed 'at random' in $\mathcal{U}$ (since if the pin is to be constrained in any further precise way, this would constitute another 'physical law'). But the notion of 'at random' requires that some appropriate *measure* be placed on $\mathcal{U}$. Is 'at random' the same concept when applied to initial conditions as when applied to final conditions? To put the question another way: is the phase-space measure that is naturally defined on $\mathcal{U}_0$ the same as that defined on $\mathcal{U}_\infty$ (or indeed on each $\mathcal{U}_t$) under the isomorphisms? Liouville's theorem tells us that it *is*, provided that we are adopting conventional Hamiltonian physics – and I am not proposing to be

† The concept 'time' is being used very loosely here. In the right-hand picture of figure 12.12 (where indefinite expansion need not be assumed) t could be some suitable parameter ranging from 0 to ∞ and measuring the 'height' up the picture.

'unconventional' in this respect here. I am, furthermore, going to ignore any difficulties† that might arise from a possible infinite-dimensionality, or infinite total measure, of the manifold $\mathcal{U}$ – not to mention all the very serious 'gauge' problems (etc.) which would arise in a proper general-relativistic treatment.

How are we to envisage the entropy of the universe according to this picture? A standard procedure is to coarse-grain each $\mathcal{U}_t$ by dividing it up into compartments, where the different members of any one compartment correspond to states macroscopically indistinguishable from one another at time t. If the Creator's pin pierces $\mathcal{U}_t$ at a point belonging to a compartment of phase-space volume V (measured in units where $\hbar = 1$) then the (Boltzmann) entropy at time t is k log V (cf. section 12.2.3). This entropy can fluctuate‡ or progressively change with time, because the coarse-grainings of the different $\mathcal{U}_t$ manifolds do not map to one another under the isomorphisms. Low entropy at one time (point in a small compartment) can correspond to high entropy at another (point in a large compartment), where the 'specialness' of the state has now become that of macroscopically indiscernible correlations.

In the present context, however, this description of entropy is not yet satisfactory. The points of $\mathcal{U}$ correspond only to those universe-histories that are compatible with the physical laws at *all* times. It may not be macroscopically discernible at a particular time t whether the laws are satisfied at all other times. I am hypothesizing, here, that there are in fact (local) physical laws which only become important near spacetime singularities, these being asymmetric in time and such as to force the Weyl curvature to vanish at any initial singular point (i.e. point of $\partial\mathcal{M}$). The effect of these laws is that each manifold $\mathcal{U}_t$ turns out to be much smaller than might otherwise have been expected. Only those motions and configurations at time t which are also compatible with the constraints ($C_{abcd} = 0$) at $t = 0$ are allowed. But since these implied constraints on each $\mathcal{U}_t$ are not macroscopically discernible, it is not reasonable, when calculating the entropy at time t, simply to use the phase-space volumes within $\mathcal{U}_t$. Instead, one must consider extended volumes within a certain

† I must apologize, particularly to the experts, for my crude and cavalier treatment of the delicate matters that I am embarking upon. My excuse is that for the questions with which I am now concerned, I do not believe that general-relativistic or thermodynamic sophistication is a key issue.

‡ If we prefer the *ensemble* picture of the world, we can to some extent avoid these fluctuations, envisaging that the Creator uses a *blunt* pin! So long as the diameter of the pin-point is large compared with the coarse-graining, fluctuations are smoothed over.

larger manifold $\mathcal{W}_t$ – defined in the same way as $\mathcal{U}_t$ but for which these (initial) constraints are not required to hold.

The equations of motion again give isomorphisms between the different $\mathcal{W}_t$ manifolds at different times t (locally at least – and so long as the extra constraining laws remain physically insignificant) and there is a corresponding abstract manifold $\mathcal{W}$ representing the totality of unconstrained universe-histories. The imbedding of $\mathcal{U}$ in $\mathcal{W}$ has a very special relation to the coarse-graining at $t = 0$, because the vanishing of Weyl curvature *is* a macroscopically discernible property. Thus, no $t = 0$ compartment of $\mathcal{U}$ extends outside $\mathcal{U}$, into $\mathcal{W}$. But as t increases, the corresponding coarse-graining compartments of $\mathcal{U}$ extend more and more into $\mathcal{W}$, and accordingly acquire larger and larger extended volumes.† In this way, we regard the 'specialness' of the actual state of the universe (arising from its having started out with $C_{abcd} = 0$) as being more and more of the 'precise-correlations' kind, and less and less of the 'low-entropy' kind, as time progresses.

This is what gives us compatibility with the entropy-increase phenomenon of our universe. In this view, we do not impose any statistical low-entropy assumption at the big bang, but, instead, a precise *local condition* ($C_{abcd} = 0$). Aside from this constraint, there is to be complete randomness, that is, the Creator's pin is placed at random in the manifold $\mathcal{U}$. With this randomness assumption, we can attribute the 'reason' for the absence of initial correlations between particle motions in the initial state (i.e. for the *law of conditional independence*[40]) to the fact that at no time *other* than $t = 0$ is a new local constraint imposed (e.g. there is none imposed at the final singularity $t = \infty$). Correspondingly the 'reason' for the presence of increasing correlations as t increases, and for the increasing entropy, is the initial $C_{abcd} = 0$ constraint. In this way, the problem of time's arrow can be taken out of the realm of statistical physics and returned to that of determining what are the precise (local?) physical laws. I shall briefly discuss this question in section 12.4.

I should make mention, at this point, of the much discussed *anthropic principle*,[106] which is often invoked in connection with the matters I have been raising. This principle would, in effect, imply that the Creator's pin is placed in $\mathcal{U}$, not just at random, but with a weighting factor, weighted in

† Curiously, if it were not for extensions into $\mathcal{W}$, the volumes of the largest compartments would *decrease* with time owing to the profusion of differing macroscopic geometries that would be produced. This seems to correspond to the fact that the universe can get more 'interesting' even though the entropy increases!

favour of universes containing (many?) *conscious observers.* (Furthermore, the pin could also pierce other manifolds $\mathscr{V}$, $\mathscr{W}$, ... corresponding to all the possible consistent alternative sets of pure-number physical constants, and to all the possible alternative laws of physics. I shall leave aside a discussion of this extended question as being 'beyond the scope of this article'!) Such an anthropic viewpoint has occasionally been invoked in an attempt to explain the entropy imbalance of the observed universe in terms of weighting in favour of a huge 'fluctuation' which might have been needed to produce the conditions necessary for life.[107] The trouble with this is that it is vastly 'cheaper' (in terms of negative entropy) simply to produce a few conscious beings out of some carefully organized particle collisions than it is to build, out of a fluctuation, such entropy imbalance as is familiar on Earth throughout the *entire*, apparently unending, universe – as revealed by the most powerful telescopes!

This is not to say that I believe 'randomness' for the Creator's pin will always remain the best explanation for the state of the world. But, with the addition of the initial $C_{abcd} = 0$ assumption, it seems to work remarkably well, and it saves our having to worry about what 'consciousness' means in physical terms – at least for the time being!

So what, indeed, are the implications of the world-view that I am proposing? With the Weyl curvature initially zero and thermal equilibrium for the matter (and radiation), we shall have something very close to spatial-isotropy and homogeneity for the initial state. Thus the discussions of Friedmann, Robertson and Walker hold good (initially), leading to a striking consistency with various remarkable observations: the uniformity of the 2.7 K black-body radiation (to one part in $\sim 10^3$),[108] the lack of measurable rotation of the universe relative to inertial frames ($< 10^{-16}\ \mathrm{s}^{-1}$),[77] the large-scale gross uniformity of galactic clusters. The big bang singularity itself must have been closely of the Robertson–Walker type.† *Some* fluctuations in the matter distribution are allowed in this view – and, indeed, *must* occur – because the initial restrictions on R_{ab} are statistical, unlike those on C_{abcd}. However, the initial vanishing of C_{abcd} imposes considerable constraints on such initial density and velocity fluctuations. Yet, particle physics will have to be better understood before these fluctuations can be calculated in detail.[109]

White holes are excluded at all times, because their singularities are 'initial' singularities (i.e. points of $\partial\mathscr{M}$) which do not remotely satisfy the constraint $C_{abcd} = 0$. Black holes are allowed, of course, provided that

† This is independently supported by the accuracy of the helium-production calculations.[77,78]

they are formed in the normal way as a consequence of gravitational collapse of a massive body or bodies. But mini-holes are presumably *not* to be expected because they require an initial state with a chaotic geometry. The non-existence of such primordial black holes is consistent with present observations.[110]

Finally, the extraordinary observed behaviour of the entropy of our world – permeating so much of our everyday experience that we tend to take it quite for granted – is the major and most striking consequence of this world-view.

12.4 Asymmetric physics?

Some readers might feel let down by this. Rather than finding some subtle way that a universe based on time-symmetric laws might nevertheless exhibit gross time-asymmetry, I have merely asserted that certain of the laws are not in fact time-symmetric – and worse than this, that these asymmetric laws are yet unknown! But this is not so negative as it might seem. In particular, it tells us to look out for such asymmetries in other places in physics. Where do we look? Ultimately there must be some connection with gravity, since it is the Weyl tensor that describes gravitational degrees of freedom. Classical general relativity is a time-symmetric theory, but one may ask whether this time-symmetry will persist when finally the link with quantum mechanics is appropriately forged. Indeed, if one believes that virtual black holes at the Planck length (10^{-33} cm) are physically important,[73] then the arguments of section 12.2.7 suggest that the vacuum could be time-asymmetric in a significant quantum-gravitational way.

This is not greatly helpful, however, and there is another possible tentative connection between quantum mechanics and gravity that might be more relevant, namely the question of *quantum-mechanical observations*, that was left hanging in section 12.2.2. Certain associations with the Bekenstein–Hawking discussion of entropy should be pointed out. One regards a quantum-mechanical observation to have been 'made', after all, only when something 'irreversible' has taken place. But 'irreversible' here refers to the fact that an essential increase in the entropy has occurred. And entropy, as we have seen, seems to depend on the rather subjective concept of coarse-graining. For a quantum-mechanical observation 'actually' to take place, effecting a real change in the state of the world, one would seem to require *objectivity* for this entropy increase. Now recall that in the Bekenstein–Hawking formula, an entropy measure

is directly put equal to a precise feature of spacetime geometry, namely the surface area of a black hole. Is it that this geometry is now subjective, with the implication that *all* spacetime geometry (and therefore all physics) must in some measure be subjective? Or has the entropy, for a black hole, become objective? If the latter, then may not entropy also become objective in less extreme gravitational situations (cf. section 12.3.1)? Moreover, if it is only with (quantum) gravity that such a passage from subjectivity to objectivity of entropy can occur, then it is with (quantum) gravity that the linear superposition of von Neumann's chain[111] finally fails!

Perhaps, then, the new laws that seem to be needed to extend quantum mechanics (cf. section 12.2.2), so that observations can be incorporated within the theory, constitute some form of *quantum gravity* – by which I mean a theory having quantum mechanics and general relativity as two appropriate limits. I would contend, in any case, that the arguments I have been presenting (notably those of sections 12.2.7 and 12.3.3 which most directly relate to the Bekenstein–Hawking formula) point towards some new theory which is *time-asymmetric.* Accordingly, whatever nonlinear physics† eventually replaces suddenly collapsing wave functions may well turn out to involve an essential time-asymmetry.

At the present state of understanding, such considerations are highly speculative. Yet we know that there *is* a physical law which is time-asymmetric! Somewhere, hidden among the more familiar time-symmetric forces of Nature is one (or perhaps more than one) whose tiny effect has been almost completely masked by these others and has lain undetected in all physical processes bar one: that delicately poised decay of the K^0-meson. I am not suggesting that quantum gravity need be involved here. But, evidently, it is *not* one of Nature's inviolable rules that time-symmetry must always hold!

Moreover, the relative strength of the T-violating (or CP-violating) component in K^0-decay, perhaps 10^{-9}, is possibly suggestive. According to section 12.3.3, we need some explanation from particle physics for the observed entropy-per-baryon number of 10^9. It has, accordingly, been put forward[109,67,62] that perhaps in the early expansion there was a

† An interesting recent suggestion for a nonlinear modification of Schrödinger's equation is that of Bialynicki-Birula and Mycielski[112] according to which an additional term $(b \log |\psi|^2)\psi$ is incorporated. Though not time-asymmetric, there is the link with the present discussion that the constant b (a temperature) has the value $\sim 10^{-8}$ K, which is the Bekenstein–Hawking temperature for a $10\,M_\odot$ black hole. These are the smallest black holes that one has physical reason to believe in.

process of production of baryons and antibaryons, where baryons outnumbered antibaryons by about $1+10^{-9}:1$. Then, in the subsequent annihilation, not only would the (apparently) required nonzero net baryon number be produced, but also the observed entropy per baryon (i.e. photons per baryon). This initial production process, if it were to be explained as arising from something with the vacuum's quantum numbers, would need to violate baryon conservation and (for an imbalance to occur) both *CP*- and *C*-symmetries.† Gross violation of *C* has long been known to be a feature of weak interactions,[20] while *CP*-violation also occurs in the K^0-decay at the apparently required low level. Moreover, baryon non-conservation[76] is *somewhere* to be expected, on the basis of Hawking's black hole evaporation.[56,67]

But to explain the law that $C_{abcd} = 0$ initially, we would need something more, namely a violation of each of *T*, *PT*, *CT* and *CPT*. If, for example, *CPT* were not violated, we could take an allowed collapse to a singularity for which $C_{abcd} \neq 0$ and apply the *CPT*-symmetry to obtain a disallowed initial singularity. The symmetries *PT* and *CT* are maximally violated in weak interactions and *T* marginally violated in K^0-decay, but *CPT*-violations have not yet been detected. Of course there is the *CPT*-theorem[113] that lends some theoretical support to universal *CPT*-conservation. But it must be borne in mind that Poincaré covariance is an important assumption of that theorem, whereas I am envisaging situations (singularities, quantum gravity) for which this assumption is explicitly violated. I would contend that somehow, in experimental physics, a *CPT*-violating effect ought eventually to be discernable. But these are early days yet, and it is in no way surprising that such effects have not yet been seen.

I have presented, here, little in the way of quantitative detail. However, many of the phenomena I have been concerned with are of so gross and blatant a nature that much can yet be gleaned even without such detail. The most blatant phenomenon of all these is the statistical asymmetry of the universe. It is, to me, inconceivable that this asymmetry can be present without tangible cause. An explanation by way of the anthropic principle seems very wide of the mark (cf. section 12.3.3). So does an explanation of the 'symmetry-breaking'[114] type, according to which the most probable states of the universe might not share the symmetries of the laws that govern it. (It is difficult to see how our vast universe could

† Or, conceivably, *CPT*-violation in place of *CP*-violation, or *CT* in place of *C*, since there is a time-asymmetry in the universe expansion.

just 'flop' into one or the other of these states when it doesn't even know which temporal direction to start in!) In my own judgment, there remains the *one* ('obvious') explanation that the precise physical laws are actually *not* all time-symmetric!

The puzzle then becomes: why does Nature choose to hide this time-asymmetry so effectively? As we do not yet know the principles that govern Nature's choices of physical law, we cannot yet answer this question. But perhaps we should not be so surprised at a situation in which a fundamental asymmetry lies hidden deep beneath a facade of apparent symmetry. The fauna of this Earth, after all, exhibits with but few exceptions a superficial external bilateral symmetry. How could one have guessed that in the nucleus of every reproducing cell lies a helix whose structure governs the growth of these magnificent symmetrical creatures – yet every one of which is right-handed?

Acknowledgements

I am grateful to Dennis Sciama and, particularly, to Amelia Rechel-Cohn for critically reading the manuscript and drawing my attention to a number of references. My thanks go also to Stephen Hawking for several enlightening conversations.