In the Beginning, At the End, and All in Between: Cosmological Aspects of Time

JOHN EARMAN, PITTSBURGH

1. Introduction

In recent decades, cosmology has transformed itself from a largely speculative enterprise, with often weak or non-existent observational constraints, into something approaching an empirical science. While there is still considerable room for speculation, the range and shape of possible answers to questions about large scale features of time—e.g., Is time infinite/finite in the past/future?—can be discussed with increasing specificity and with a growing hope that they can be settled in the foreseeable future. The purpose of this paper is to provide a framework for the discussion and to deliver some admittedly biased opinions about the likely outcomes of the discussion.

The paper is organized into three main sections. Section 2 deals with the discovery of the accelerating expansion of the universe and its implications for the fate of the universe. Section 3 takes up the question of whether time has a beginning. Section 4 deals with cosmological features of time in the middle ground. Conclusions are presented in Section 5.

2. A finite or infinite future?

2.1 Geometry as destiny

Not so long ago the discussion of the question that forms the title of this section was straightforward, at least in orthodox cosmology; and the answer was (to borrow a phrase from Krauss and Turner 1999) that geometry is destiny. Orthodox Big Bang cosmology described the large scale features of the universe using Friedmann-Robertson-Walker (FRW) models. The line-element for the spacetimes in these models can be written as

$$ds^2 = a(t)[\frac{dr^2}{1-kr^2} + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2] - dt^2 \tag{1}$$

where $a(t)$ is the scale factor (sometimes called the radius of the universe) and $k = 0, -1$, or $+1$, corresponding respectively to space sections of zero curvature, constant negative curvature, and constant positive curvature. The symmetries of the spacetime metric together with Einstein's gravitational field equations (EFE) force the stress-energy tensor to have the form that arises from a perfect fluid characterized by a density ρ and a pressure p. It was assumed that the matter fields that make up this fluid are normal in that they obey both the weak and strong energy conditions, which in the case of perfect fluid will be satisfied if $\rho \geq 0$ and there are no negative pressures larger than ρ.[1] The behavior of the scale is shown in Fig. 1, from which one sees that: if the universe is spatially open—the $k = 0$ or -1 cases—the universe expands forever, making time infinite in the future; and if the universe is spatially closed—the $k = +1$ case—the universe recollapses to a big crunch, making time finite in the future. Thus, the geometry of the spatial sections of the universe dictates its fate.
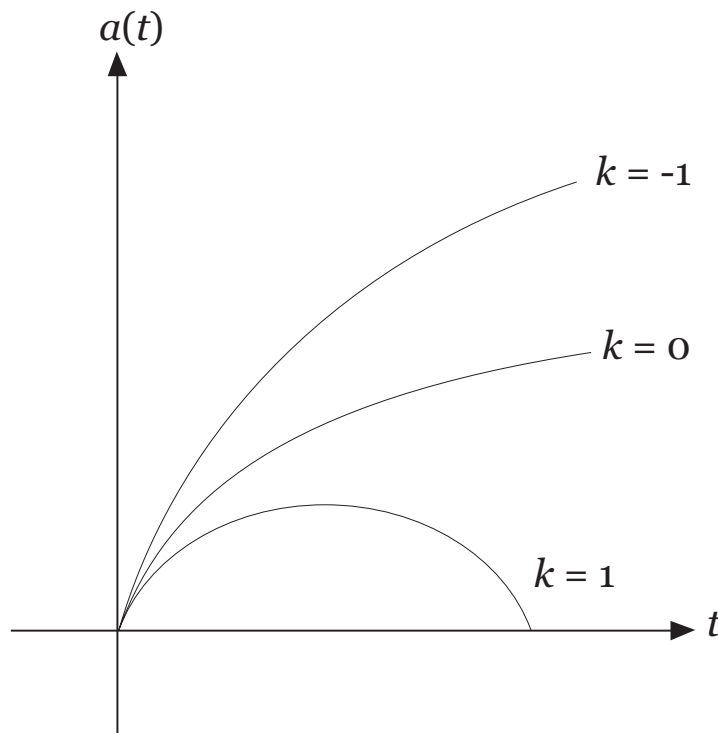

Figure 1: *Behavior of the scale factor in FRW models.*

---

1   The stress-energy tensor T_{ab} satisfies the *weak energy condition* iff $T_{ab}V^aV^b \geq 0$ for any timelike vector V^a; for a perfect fluid this will hold just in case $\rho \geq 0$ and $\rho + p \geq 0$. The *strong energy condition* (SEC) requires that $T_{ab}V^aV^b \geq \frac{1}{2}\text{Tr}(T_{ab})$ for any unit timelike V^a; for a perfect fluid this will hold just in case $\rho + 3p \geq 0$. See Wald (1984, pp. 219ff) for details.

The situation was even simpler in non-orthodox steady-state cosmology. As will be discussed in Section 3, steady-state cosmology describes the large scale structure of spacetime using (a portion of) de Sitter spacetime, which expands forever.

The simplest versions of inflationary cosmology modifies the standard hot Big Bang model by postulating that shortly after the Big Bang—say, at 10^{-35} sec—the universe passed through an era of accelerating expansion where $a_{final}/a_{initial} \sim e^{60}$. But since the universe is supposed to exit this inflationary era by, say, 10^{-33} sec and resume its more leisurely expansion rate, the slogan that geometry is destiny retains its validity.

This situation was profoundly changed by the recent discovery of "dark energy". Since this mysterious form of energy is a reincarnation of Einstein's cosmological constant, I will give a brief review of the history of this constant that refuses to die.

2.2 A very brief history of Λ

Einstein (1917) had two motives for introducing the cosmological constant: first, he wanted to harmonize his general theory of relativity (GTR) with Mach's principle; and second, since he believed that the actual universe is uniform and static at large scales, he needed to modify his original field equations in order to permit cosmological solutions with a uniform static distribution of matter. When Hubble's measurements of the red shifts of galaxies revealed that the universe is not static but is in fact expanding, Einstein abandoned the cosmological constant and would have nothing further to do with it. Reportedly, he called Λ his "biggest blunder".[2]

Einstein's orphan was adopted by two champions, Georges Lemaître and Arthur Stanley Eddington, who had rather different reasons for the adoption. Hubble's early reports of the present value H_o of the (non-constant!) Hubble constant $H := \dot{a}/a$ translated into an age of the universe of 2 billion years, which is less than the age of stars, as determined from theories of stellar evolution, and even less than the age of the earth, as determined by radioactive decay. Lemaître found that EFE with a positive cosmological constant permit homogeneous, isotropic cosmological models that expand from an initial singularity, then they enter a coasting phase in which the expansion rate slows down to a crawl, and finally they exit the coasting phase

---

2   This is not a direct quotation but a remark attributed to Einstein by Gamov (1958, pp.66–67).

and resume an expansion rate that can be fitted to Hubble's measurements. By tuning the parameters of the model, the coasting phase can be adjusted to make the age of the universe as large as desired.

Eddington's enthusiasm for Λ was based partly on a view of scientific explanation that anticipated by half a century one of the motivations for inflationary cosmology; namely, the notion that a good scientific explanation should not rely on special initial conditions. GTR can produce the observed recession of the galaxies without invoking the effective repulsive force of a positive Λ; but the large velocities of recession have to be, so to speak, built into the model from the beginning. Eddington opined that "This … can scarcely be called an *explanation* of the large velocities" (1933, p. 37).

Eddington's quasi-philosophical motivations for favoring a cosmological constant do not seem to have persuaded many of his scientific contemporaries. And the need for Lemaître's models seemed to fade as time wore on and Hubble's estimated value of H_o declined, raising the estimate of the age of the universe. In the late 1960s and early 1970s there was a brief revival of the Lemaître models as a means of explaining why the preponderance of quasistellar objects (QSOs) had redshifts near $z = 2$, a value that will fall out of the models if these objects were emitting during the coasting phase and the scale factor at that phase was ⅓ of its present value. However, the models predicted that there should be a rapid fall off of the intensity of QSOs when $z > 2$. When several bright QSOs for $z > 2.2$ were detected, Λ was put back on the shelf.

Still Λ refused to die. The 1980s saw the rise in the popularity of inflationary cosmology, which postulates an inflaton field that gives rise to an effective cosmological constant in the very early universe. The initial motivations for this postulation came from the perceived explanatory inadequacies of the standard hot Big Bang model in the guise of the "horizon problem" and the "flatness problem". But inflationary cosmology has scored a major empirical success in offering a natural explanation of the nearly scale-free spectrum of density perturbations revealed by measurements of the cosmic microwave background radiation. However, the inflationary paradigm seemed to have an Achilles heel: the simplest and most natural inflationary scenarios imply that the current value of the density parameter (defined below) should be near unity, whereas estimates from the measurements of the motion of galaxies and galactic clusters give an estimate of a contribution of matter of about 0.3. Thus, either much more complicated and artificial models of inflation are needed, or else the inflationary paradigm requires

the presence of a positive Λ or some surrogate for Λ to make up for the missing mass. The prayers of inflationary cosmologists were answered at the end of the 1990s.[3]

2.3 Lambda redux

In 1998 two independent teams of astrophysicists reported that observations of Type Ia supernovae indicated the rate of expansion of the universe is speeding up. To understand what this implies about the energy driving the accelerating expansion, it is helpful to work through some definitions. The deacceleration parameter q for the FRW models is given by

$$q := -\frac{a\,\ddot{a}}{\dot{a}^2} \tag{2}$$

The density parameters Ω^M and Ω^Λ associated respectively with matter and the cosmological constant are

$$\Omega^M := \frac{8\pi\rho}{3H^2}, \qquad \Omega^\Lambda := \frac{\Lambda}{3H^2} \tag{3}$$

where H is the Hubble constant introduced above. In FRW models

$$q = \frac{1}{2}\Omega^M(1 + 3w) - \Omega^\Lambda \tag{4}$$

where $w := p/\rho$ is the equation of state of matter. Since at the present time $\dot{a}_o > 0$, an accelerating expansion $\ddot{a}_o > 0$ means that $q_o < 0$. Equation (4) then implies that either there is a positive Λ or else a strange form of matter with $w < -\frac{1}{3}$, violating the strong energy condition (SEC).

To discuss the latter possibility in more detail, rewrite the first term on the rhs of (4) as the sum of two terms, one for normal matter ('N') satisfying the weak and strong energy conditions, and one for strange matter ('S') satisfying the weak energy condition but violating the SEC:

$$q = \frac{1}{2}[\Omega^{M_N}(1 + 3w_N) + \Omega^{M_S}(1 + 3w_S)] - \Omega^\Lambda \tag{5}$$

Since at the present time $w_N << 1$,

$$q_o \approx \frac{1}{2}[\Omega_o^{M_N} + \Omega_o^{M_S}(1 + 3w_S)] - \Omega^\Lambda \tag{6}$$

---

3 For a more detailed history of the cosmological constant, see Earman (2001).

So to get $q_o < 0$ requires either $\Lambda > 0$ or a strange form of matter-energy in sufficient quantity to dominate normal matter. The term "dark energy" has been coined to cover both possibilities.

Because of the radical nature of these conclusions it is important to note that they are supported by multiple independent lines of evidence, of which I will mention a few: X-ray observations of galactic clusters confirms the conclusion that $\ddot{a}_o > 0$; CMB-measurements give $\Omega_o^{tot} := \Omega_o^{M_N} + \Omega_o^{M_S} + \Omega_o^{\Lambda} \approx 1$, whereas observations of the dynamics of galaxies gives $\Omega_o^{M_N} \approx 0.3$, which together imply that 70% of matter-energy is in the form of dark energy; and measurements of the integrated Sacks-Wolf effect confirm that dark energy is a large fraction of the total.

2.4 Possibilities for explaining accelerating expansion

Going beyond orthodox GTR opens a Pandora's box of possibilities for explaining the accelerating expansion of the universe. For example, adding an $1/R$ term to the Hilbert action for GTR leads to field equations that allow accelerating expansion without dark energy (see Carroll, Duvvuri, Trodden, and Turner 2003). While this particular possibility appears to be ruled out by solar system tests (see Chiba 2003), there are no doubt more subtle modifications of GTR that give accelerating expansion without dark energy and that are in accord with other extant observational constraints. Going further afield to speculative theories of quantum gravity, an effective cosmological constant can result from the extra spatial dimension postulated by M-theory (see Gu and Hwang 2002); and the causal sets approach to quantum gravity leads to a fluctuating value for Λ (see Ahmed at al. 2004). I will not open this Pandora's box here but will confine myself to orthodox GTR.

If the possibilities are confined to those that can be described within orthodox GTR and if the FRW cosmologies are used to model the universe, then as explained above, the cause of the accelerating expansion can be identified as dark energy. But the second "if" necessitates a note of caution since the actual cosmos is not exactly homogeneous and isotropic (as is assumed in the FRW models), although suitable averaging over large volumes will produce a model with these features. Kolb et al. (2005) claim that "back reaction" effects of smoothing inhomogeneities over a volume on the order of the size of the current Hubble volume can produce accelerated expansion. Let g_{ab} and T_{ab} be respectively the actual, exact metric and stress-energy tensor for an inhomogeneous universe, and let $\bar{g}_{ab}$ and $\bar{T}_{ab}$ denote respectively the smoothed-out metric and the smoothed-out stress-tensor.

The Einstein field equations may not be satisfied by the pair consisting of the Einstein tensor $\bar{G}_{ab}$ (computed from $\bar{g}_{ab}$) and $\bar{T}_{ab}$. If not, the difference $\bar{G}_{ab} - \bar{T}_{ab}$ can be interpreted as an extra source term, which may violate the SEC even though the exact T_{ab} does satisfy it. There is no doubt that this phenomenon is mathematically possible (as was first noted by Ellis 1984). But care needs to be exercised in order to separate out real effects from gauge-dependent artifacts since the computation of the back reaction relies on the choice of gauge (see Ellis and Buchert 2005). Then it needs to be demonstrated that the magnitude of a gauge-independent back reaction effect can be great enough to explain the actually observed accelerated expansion while being consistent with other observational constraints. Ishibashi and Wald (2005) argue that it is implausible that such a demonstration will be forthcoming.

These issues can also be investigated by studying features of inhomogeneous cosmological models satisfying Einstein's field equations. Hirata and Seljak (2005) show that within this class of models, if the vorticity vanishes (as indicated by observations and as required by inflationary cosmology), the deacceleration parameter $\hat{q}$ cannot be negative unless the SEC is violated, where $\hat{q}$ (which is now a function of spatial position as well as time) is defined in terms of the time rate of change of the local expansion of matter. It could be objected that what matters to the supernovae Type Ia observations is not the deacceleration parameter $\hat{q}$ but rather the deacceleration parameter $\tilde{q}$ that appears in the luminosity distance(d_L)-redshift (z) relation. Expanding d_L in a Taylor series in z gives $d_L = \dfrac{z}{\tilde{H}} + \dfrac{1-\tilde{q}}{2\tilde{H}} z^2 + O(z^3)$, where $\tilde{q}$ and the Hubble parameter $\tilde{H}$ are now functions of the direction of observation as well as of the spacetime location of the observer. While $\tilde{q}$ can be negative without a violation of the SEC, Hirata and Seljak (2005) argue that $\tilde{q} < 0$ requires anisotropic expansion and, thus, that if there are accelerating directions ($\tilde{q} < 0$) then there are also deaccelerating directions ($\tilde{q} > 0$). There is no observational support for the latter. Furthermore, it is a mystery how such a variation in $\tilde{q}$ can be reconciled with the observed isotropy of the CMB.

While the controversy over whether dark energy is needed to explain accelerating expansion is apt to continue for the foreseeable future, at this juncture it seems that the weight of the evidence strongly favors the presumption that dark energy is the cause, and I will discuss the fate of the universe under this presumption.

2.5 The fate of the universe

Three categories of dark energy can be distinguished in terms of the equation of state: a true cosmological constant corresponds to the case $w_\Lambda = -1$; so-called *quintessence* corresponds to the case $-1 < w_Q < -\frac{1}{3}$; and so-called *phantom matter* corresponds to the case $w_P < -1$. The implications for the fate of the universe can be discussed by noting that if the equation of state w_X does not change with time, then the matter-energy density ρ_X scales as $a^{-3(1+w_X)}$. The major possibilities can be divided as follows.

1) A true Λ is entirely responsible for $\ddot{a}_o > 0$. Then, regardless of whether $k = 0, -1$, or $+1$, the universe will expand forever. However, since the scale factor increases faster than the horizon distance, an observer comoving with the expansion will see a universe that becomes increasingly empty, dark, and cold. In sum, we get Eternity with a Big Chill. It is dubious that critters such as ourselves can survive the Big Chill to experience Eternity (see Krauss and Starkman 2000).

2) $\Lambda = 0$ and quintessence is entirely responsible for $\ddot{a}_o > 0$. Then if the equation of state w_Q does not change we again get Eternity with a Big Chill. If w_Q does change as the universe expands, then the fate of the universe has to be discussed in terms of particular models for quintessence.

3) $\Lambda = 0$ and phantom matter is entirely responsible for $\ddot{a}_o > 0$. If the equation of state w_P does not change, $\rho_P(t) \to +\infty$ in a finite time, and the universe ends in a Big Smash. Before the end, there is a Big Rip in which gravitationally bound systems are ripped apart (see Caldwell et al. 2002; Caldwell et al. 2003). Again if w_P does change as the universe expands, then the fate of the universe has to be discussed in terms of particular models for phantom matter. It is far from clear how seriously to take phantom matter since it violates not only the SEC but it also the dominant energy condition[4], leading to the possibility of acausal propagation and making it difficult to build stable models in QFT (see Carroll, Hoffman, and Trodden 2003).

---

4   The stress-energy tensor T_{ab} satisfies the *dominant energy condition* iff for any future directed timelike vector V^a, $-T_{ab}V^a$ is a future directed timelike or null vector. This condition conjoined with the conservation of T_{ab} (i.e. $\nabla^a T_{ab} = 0$, which is entailed by EFE) implies that matter-energy is not transmitted faster than light: if T_{ab} vanishes in some achronal spacetime region, then it vanishes throughout the domain of dependence $D(S)$, which consists of all spacetime points p such that every endless non-spacelike curve that passes through p meets $D(S)$. See Hawking and Ellis (1973, Lemma.4.3.1).

4) There are many mix-and-match possibilities. For example, suppose that quintessence is entirely responsible for $\ddot{a}_o > 0$ and that there is small negative Λ, so small that it cannot be detected by current observations. Then as the universe expands, the attractive force due to the negative Λ will eventually dominate the repulsive force of quintessence, and the universe will begin to contract. On the other hand, a negative Λ will never come to dominate phantom matter and prevent a Big Rip followed by a Big Smash.

A major challenge for observational and theoretical astrophysics is to sort through these possibilities to determine the equation of state of dark energy and, thus, the fate of the universe.

3. Does time have a beginning?

Specifying what it means for a general relativistic spacetime to be finite in the past turns out to be a more difficult task than might be imagined. To illustrate the point, one might take the following as a necessary and sufficient condition for such finiteness: for any global time slice S (i.e. connected spacelike hypersurface without edges)—which may be thought of as representing "now"—any past directed timelike curve with future endpoint on S and no past endpoint has a proper length less than or equal $b(S) < \infty$. This condition suffers from multiple deficiencies. First, there are spacetimes, such as Gödel spacetime, which do not intuitively count as being temporally finite in the past but which vacuously satisfy the condition because they possess no global time slices. Second, unless restrictions are put on the choice of S the condition in question would yield incorrect results. For example, with (x, t) an inertial coordinate system for $(1 + 1)$-dim Minkowski spacetime, delete all the points on or below $t = 0$. This spacetime should count as finite in the past direction, but the condition in question fails if S is given by $t = \sqrt{1+x^2}$. Third, the condition wrongly counts the surgically mutilated Minkowski spacetime shown in Fig. 2 as being temporally finite in the past.
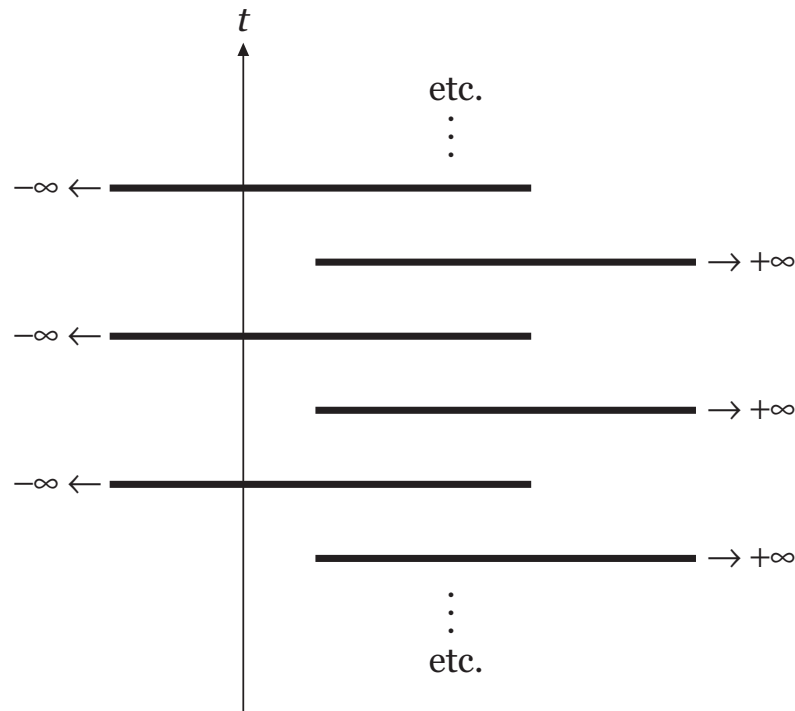
Figure 2: *(1 + 1)-dim Minkowski spacetime with spacelike strips (heavy lines) removed. Light cones are at 45°. Every timelike curve has finite proper length.*

Coming up with a more adequate analysis that is applicable to any general relativistic spacetime is a ticklish problem. But the problem need not be tackled here since the cosmological models under consideration belong to a special class of models having features that make it obvious how to decide whether time in a model in the class is finite or infinite in the past/future. All of these models have a global time function t whose level surfaces $t = const$ are Cauchy surfaces (i.e. each is intersected exactly once by every timelike curve without endpoint). Moreover, level surfaces $t = const$ are orthogonal to a congruence of timelike geodesics, and the lengths of geodesic segments between any two levels are all the same; these geodesics are the longest timelike curves between the levels. Such a t induces a natural metric of time, and one can say that time is finite in the past (respectively, the future) iff the range of t is bounded from below (respectively, the future). The class of FRW models is a subclass of the class whose members admit a natural time metric, and the cosmic time t of the line element (1) has the properties enumerated above. If this t is bounded (from above or below), then the scale factor $a(t)$ approaches 0 as the t approaches the bound. For the standard FRW Big Bang model that best fits the actual universe, the

lower bound is around 14 billion years. The bound rises by only a few seconds for the simple inflationary models that insert an inflationary era after the Big Bang. Thus, assuming the validity of the models, the actual universe had a beginning, not in the sense that there is a first instant of time, but in the sense that time in the actual universe is finite in the past.

Within orthodox classical GTR speculation about what happened before the Big Bang is physically meaningless since there is no way to extend the FRW models as solutions to EFE even in the distributional sense of solution. Of course, this didn't stop some cosmologists from speculating. For example, R. C. Tolman (1934) speculated about oscillating models in which a spatially closed universe cycles through expanding and contracting phases. He admitted that "our differential equations for the motion of the model are not sufficient to describe the mechanism of passage through the lower limit of contraction" (1934, p. 428). Nevertheless, he thought that "the existence of which [i.e. the passage through $a = 0$] is physically inevitably necessary" (ibid.). His confidence was based on the notion that the initial and final singularities were artifacts of the idealizations of the models. The singularity theorems of Hawking and Penrose undermine this confidence by showing that past timelike or null geodesic incompleteness is to be expected under quite generic conditions, at least if energy conditions discussed in the preceding section are assumed to hold. It is interesting to note that Tolman thought that in his $\Lambda = 0$ oscillating model the entropy of successive cycles would increase. Reading this backwards in time, Steinhardt and Turok (2002, p. 126003–1) argue that the lengths of the cycles decreases sufficiently rapidly in the past direction that the sum of the cycles lengths is finite, giving a beginning of time.

Steady state cosmology, despite the name and despite the fact that it does not involve a Big Bang[5], does imply that time is finite in the past. This can be seen from the fact that the line element used in steady state cosmology is just the de Sitter metric written in a coordinate system that is adapted to flat space slices of de Sitter spacetime. These flat slices cover only a portion of the spacetime, a portion that is timelike geodesically incomplete in the past (see Fig. 3).

---

5   This phrase was invented by the steady state cosmologists as a term of derision.
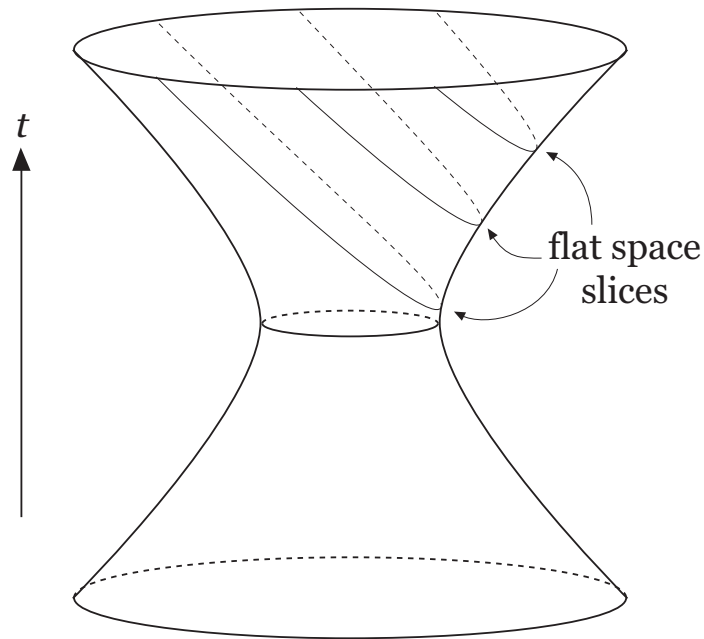
Figure 3: *De Sitter spacetime as an hyperboloid embedded in a flat space. The upper portion can be covered by a family of flat spacelike slices. This portion is past timelike geodesically incomplete.*

In addition to the simple inflationary models referred to above, there are also more elaborate models that entail "eternal inflation" in the future direction. The picture presented by the latter is that of thermalized regions set in an inflating background. As time goes on, the thermalized regions expand and new ones come into existence; but since the background expands so much faster than the thermalized regions, the universe is never completely thermalized and inflation continues forever. Could it be that inflation is also eternal in the past direction and that the initial Big Bang singularity and, indeed, all cosmological singularities in the past are avoided? Some hope for a positive answer comes from the fact that some of the Hawking-Penrose singularity theorems use the SEC, which is violated by inflationary mechanisms. However, a negative answer has been claimed by inflationary cosmologists who have extended the Hawking-Penrose theorems by proving past timelike or null geodesic incompleteness using only the weak energy condition or no energy condition at all (see Borde and Vilenkin 1994, 1997; Borde, Guth, and Vilenkin 2003).

On the other side, Ellis and Maartens (2002, 2004) have argued that these theorems are based on overly restrictive assumptions, in particular that $k = 0$ or -1 and/or that $H > 0$ for all past times. As for the first assump-

tion, it is not excluded by inflation that $k = +1$ even though inflation does drive k/a^2 towards 0 and the density parameter towards 1. And in any case, $k = +1$ or $\Omega_o^{tot} > 0$ is not excluded by present observations which give $\Omega_o^{tot} = 1.02 \pm 0.02$. Ellis and Maartens show that if $k = +1$, eternal inflation to the past can be achieved in that, without having to invoke negative energy densities, there is no past singularity. They explicitly exhibit three models of this type, whose scale-factor behavior is illustrated in Fig. 4. In two of these models the scale factor has a minimum value $a_* > 0$, which can be set large enough that Planck-scale quantum gravity effects are avoided. Just how physically plausible these models are remains to be seen since some of them seem to require fine tuning of initial conditions. And in any case these models do not conform to the most recent astronomical observations that indicate that before entering the present accelerating phase, the universe was expanding ($\dot{a}>0$) but deaccelerating ($\ddot{a}<0$).
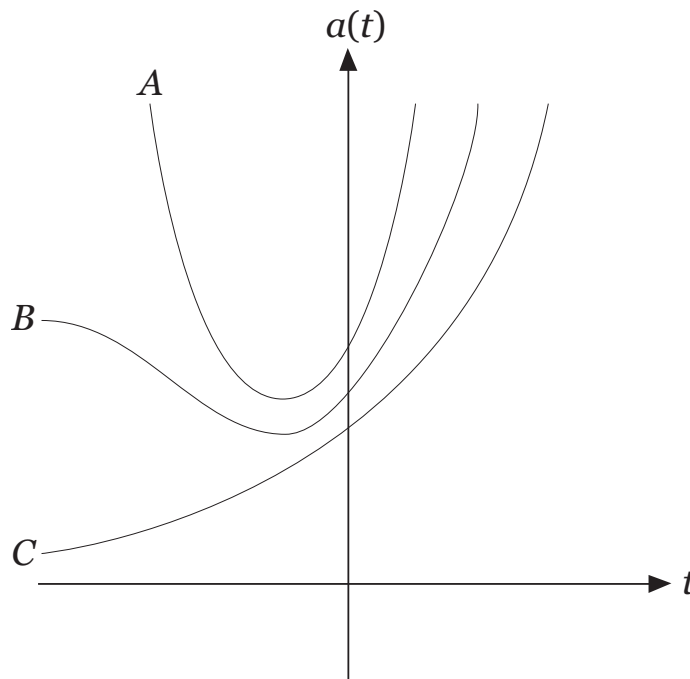

Figure 4: *The behavior of the scale factors in three models of eternal past inflation: Phoenix (A), Awakening (B), Emergent (C). From Ellis and Maartens (2002).*

More interesting, but also more speculative, are the implications of quantum gravity for the beginning of time. For example, the extra spatial dimensions postulated by string theory or M-theory may permit a smooth continuation of time through what on orthodox GTR is seen as an impass-

able initial spacetime singularity (see Khoury et al. 2001, 2002; Steinhardt and Turok 2002). I will concentrate here on the implications of the other main line of research on quantum gravity, referred to as loop quantum gravity (LQG), which attempts to quantize the gravitational field by applying the canonical quantization program to classical GTR.

In LQG version of the FRW models, there is a self-adjoint operator corresponding to the classical quantity $1/a$ and acting on the Hilbert space of spatially homogeneous, isotropic quantum kinematical states. It is found that the spectrum of this operator is bounded from above, giving a first indication that the classical singularity has been removed (see Bojowald 2001 and Ashtekar et al. 2003).[6] A complete proof of removal would require that the quantum dynamics gives an unambiguous evolution through the classical singularity. And it is just here that we meet a seemingly paradoxical implication of the approach of LQG. Written in Hamiltonian form, Einstein's field equations comprise a constrained Hamiltonian system. For such a system a subset of these constraints, called the first class constraints, generate gauge transformations—that is, transformations that connect two state descriptions that are regarded as descriptions of the same physical situation. In the case of GTR one of the first class constraints is the Hamiltonian constraint. Since this constraint generates the motion it appears that motion is pure gauge, with the changing state descriptions not describing any genuine change of physical state but just redescriptions of the same physical state. I will not pause to comment on this "frozen time" picture but simply push on to summarize the rest of the story.[7]

In LQG the "dynamics" is obtained by solving the Hamiltonian constraint equation, which restricts the physically allowed states. For the case at issue this constraint equation comes in the form of a difference equation rather than a differential equation. If the scale factor a is regarded as a "clock variable", then the constraint equation provides a "time evolution" of the quantum state through discrete steps of the clock variable. The key point is that this evolution equation does determine a unique continuation through the classical singularity, at least in the $k = 0$ case (but see Green and Unruh (2004) where it is shown that in a spatially closed FRW model,

6   The issue of singularity resolution in loop quantum gravity turns out to be much more complicated than indicated above; see Bunneman, J. and Thiemann, T. 2006 "On (cosmological) singularity avoidance in loop quantum gravity". *Classical and Quantum Gravity* 23: 1395–1427.
7   See Isham (1992) and Kuchař (1992) for a discussion of this and related issues.

the use of the scale factor as a "clock variable" is problematic). Note, however, that the continuation involves the passage through the Planck regime where, according to LQG, the spacetime of classical GTR is not valid even as an approximation. Thus, although the continuation through the classical singularity tempts one to say that LQG justifies speaking of what happens prior to the Big Bang, "prior to" here does not refer to a time order in the sense of classical special and general relativity. The verdict of LQG is that there is a beginning for classical relativistic time in our universe (insofar as our universe is described by the LQG quantization of a an FRW model).

4. In the meantime

4.1 Asymmetries of time and asymmetries in time[8]

Many philosophers and physicists subscribe to the notion that temporal asymmetries of physical processes and perhaps also asymmetries of time itself are grounded in entropic behavior. This viewpoint must cope with twin puzzles which Boltzmann was forced to confront but which he did not successfully resolve. Grant that the microdynamics of the physical processes we encounter are such that if a system is in a state of low Boltzmann entropy $S_B(t)$ at t, then it is very likely that at $t + \Delta$, for some suitable $\Delta > 0$, $S_B(t + \Delta) > S_B(t)$. But what is the justification for the posit of initially low entropy? And even if this posit is granted, what justifies the normal expectation that at $t - \Delta$, $S_B(t - \Delta) < S_B(t)$? The latter question becomes pressing when it is realized that if the laws governing the microdynamics are time reversal invariant, then the very apparatus that leads to the prediction that $S_B(t + \Delta) > S_B(t)$ also leads to the prediction that $S_B(t - \Delta) > S_B(t)$. Thus, if we are told that a thermally isolated system consists at t of an ice cube in a glass of lukewarm water, we would normally predict that a few minutes after t the ice cube will have partly melted and the temperature of the water will have decreased, and also that a few minutes prior to t the ice cube would have been less melted and the temperature of the water would have been higher. But according to what was just said, the Boltzmann apparatus does not underwrite this asymmetry of inference.

The current dogma is that modern cosmology comes to Boltzmann's rescue in that the answers to the initial state puzzle and the asymmetry puz-

---

8   This section summarizes some of the conclusions of Earman (2006), to which the interested reader is referred for details.

zle lie in the fact that the very early universe was in an extraordinarily low entropy state. Here "very early" means shortly after the Big Bang, if one is using a standard hot Big Bang model, or shortly after the universe reheats, if one is using an inflationary model. It might seem counterintuitive to say that a thermalized state of a homogeneous and isotropic universe has low entropy, but (the story goes) intuitions are misleading because they neglect the gravitational contribution to entropy—since gravity tends to clump matter, a smooth state is very improbable.[9]

I have two objections to this dogma. The first is that it is very likely that it is (to echo Pauli) not even false. The Boltzmann entropy for a deterministic dynamical system is defined by choosing a coarse graining and then setting the value of the entropy for a coarse-grained state to be proportional to the log of the measure $\mu(V)$ of the volume V of the micro-state space compatible with the coarse-grained state in question, where μ is a normed measure on the micro-state space that is invariant under the deterministic flow. In the cosmological setting, however, it is doubtful that this formalism yields coherent results. An intimation of the difficulties can be gleaned from features of the model that Hawking and Page (1988) used to try to answer the question of how probable it is that inflation initiates. They investigated a FRW-ϕ model in which the matter content is given by a minimally coupled massive scalar field ϕ. The state space X for this system is four-dimensional and can be coordinatized using the scale factor a of the FRW model, the field ϕ, and their respective conjugate momenta p_a and p_ϕ. As the discussion from the previous section would lead one to expect, the equations of motion have the form of a constrained Hamiltonian system with the one and only constraint being the vanishing of the Hamiltonian $\mathcal{H} = 0$. The three-dimensional subspace $\mathcal{C} \subset X$ where the constraint is satisfied is called the *constraint surface*. A reduced phase space free of gauge redundancy can be formed by choosing a two-dimensional surface Σ that is transverse to the dynamical trajectories on $\mathcal{C}$. Then the pullback of the (degenerate) symplectic form $\omega = dp_a \wedge da + dp_\phi \wedge d\phi$ defines a volume measure $^{(2)}\mu$ on Σ that is invariant under dynamical evolution.[10]

---

9   This idea has been promulgated by Penrose (1979, 1989, 2004). It is accepted by physicists who worry about the foundations of statistical mechanics, e.g. Lebowitz (1993, 1999) and by philosophers of science, e.g. Price (2004).

10 Alternatively, as noted by Hollands and Wald (2002), an invariant volume measure $^{(3)}\mu$ can be defined on $\mathcal{C}$ by the volume element $^{(3)}\epsilon$ given by the condition $d\mathcal{H} \wedge {}^{(3)}\epsilon = {}^{(4)}\epsilon$, where $^{(4)}\epsilon := dp_a \wedge da \wedge dp_\phi \wedge d\phi$ is the Liouville volume element for X.

Now a crucial fact is that $^{(2)}\mu(\Sigma)$ does not normalize.[11] Since this result holds for the $k = +1$ FRW model, the non-normalizability cannot be blamed on the infinity of space implied by the $k = 0$ and $k = -1$ models. One could still try to assign finite measures to coarse grained states as follows. Let $M \subseteq \Sigma$ be the region corresponding to a coarse grained state m. Set $^{(2)}\Pr(m) = 0$ in case $^{(2)}\mu(M) < \infty$; set $^{(2)}\Pr(m) = 1$ in case $^{(2)}\mu(M) = \infty$ and $^{(2)}\mu(\Sigma - M) < \infty$; and set declare $^{(2)}\Pr(m)$ undefined otherwise. Then it is easy to show that $^{(2)}\Pr$ is a finitely additive, partial probability measure on the coarse graining. Hawking and Page found that the third case of ill-definedness holds for the probability of inflation. And one should be prepared to find that $^{(2)}\Pr(m_i)$ is ill-defined for the contemplated smooth, thermalized initial coarse grained state m_i of the universe. Even if $^{(2)}\Pr(m_i)$ is well-defined, the best that one can hope for as an explication of the notion that the universe begins with a low entropy state is that $^{(2)}\Pr(m_i) = 0$, which makes conditional probabilities $^{(2)}\Pr(m_l / m_i)$ of later coarse-grained states m_l ill-defined if conditional probability is given its usual interpretation.

My second objection is that even if there were an appropriate normalized measure and even if the Boltzmann entropy of the very early universe is well-defined and has a low value, it seems dubious that the latter explains the thermodynamic asymmetries of the kinds of systems typically encountered by critters like us. It is, of course, a truism that if some asymmetry does not follow from the laws of physics, then its origins must be sought in initial/boundary conditions. But it is a fallacy to reason: The thermodynamic asymmetries we observe today are traceable to the conditions of the early universe; the early universe was in a low entropy state; therefore, the low entropy of the early universe is the key ingredient in explaining the thermodynamic asymmetries we observe today. I will mention two among many reasons for thinking that there is a gap in the argument. First, Boltzmann entropy is not a spatially localized quantity but rather a global property of the system—in this case the system consisting of the entire universe—and the value of this entropy places only weak constraints on the entropy of the small subsystems of the kind we are interested in (e.g. the ice cube in the glass of water), and this is especially so since these subsystems are not dynamically isolated. Second, if the argument worked, it should work even better if the entropy of the early universe were even lower than it is in the actual universe because, for example, the state is completely homogeneous

---

11 The same holds for the $^{(3)}\mu(\Sigma)$ of Hollands and Wald (2002).

and isotropic. But in this case the galaxies and solar systems would not have formed and the universe would not develop localized systems that are in thermal disequilibrium and that display the familiar thermodynamic arrow of time.

In sum, I think that, despite the rosy pronouncements to the contrary, we are far from an understanding the role that cosmology plays in the local physics of thermal systems.

4.2 Black holes and black hole evaporation

The FRW models describe only the large scale structure of the universe—the homogeneity and isotropy assumed by the models appears only at scales above 10 mega parsecs. At smaller scales many interesting and disturbing things can happen—in particular, the formation and evaporation of black holes.

Classical GTR predicts the formation of black holes through the process of gravitational collapse of stars of up to 10 $M_\odot$. In addition, the cores of galaxies can collapse to form black holes with masses ranging from millions to billions of solar masses. There is very strong evidence that our universe is well populated by both solar mass and supermassive black holes. It is also possible that tiny black holes with masses of the order of $10^{-19}\,M_\odot$ could have formed in the early universe. There is currently no observational confirmation of these primordial black holes, but if they do exist they would provide a test for predictions about black hole evaporation described below.

An observer who falls through the horizon of a black hole has his fate sealed: twist and turn how he will, his world line can be extended for only a finite amount of proper time (see Fig. 5)[12], and even before the theoretical bound on his time expires, he will expire by being torn apart by the tidal forces of the black hole. Thus, even though time itself may be infinite, the fact that the universe is well populated with black holes means that even observers idealized to escape the normal effects of aging can have an in-principle bound to their future existence.

---

12 The standard conventions of conformal diagrams are in effect, e.g. light cones are at 45°. $\mathfrak{J}^+$ denotes future null infinity, the terminus of outgoing light rays; and ι° denotes spatial infinity.
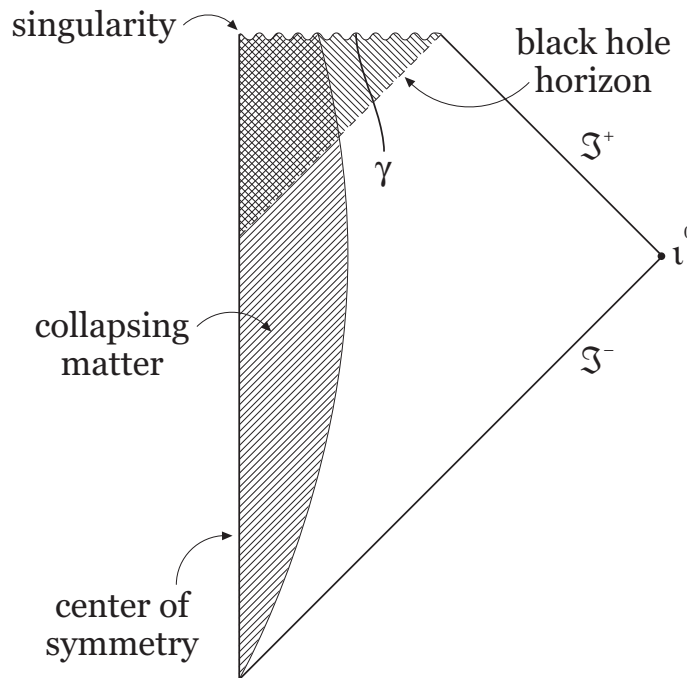
Figure 5: *Conformal diagram of black hole formation in spherically symmetric gravitational collapse. γ is the worldline of an observer who falls through the black hole horizon.*

Suppose that the absence of evidence for white holes—the time reverses of black holes—is evidence for their absence. Then the combination of this evidence with the evidence for the existence of black holes would be evidence for a pervasive temporal asymmetry. It is currently a matter of speculation as to whether this black hole-white hole asymmetry is purely *de facto* or whether it is indicative of some deep lawlike asymmetry.[13]

The discussion to this juncture was limited to classical GTR. But now we must consider quantum effects. Stephen Hawking showed that in the presence of a black hole a quantum field will thermalize so that, effectively, a black hole radiates with a thermal spectrum. This prediction of black hole radiation does not depend on any nascent quantum theory of gravity but simply uses quantum field theoretical calculations for a quantum field propagating on a spacetime background supplied by classical GTR. The next step does take a step towards quantum gravity by calculating the back reaction of the radiation on the spacetime geometry of the black hole. The prediction is that the black hole will lose mass and its horizon area will shrink. Presumably these semi-classical calculations will break down when the Planck

13   Penrose (2004, Sec. 30.9) takes the latter alternative.

regime is reached. But if the black hole does evaporate completely and if the result can be described to good approximation in classical GTR, then the upshot will be something like that pictured in Fig. 6.[14] The time required for a solar mass black hole to evaporate is very long indeed—on the order of 10^{67} years; but if the universe expands forever, there is no shortage of time.
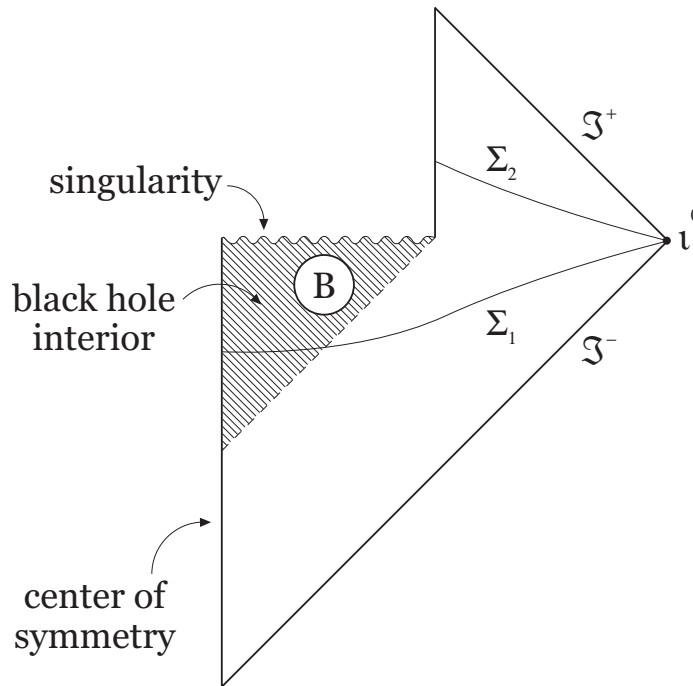

Figure 6: *Conformal diagram of conventional treatment of black hole evaporation.*

An observer who falls through the horizon of an evaporating black hole is still doomed. But two even more ominous things emerge. First, the evaporation results in a naked singularity, i.e. a singularity visible from future null infinity $\mathfrak{I}^+$. Second, a quantum field propagating on this spacetime undergoes a transition from a pre-evaporation pure state to a mixed post-evaporation state, at least assuming that the pre-evaporation state was pure. (Consider the subalgebra of observables $\mathcal{A}(\Sigma_2)$ associated with a thin neighborhood of the post-evaporation time slice Σ_2 of Fig. 6. Correlations between observables belonging to $\mathcal{A}(\Sigma_2)$ and observables belonging to the subalgebra $\mathcal{A}(B)$ associated with a region B in the interior of the black hole are established by the presence of a common cause in the form of the Hawking radiation. Since B and a thin sandwich of Σ_2 are relatively spacelike, $[\mathcal{A}(\Sigma_2), \mathcal{A}(B)] = 0$.

---

14 See Wald (1994, Ch. 7) for an overview.

Consider a state ω on the global algebra of observables that encodes the correlations between $\mathcal{A}(\Sigma_2)$ and $\mathcal{A}(B)$. It follows that the restriction of ω to $\mathcal{A}(\Sigma_2)$ is a mixed state.) Such a pure-to-mixed transition is necessarily non-unitary, which is the precise content of what is called the Hawking "information loss paradox".[15] So desperate are some commentators to resolve this paradox that they are prepared to believe six contradictory things before lunch, such as "black hole complementarity". Part of the desperation is based on the misguided notion that loss of unitarity is a disaster that must be avoided at all costs. On the contrary, unitarity is not essential to conservation of probability, and a respectable amount of QFT can be done without it.[16] One result of coming to terms with the pure-to-mixed state transition is that black hole evaporation involves a violation of time reversal invariance and, thus, is a possible source for time's asymmetries.[17]

It must be emphasized, however, that the preceding discussion is based on semi-classical quantum gravity and that a full theory of quantum gravity may present quite a different picture of black hole evaporation. And in fact, the proponents of LQG have recently claimed that on the basis of the resolution of black hole singularities in LQG, it is reasonable to think that the depiction of black hole evaporation in Fig. 6 will be replaced by something more like of Fig. 7 (see Asktekar and Bojowald 2005). It is argued that, analogously to the FRW case, the quantum evolution continues through the classical singularity and that a pure state remains pure and, in this sense, no information is lost.[18] Note, however, that as in the FRW case, the classical singularity is not resolved in the sense of being replaced by smooth classical relativistic spacetime structure; rather it is replaced by quantum spin-foam where a description in terms of classical relativistic spacetime does not have even approximate validity. It is but cold comfort to tell an observer who falls into a black hole that, within a finite time, he will dissolve into quantum spin-foam rather than splatting on a classical spacetime singularity. On the other hand, it is real comfort to tell an external observer that the strange things he is seeing are the results of quantum spin-foam rather than a naked singularity of classical GTR; for the latter is a source of genuine unpredict-

---

15 See Wald (1994, pp. 178–181) and Belot et al. (1999).

16 See Wald (1994, pp. 181–183) for a defense of this viewpoint.

17 See Wald (1980) and Earman (2002).

18 As in the FRW case, "quantum evolution" means that some suitable clock variable is chosen and that the Hamiltonian constraint equation is used the track the quantum state through discrete steps of the clock variable.

ability since classical GTR places no restrictions on the pathologies that can devolve from such singularities, whereas the former is subject to the laws of LQG.
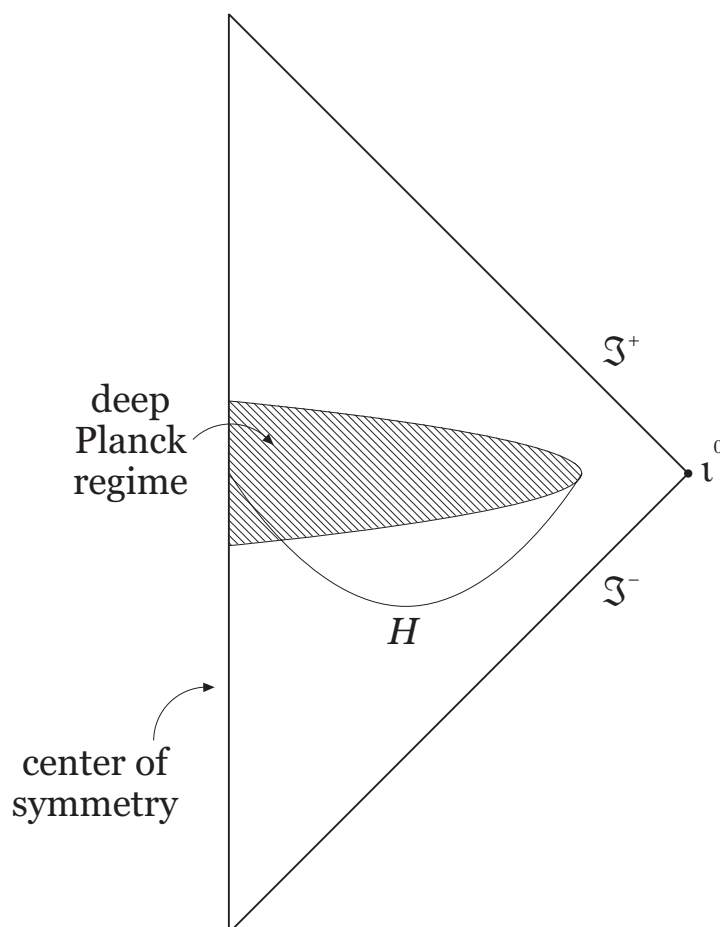

Figure 7: *Loop quantum gravity paradigm for black hole evaporation.* H *is the dynamical horizon. From Ashtekar and Bojowald (2005).*

5 . Conclusion

For what it is worth, I list my bets regarding cosmological aspects of time.

1. I would offer better than even odds that time is infinite in the future because the universe will expand forever; but I would also bet that no one will be around to collect the stakes since critters like us will perish in the Big Chill.
2. I would bet that time in the sense of classical special and general relativity does have a beginning because classical spacetime dissolves into quantum foam in the finite past. I would not be surprised to learn—but would not

bet—that the correct quantum theory of gravity resolves the initial Big Bang singularity. But until the shape of this sought-after theory is more clearly discerned, I would decline bets about the sense in which the resolution of the Big Bang singularity legitimates talk of what happened "before" the Big Bang.

3. I would offer high odds that the explanation of the local temporal asymmetries of concern to us is not to be found in a low entropy state for the very early universe.

4. I would bet that our universe is populated with black holes and that these objects are not really black but radiate with a thermal spectrum, as Hawking predicted. But I would try to beg off bets on the upshot of black hole evaporation, except insofar as I could bet against such expediencies as black hole complementarity.

Bets aside, it is satisfying to reflect on how much we have learned over the past few decades about cosmological aspects of time. It is humbling to realize how much we still don't know. And it is inspiring to learn that we can discern pathways that will lead—we can hope—to a resolution of our ignorance.

References

Ahmed, M., Dodelson, S., Greene, P.B., and Sorkin, R. 2004 "Everpresent Λ", *Physical Review D* 69: 103523-1-8.

Ashtekar, A. and Bojowald, M. 2005 "Black Hole Evaporation: A Paradigm", gr-qc/0504029.

Ashtekar, A., Bojowald, M., and Lewandowski, J. 2003 "Mathematical Structure of Loop Quantum Cosmology", *Advances in Theoretical and Mathematical Physics* 7: 233–268; gr-qc/0304074.

Belot, G., Earman, J., and Ruetsche, L. 1999 "The Hawking Information Loss Paradox: The Anatomy of a Controversy", *British Journal for the Philosophy of Science* 50: 189–229.

Bojowald, M. 2001 "Absence of a Singularity in Loop Quantum Cosmology", *Physical Review Letters* 86: 5227–5230.

Borde, A. and Velinkin, A. 1994 "Eternal Inflation and the Initial Singularity", *Physical Review Letters* 72: 3305–3308.

— 1997 "Violation of the weak energy condition in inflating spacetimes", *Physical Review D* 56: 717–723.

Borde, A., Guth, A., and Vilenkin, A. 2003 "Inflationary spacetimes are not past-complete", gr-qc/0110012 v2.

Caldwell, R.R. 2002 "A Phantom Menace? Cosmological consequences of a dark energy component with super-negative equation of state", *Physics Letters B* 545: 23–29.

Caldwell, R.R., Kamionkowski, M., and Weinberg, N.N. 2003 "Phantom Energy and Cosmic Doomsday", astro-ph/0302506.

Carroll, S.M., Duvvuri, V., Trodden, M., and Turner, M.S. 2003 "Is Cosmic Speed-Up Due to New Gravitational Physics?" astro-ph/0306438.

Carroll, S.M., Hoffmann, and Trodden, M. 2003 "Can the Dark Energy Equation-of-State Parameter w be less than -1?" astro-ph/0301273.

Chiba, T. 2003 "1/R Gravity and Scalar-Tensor Gravity", astro-ph/0307338.

Earman, J. 2001 "Lambda: The Constant That Refuses to Die", *Archive for History of Exact Sciences* 55: 189–220.

— 2002 "What Time Reversal Invariance Is and Why It Matters", *International Journal for the Philosophy of Science* 16: 245–264.

— 2006 "Note Even False: The 'Past Hypothesis'", *Studies in History and Philosophy of Science*, in press.

Eddington, A.S. 1933 *The Expanding Universe*. New York: Macmillan.

Einstein, A. 1917 "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*, 142–152. English translation as "Cosmological Considerations on the General Theory of Relativity", in W. Perrett and G. B. Jeffrey (eds.), *The Principle of Relativity*, 177–188. New York: Dover Books, 1952.

Ellis, G.F.R. 1984 "Relativistic Cosmology", in B. Bertotti, F. de Felice, and A. Pascolini (eds.), *General Relativity and Gravitation*, 215–288. Dordrecht: D. Reidel.

Ellis, G.F.R. and Buchert, T. 2005 "The universe at different scales", gr-qc/0506106.

Ellis, G.F.R. and Maartens, R. 2002 "Eternal inflation without quantum gravity", gr-qc/0211082.

Ellis, G.F.R. and Maartens, R. 2004 "The emergent universe: inflationary cosmology with no singularity", *Classical and Quantum Gravity* 21: 223–232.

Gamov, G. 1958 "The Evolutionary Universe", in *The Universe: A Scientific American Book*. London: G. Bell and Sons.

Green, D. and Unruh, W.G. 2004 "Difficulties with Re-collapsing Models in Closed Isotropic Loop Quantum Cosmology", *Physical Review D* 70: 103501-1-7.

Gu, Je-An and Hwang, W-Y.P. 2002 "Accelerating Universe as From the Evolution of Extra Dimensions", *Physical Review D* 66: 024003-1-6.

Hawking, S.W. and Ellis, G.F.R. 1973 *The Large Scale Structure of Space-Time*. Cambridge: Cambridge University Press.

Hawking, S.W. and Page, D.N. 1988 "How Probable Is Inflation?" *Nuclear Physics B* 298: 789–809.

Hirata C.M. and Seljak, U. 2005 "Can surperhorizon cosmological perturbations explain the acceleration of the universe?" astro-ph/0503582.

Hollands, S. and Wald, R.M. 2002 "Comment on Inflation and Alternative Cosmology", hep-th/0210001.

Ishibashi, A. and Wald, R.M. 2005 "Can the Acceleration of Our Universe Be Explained by the Effects of Inhomogeneities?" gr-qc/0509108.

Isham, C.J. 1992 "Canonical Quantum Gravity and the Problem of Time", in L.A. Ibot and M.A. Rodríguez (eds.), *Integrable Systems, Quantum Groups, and Quantum Field Theories*, 157–287. Boston: Kluwer Academic.

Khoury, J., Ovrut, B.A., Steinhardt, P.J., and Torok, N. 2001 "Ekpyrotic universe: Colliding branes and the origin of the hot big bang", *Physical Review D* 64: 123522-1-24.

Khoury, J., Ovrut, B.A., Seiberg, N., Steinhardt, P.J., and Torok, N. (2002) "From big crunch to big bang", *Physical Review D* 65: 086007-1-8.

Kolb, E.W., Matarrese, S. and Riotto, A. 2005 "On cosmic acceleration without dark energy", astro-ph/0506534.

Krauss, L.M. and Starkman, G.D. 2000 "Life, the universe, and nothing: Life and death in an ever-expanding universe", *Astrophysical Journal* 531: 22–30.

Krauss, L.M. and Turner, M.S. 1999 "Geometry and Destiny", *General Relativity and Gravitation* 31: 1453–1459.

Kuchař, K. 1992 "Time and the Interpretation of Quantum Gravity", in G. Kunsatter, D. Vincent, and J. Williams (eds.), *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, 211–314. Singapore: World Scientific.

Lebowitz, J.L. 1993 "Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy", *Physica A* 194: 1–27.

— 1999 "Statistical mechanics: A selective review of two central issues", *Reviews of Modern Physics* 71: S346–S357.

Penrose, R. 1979 "Singularities and time asymmetry", in S.W. Hawking and W. Israel (eds.), *General Relativity: An Einstein Centenary*, 581–638. Cambridge: Cambridge University Press.

— 1989 *The Emperor's New Mind: Concerning Computers, Minds, and the laws of Physics*. Oxford: Oxford University Press.

— 2004 *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: Jonathan Cape.

Price, H. 2004 "On the Origins of the Arrow of Time: Why There is Still a Puzzle about the Low-Entropy Past", in C. Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*, 219–239. London: Blackwell.

Steinhardt, P.J. and Turok, N. 2002 "Cosmic evolution in a cyclic universe", *Physical Review D* 65: 126003-1-20.

Tolman, R.C. 1934 *Relativity, Thermodynamics and Cosmology*. Oxford: Oxford University Press.

Wald, R.M. 1980 "Quantum Gravity and Time Reversibility", *Physical Review D* 21: 2742–2755.

— 1984 *General Relativity*. Chicago, IL: University of Chicago Press.

— 1994 *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago, IL: University of Chicago Press.