BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 6, November 1976

BOOK REVIEWS

The large scale structure of space-time, by S. W. Hawking and G. F. R. Ellis, Cambridge Monographs on Mathematical Physics, Cambridge University Press, New York, 1975, 391 + xi pp., $10.95 (paper), $28.50 (cloth).

This is an exciting and important volume since it is the first comprehensive presentation of a theory of cosmology taking into account the discoveries of the past quarter century in particle physics, radio astronomy, and differential topology. The astronomical universe or cosmos is examined within the framework of general relativity and global differential geometry. The exposition is authoritative and painstaking, although in the search for logical completeness sometimes a bewildering tangle of alternatives and complexities is introduced (see, for instance, Chapter 6 on causal structure). The authors assume a basic knowledge of the physical aspects of general relativity theory, and write for the reader who is skilled in tensor calculus but who wishes to see the appropriate concepts defined in an intrinsic coordinate-free manner suitable for a global geometry.

Concerning the central thesis of their treatise, the authors write in the preface:

"The subject of this book is the structure of space-time on length scales from 10^{-13} cm, the radius of an elementary particle, up to 10^{28} cm, the radius of the universe. For reasons explained in Chapters 1 and 3, we base our treatment on Einstein's General Theory of Relativity. This theory leads to two remarkable predictions about the universe: first, that the final fate of massive stars is to collapse behind an event horizon to form a 'black hole' which will contain a singularity; and secondly, that there is a singularity in our past which constitutes, in some sense, a beginning to the universe. Our discussion is principally aimed at developing these two results. They depend primarily on two areas of study: first, the theory of the behaviour of families of timelike and null curves in space-time, and secondly, the study of the nature of the various causal relations in any space-time."

Thus, for the authors, a mathematical model of the space-time universe consists of

(i) A differentiable 4-manifold $\mathfrak{M}$ (connected, Hausdorff, paracompact, C^∞-manifold without boundary)–this represents the amorphous qualitative structure of the cosmos.

(ii) A Lorentz metric tensor $\mathbf{g}$, with components $g_{ab}(x)$ in any local chart $(x^a), a = 1, 2, 3, 4$, of $\mathfrak{M}$ (a symmetric covariant 2-tensor field of class C^r, $r \geqslant 2$, with the relativistic signature 2, or $(+ + + -)$)–this represents the special relativistic or Minkowski structure on each tangent space T_p of $\mathfrak{M}$, and permits the construction of spacelike and nonspacelike (timelike and lightlike or null) tangent vectors in T_p. The nonspacelike vectors fill the lightcones in each T_p and these define the basic causal structure on $\mathfrak{M}$. Timelike and null geodesics define the world-trajectories or histories of free particles and light rays.

(iii) A matter energy-momentum tensor **T**, with components $T_{ab}(x)$ in any local chart (x^a) on $\mathfrak{M}$ (a symmetric covariant 2-tensor field of class C^r, satisfying the "conservation condition" $T^{ab}{}_{;b} = 0$)–this represents the local effect of matter and energy on the quantitative or inertial structure of the universe $\mathfrak{M}$. The most commonly assumed form for **T** is that of a perfect fluid

$$T^{ab} = (\mu + p)V^aV^b + pg^{ab}.$$

Here the unit timelike 4-velocity vector is V, the scalar energy density is $\mu > 0$, and the pressure is p. If $p = 0$, the fluid reduces·to a stream of dust particles of density μ. In empty space-time $\mathbf{T} = 0$.

(iv) An Einstein field equation for **g** on $\mathfrak{M}$ (a quasi-linear hyperbolic partial differential equation to determine g_{ab} from T_{ab} and possible boundary, asymptotic, and symmetry hypotheses)

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = 8\pi T_{ab}.$$

Here the Ricci curvature tensor $R_{bd} = R^a_{bad}$ is the contracted Riemann curvature tensor, the scalar curvature $R = R^a_a$, and Λ is the "cosmological constant" often assumed zero. See Chapter 7 for a discussion of this partial differential equation. (Recall that R_{ab} involves second derivatives of g_{ab}.)

In this review I shall first tabulate the 10 chapters with a brief comment on the contents of each (there are also two appendices: A, an essay by P. S. Laplace on the trapping or capture of light by a massive Newtonian gravitational body; B, an exposition of the theory of G. D. Birkhoff characterizing the Schwarzschild solution). Following this tabulation I shall attempt a critique of some of the major physical and mathematical results expounded in the text. Since the authors claim nothing less than an explanation of the birth and death of the cosmos, some philosophical comments are also in order.

Chapter 1. The role of gravity. This is an exposition of the historical background for the theory, an insight into the geometrical significance of geodesics and conjugate points, and an outline for the remainder of the text. Accordingly, an extensive quote is useful here:

"Not only is gravity the dominant force on a large scale, but it is a force which affects every particle in the same way. This universality was first recognized by Galileo, who found that any two bodies fell with the same velocity. This has been verified to very high precision in more recent experiments by Eotvos, and by Dicke and his collaborators (Dicke 1964). It has also been observed that light is deflected by gravitational fields. Since it is thought that no signals can travel faster than light, this means that gravity determines the causal structure of the universe, i.e., it determines which events of space-time can be causally related to each other.

These properties of gravity lead to severe problems, for if a sufficiently large amount of matter were concentrated in some region, it could deflect light going out from this region so much that it was in fact dragged back inwards. This was recognized in 1798 by Laplace, who pointed out that a body of about the same density as the sun but 250 times its radius would exert such a strong gravitational field that no light could escape from its surface. That this should have been predicted so early is so striking that we give a translation of Laplace's essay in an appendix.

One can express the dragging back of light by a massive body more precisely using Penrose's idea of a closed trapped surface. Consider a sphere $\mathcal{T}$ surrounding the body. At some instant let $\mathcal{T}$ emit a flash of light. At some later time t, the ingoing and outgoing wave fronts from $\mathcal{T}$ will form spheres $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively. In a normal situation, the area of $\mathcal{T}_1$ will be less than that of $\mathcal{T}$ (because it represents ingoing light) and the area of $\mathcal{T}_2$ will be greater than that of $\mathcal{T}$ (because it represents outgoing light; see Figure 1). However if a sufficiently large amount of matter is enclosed within $\mathcal{T}$, the areas of $\mathcal{T}_1$ and $\mathcal{T}_2$ will *both* be less than that of $\mathcal{T}$. The surface $\mathcal{T}$ is then said to be a closed trapped surface. As t increases, the area of $\mathcal{T}_2$ will get smaller and smaller provided that gravity remains attractive, i.e. provided that the energy density of the matter does not become negative. Since the matter inside $\mathcal{T}$ cannot travel faster than light, it will be trapped within a region whose boundary decreases to zero within a finite time. This suggests that something goes badly wrong. We shall in fact show that in such a situation a space-time singularity must occur, if certain reasonable conditions hold."

"In Chapter 8 we discuss the definition of space-time singularities. This presents certain difficulties because one cannot regard the singular points as being part of the space-time manifold $\mathcal{M}$.

We then prove four theorems which establish the occurrence of space-time singularities under certain conditions. These conditions fall into three categories. First, there is the requirement that gravity be attractive. This can be expressed as an inequality on the energy-momentum tensor. Secondly, there is the requirement that there is enough matter present in some region to prevent anything escaping from that region. This will occur if there is a closed trapped surface, or if the whole universe is itself spatially closed. The third requirement is that there should be no causality violations. However this requirement is not necessary in one of the theorems. The basic idea of the proofs is to use the results of Chapter 6 to prove there must be longest timelike curves between certain pairs of points. One then shows that if there were no singularities, there would be focal points which would imply that there were no longest curves between the pairs of points."

Chapter 2. Differential geometry. This is a standard treatment of the fundamental properties of differentiable manifolds, differentiable maps, vector and tensor fields, and fiber bundles. The definitions are precise and intrinsic, and the calculations are displayed in coordinate notation. The only unusual features are the attention to the Weyl conformal curvature tensor (as the trace-free part of the Riemann curvature tensor), and the theory of hypersurfaces and Gauss' theorem in the case of a Lorentz metric.

Chapter 3. General relativity. "The mathematical model we shall use for space-time, i.e. the collection of all events, is a pair $(\mathcal{M}, \mathbf{g})$ where $\mathcal{M}$ is a connected four dimensional Hausdorff C^∞ manifold and $\mathbf{g}$ is a Lorentz metric (i.e. a metric of signature $+2$) on $\mathcal{M}$."

Two models will be taken equivalent if they are globally isometric. Further it is assumed that $(\mathcal{M}, \mathbf{g})$ is inextendible, that is, it is not isometric to a proper open submanifold of some extension $(\mathcal{M}', \mathbf{g}')$. This last hypothesis is made to try to make certain that all singularities, evidenced by inextendible yet incomplete nonspacelike geodesics, arise from intrinsic physical causes rather than from mathematical devices such as excising a closed subset of $\mathfrak{M}$.

The postulates for the matter field T^{ab} are local causality (essentially that the matter dynamics satisfies hyperbolic partial differential equations whose characteristic cones are compatible with the light cones of **g**), and local conservation $T^{ab};_b = 0$. This latter condition implies a strict conservation law only when some symmetry of space-time (say, a Killing vector field) is present.

The "derivation" of the format for special energy-momentum tensors follows a standard program of variational techniques for appropriate Lagrangians L.

The Einstein field equation is defended in the usual ways as an analogue of the Laplace-Poisson equation for the Newtonian potential, and also by the variation of the integral $\int[(8\pi)^{-1}(R - 2\Lambda) + L]\,dv$.

Certainly this field equation

$$R_{ab} - \tfrac{1}{2}Rg_{ab} + \Lambda g_{ab} = 8\pi T_{ab}$$

is one of the weakest links in the entire theory. It is absolutely foundational, yet rests on historical tradition, rarified intuitive guesswork, and very limited physical observation. The authors mention the competitive theories of Nordström, Hoyle-Bondi, and also of Brans-Dicke, but proceed on the basis of Einstein's General Relativity.

Chapter 4. The physical significance of curvature. Consider a congruence of timelike smooth curves in $\mathfrak{M}$ with timelike unit tangent vectors V initially orthogonal to a spacelike hypersurface $\mathcal{H}$. These curves could represent the histories of small test particles, or the flow lines of a perfect fluid.

For simplicity of exposition we consider the case where the flow lines are timelike geodesics orthogonal to an initial spacelike hypersurface $\mathcal{H}$. Then $V_{a;b} = \chi_{ab}$ is the second fundamental form on $\mathcal{H}$. We recall some methods of the calculus of variations to measure the expansion of the fluid flow lines.

Choose an orthonormal basis $E_\alpha, \alpha = 1, 2, 3$, for the spacelike tangent space H_q at an initial point $q \in \mathcal{H}$, and propagate this basis along the flow by parallel translation to form a corresponding basis at each point $\gamma(s)$, after a proper time s, on the geodesic initiating at $\gamma(0) = q$. Suppose $\lambda(\sigma)$ is a curve in $\mathcal{H}$ with tangent vector $Z \in H_q$. Then, by transporting this curve along the fluid flow we define the tangent vector $Z(s)$ at $\gamma(s)$. That is, $Z(s)$ is a variational vector measuring the separation of infinitesimally neighboring geodesics. Let $_\perp Z(s)$ be the component of $Z(s)$ which is orthogonal to V and denote its components by $_\perp Z = Z^\alpha E_\alpha$. From the definition of this variation vector $_\perp Z$ we compute $dZ^\alpha/ds = V^\alpha_{;\beta}\, Z^\beta$. Thus, (misprint in equation 4.39), $Z^\alpha(s) = A_{\alpha\beta}(s)Z^\beta(0)$ where the 3×3 fundamental matrix $A(s) = A_{\alpha\beta}(s)$ satisfies

$$dA_{\alpha\beta}(s)/ds = V_{\alpha;\gamma}A_{\gamma\beta}(s), \qquad A_{\alpha\beta}(0) = \delta_{\alpha\beta}.$$

Upon differentiating we obtain the Jacobi variational equation

$$d^2A_{\alpha\beta}/ds^2 = -R_{\alpha 4\gamma 4}A_{\gamma\beta}$$

with the initial conditions

$$A_{\alpha\beta}(0) = \delta_{\alpha\beta}, \qquad dA_{\alpha\beta}(0)/ds = \chi_{\alpha\beta}.$$

In the case of a fluid flow the matrix $A_{\alpha\beta}$ can be regarded as representing the shape and orientation of a small element of fluid which is spherical at the initial point q. As is customary in fluid mechanics, we define the volume expansion coefficient (negative of divergence) by

$$\theta = (\det A)^{-1} \frac{d}{ds} (\det A) = \frac{d}{ds} (\ln \det A).$$

Then a direct ingenious calculation yields the basic equation discovered independently by Landau and Raychaudhuri,

$$d\theta/ds = -R_{ab}V^aV^b - 2\sigma^2 - \tfrac{1}{3}\theta^2$$

where $\sigma^2 \geqslant 0$ is the shear effect on the expansion $\theta(s)$ (and the vorticity term vanishes since $A(0) = I$). If we assume that $R_{ab}V^aV^b \geqslant 0$, then the Raychaudhuri equation yields the important differential inequality

$$d\theta/ds \leqslant -\tfrac{1}{3}\theta^2.$$

If $\theta(0) = \chi_{ab}g^{ab} < 0$, the flow lines are initially convergent, and then this nonlinear inequality forces $\theta \to -\infty$ at a point p after a finite proper time in the future. This would imply that $\det A(s) \to 0$ so that p is a conjugate point to $\mathcal{H}$ along the geodesic flow line through q. A similar argument could yield a conjugate point to $\mathcal{H}$ in the past.

It is this deduction (Propostion 4.43)–that a positive Ricci curvature produces conjugate points–which is the heart of the mathematical argument of the text. The physical justification for this curvature hypothesis is the energy condition $T_{ab}V^aV^b \geqslant 0$ which is satisfied for a dust stream, or a fluid with negligible pressure p. From the Einstein field equations (say, with $\Lambda = 0$)

$$R_{ab} = 8\pi(T_{ab} - \tfrac{1}{2}Tg_{ab}),$$

so we obtain the required timelike convergence condition

$$R_{ab}V^aV^b = 8\pi\left(T_{ab}V^aV^b + \tfrac{1}{2}T\right) \geqslant 0.$$

The final section of Chapter 4 concerns the variation of arclength, using the techniques of the calculus of variations. Using the theory of the first variation, we can characterize a geodesic in Riemannian geometry as minimizing the arc length locally. Similarly, a timelike geodesic in Lorentzian geometry maximizes the proper time duration, at least locally. In considering the second variation $L(Z_1, Z_2)$ we shall say that a timelike geodesic curve $\gamma(t)$ from a spacelike hypersurface $\mathcal{H}$ to a point p is maximal if $L(Z_1, Z_2)$ is negative semidefinite. Thus if $\gamma(t)$ is not maximal there is a small variation which yields a longer (proper-time) curve from $\mathcal{H}$ to p. By standard variational methods we now obtain:

PROPOSITION 4.5.9. *A timelike geodesic curve $\gamma(t)$ from $\mathcal{H}$ to p is maximal if and only if there is no point in $(\mathcal{H}, p)$ conjugate to $\mathcal{H}$ along γ.* (*Misprint occurs in statement in text.*)

The remaining logical step, namely the existence of a longest timelike geodesic from p to $\mathcal{H}$ when $\mathfrak{M}$ is complete, is proved in Chapter 6 (Proposition 6.7.1) after certain problems in causality are clarified.

Chapter 5. Exact solutions. Numerous classical solutions ($\mathfrak{M}$, $\mathbf{g}$) are examined as global Lorentz manifolds. In particular, completeness, causal structure, event horizons, asymptotic behavior, and symmetries are specified intrinsically.

The classical models are those of Minkowski (special relativity), de Sitter and anti-de Sitter worlds (constant curvature), Robertson-Walker-Friedmann worlds (expanding universe), Gödel space, Taub-NUT space, with some indication of the plane wave solutions in empty space.

A thorough discussion, with many diagrams, is given for the local solutions of Schwarzschild (spherically symmetric) and Kerr (axisymmetric). The singularity at $r = 0$ in the Schwarzschild world ($\mathfrak{M}$, $\mathbf{g}$) is real (an intrinsic curvature becomes infinite as $r \to 0$), but that at $r = 2m$ (Schwarzschild radius) is not physically singular since there is an extension ($\mathfrak{M}^*$, $\mathbf{g}^*$) of ($\mathfrak{M}$, $\mathbf{g}$) through $r = 2m$, as determined by Kruskal. In fact the Kruskal extension is the unique analytic and locally inextendible extension of the Schwarzschild solution. Similar results hold for the Kerr world.

The amount of detailed special information and conceptual analysis contained in these classical models and examples is overwhelming. They should be a source of scientific inspiration for decades to come. For instance, even the Minkowski space F^4 (R^4 with flat Lorentz tensor) contains new lessons. Namely, all of F^4 is conformally diffeomorphic to a proper subregion of the Einstein cylindrical space (product $S^3 \times R^1$). Using this embedding of F^4 in $S^3 \times R^1$ we can study the causal structure of F^4 at infinity. Similar conformal embeddings of de Sitter and the Robertson-Walker spaces in $S^3 \times R^1$ are possible. The Penrose diagram is a convenient graphical scheme for illuminating these embeddings.

The reviewer looked in vain for models displaying the various Petrov types of gravitational waves. Also here, or in the next Chapter 6, there could well be an indication of Zeeman's theorem that any causal automorphism of F^4, or of the Einstein cylindrical world $S^3 \times R^1$, is an isometry (after scale change).

Chapter 6. Causal structure. The causal structure of a Lorentz manifold ($\mathfrak{M}$, $\mathbf{g}$) is the collection of all Lorentz metrics conformally equivalent to $\mathbf{g}$. Thus the causal structure on $\mathfrak{M}$ can be specified by the field of light cones, or by the causal relation $q \ll p$ (q precedes p along a future-directed timelike curve). For this relation to be defined we need a time-oriented or isochronal world (which can always be achieved using a 2-fold covering space-time). We assume this is henceforth the case.

However if there were a closed timelike curve on $\mathfrak{M}$ (which is always the case when $\mathfrak{M}$ is compact), then our notions of physical causality and free will would be violated. Various kinds of causality conditions and hypotheses are possible as one rules out recurrent timelike, lightlike, or nonspacelike curves. Since the discipline of topological dynamics has created a whole hierarchy of kinds of recurrence (periodic, almost periodic, pointwise, regional, etc.) these give rise to corresponding kinds of causality violations. One sensible mental anchor in this phantasmagoric maelstrom is the concept of "structural stability" which here surfaces as the *stable causality condition*.

We say that the stable causality condition holds on $\mathfrak{M}$ if the spacetime metric $\mathbf{g}$ has an open neighborhood in the C^0 topology (compact-open topology on tensor fields on $\mathfrak{M}$) such that there are no closed timelike curves in any of the metrics belonging to this neighborhood. Then one can prove the interesting:

PROPOSITION 6.4.9. *The stable causality condition holds everywhere on $\mathfrak{M}$ if and only if there is a function f on $\mathfrak{M}$ whose gradient is everywhere timelike.*

The primary goal of this study of causality is to determine a partial Cauchy surface, that is, a spacelike hypersurface $\mathcal{S}$ which no nonspacelike curve intersects more than once. Moreover a partial Cauchy surface $\mathcal{S}$ is a global Cauchy surface in case each point $p \in \mathfrak{M}$ lies either in the past or future domain of dependence of $\mathcal{S}$, that is, $D^-(\mathcal{S}) \cup D^+(\mathcal{S}) = \mathfrak{M}$. Thus a global Cauchy surface is a spacelike hypersurface which every inextendible nonspacelike curve intersects exactly once.

Suppose $\mathfrak{M} = R^1 \times \mathcal{S}$ where $\mathcal{S}$ is a compact 3-manifold and for each $t_0 \in R^1$ the submanifold $t_0 \times \mathcal{S}$ is a Cauchy surface for $\mathfrak{M}$. From 6.7.1 we obtain the following

COROLLARY. *For each $p \in D^+(\mathcal{S})$ there exists a timelike geodesic curve $\gamma(t)$ which is orthogonal to $\mathcal{S}$, and which realizes the longest proper time from $\mathcal{S}$ to p. Moreover, from the theory of the second variation, $\gamma(t)$ contains no point conjugate to $\mathcal{S}$ between $\mathcal{S}$ and p.*

We can replace the assumption that $\mathcal{S}$ is compact by the requirement that $\mathfrak{M}$ is globally hyperbolic in the sense of Leray. This important existence theorem is proved by the direct method of the calculus of variations.

Chapter 7. The Cauchy problem in General Relativity. The Einstein field equations

$$R^{ab} - \tfrac{1}{2} R g^{ab} = 8\pi T^{ab}$$

are meant to determine the Lorentz metric tensor g^{ab} on the 4-manifold $\mathfrak{M}$, in terms of the physically prescribed energy-momentum tensor T^{ab}. There is a reduced form of the Einstein tensor, locally written

$$R^{ab} - \tfrac{1}{2} R g^{ab} = \tfrac{1}{2}\, g^{ij}\partial_{ij} g^{ab} + \big(\text{terms in } \partial_i g^{cd} \text{ and } g^{ef}\big),$$

provided we establish the "gauge harmonic conditions"

$$\psi^b = \partial_c g^{bc} - \tfrac{1}{2}\, g^{bc} g_{de} \partial_c g^{de} = 0.$$

If $T^{ab} = 0$ (empty space-time), or T^{ab} is reasonably behaved, then the field equations form a quasilinear hyperbolic system of 10 equations for the 10 components of the metric tensor. Hence the Cauchy initial value problem is well posed; developing $\mathbf{g}$ from its initial value $\mathbf{h}$ which is a Riemannian metric on some initial 3-manifold $\mathcal{S}$, and normal data $\partial\mathbf{g}$ on $\mathcal{S}$. The 6 initial components of this Riemann metric $\mathbf{h}$ are sufficient data to establish the existence of the solution $\mathbf{g}$ of the reduced Einstein equations, in view of the 4 guage conditions imposed above, provided $\mathbf{h}$ satisfies the necessary restriction of the field equations to $\mathcal{S}$. Locally every solution $\mathbf{g}$ of the full Einstein field equations (say in empty space-time) satisfies the gauge conditions, and hence the reduced Einstein equations, relative to some differentiable coordinate chart; although such local coordinates might not be related to any particular metric. Thus we expect the solution $\mathbf{g}$ to be unique only up to a diffeomorphism (not isometry) of $\mathfrak{M}$.

At least this is the classical or intuitive viewpoint. But this brief discussion leaves out many serious problems–what are the gauge conditions in an invariant description and why can they be demanded; what initial data are admissible and do they impose topological restrictions on $\mathcal{S}$; how can one define the Cauchy problem in terms of $\mathcal{S}$ when the space-time $(\mathfrak{M}, \mathbf{g})$ itself is to be developed?

In this chapter a technical proof is provided for the existence and uniqueness of the global Cauchy development $(\mathfrak{M}, \mathbf{g})$ from $\mathcal{S}$ with appropriate initial data ω. The global Cauchy problem for the Einstein field equations (first, in empty space-time) is thus specified as follows:

Given: a 3-manifold $\mathcal{S}$ with data $\omega = (\mathbf{h}, \chi)$ where $\mathbf{h}$ is a Riemannian metric on $\mathcal{S}$ and χ is a symmetric second order tensor which plays the role of the second fundamental form for the hypersurface $\mathcal{S}$. Require that the necessary compatibility or constraint equations hold on $\mathcal{S}$

$$\chi^{cd}{}_{\|d} h_{ce} - \chi^{cd}{}_{\|e} h_{cd} = 0,$$

$$\tfrac{1}{2}\left(R' + \left(\chi^{dc} h_{dc}\right)^2 - \chi^{ab}\chi^{cd} h_{ac} h_{bd}\right) = 0,$$

where R' is the curvature scalar of $\mathbf{h}$ and the covariant differentiation is relative to $\mathbf{h}$ on $\mathcal{S}$.

To find: a differentiable imbedding of $\mathcal{S}$ in a Lorentz 4-manifold $(\mathfrak{M}, \mathbf{g})$ such that $\mathbf{g}$ satisfies the Einstein field equation in $\mathfrak{M}$ and reduces to $\mathbf{h}$ on the global Cauchy hypersurface $\mathcal{S}$, for which the second fundamental form is χ.

The basic result of 7.5 asserts that there exists such a local (in time) Cauchy development $(\mathfrak{M}, \mathbf{g})$ from $(\mathcal{S}, \omega)$. The local uniqueness theorem asserts that two such local Cauchy developments $(\mathfrak{M}, \mathbf{g})$ and $(\mathfrak{M}', \mathbf{g}')$ are equivalent under a diffeomorphism of $\mathfrak{M}$ onto $\mathfrak{M}'$ (at least in a neighborhood of $\mathcal{S}$), carrying $\mathbf{g}$ to $\mathbf{g}'$, and holding $\mathcal{S}$ fixed pointwise.

The first part of the proof deals with the local solution of quasilinear hyperbolic systems, such as the reduced Einstein equations in empty space-time. The methods are standard in the theory of differential equations, but the intricate calculations are carefully presented. For linear systems with analytic data the Cauchy-Kowalevski power series methods produce the solution. For differentiable data, we approximate by analytic data and use some a priori estimates that the solutions depend continuously, in an appropriate Hilbert-Sobolev space $W^{4+\alpha}$, on the data. For quasilinear equations we are led by successive approximations through a sequence of linear equations. A fix-point theorem then establishes the required local existence theorem.

Next the authors attack the problem of the global existence of the Cauchy development $(\mathfrak{M}, \mathbf{g})$ from appropriate initial data $(\mathcal{S}, \omega)$. Here the exposition becomes rather chaotic, as though the writers and the readers are equally exhausted.

The global construction begins with the differentiable product $\hat{\mathfrak{M}} = \mathcal{S} \times R$, and we introduce an auxiliary Lorentz metric $\hat{\mathbf{g}}$ on $\hat{\mathfrak{M}}$, at least in a neighbourhood of $\mathbb{S}$. The gauge conditions are now formed using the covariant derivatives for $\hat{\mathbf{g}}$.

Usually $\hat{\mathbf{g}}$ is not locally flat, but the Christoffel symbols are nearly constant in a sufficiently small region. Such an approximation is now adequate to employ the above local existence theorem for $\mathbf{g}$.

For the global extension of $\mathbf{g}$ (over the global 3-space $\mathbb{S}$, but for short time t) we must consider the solution $\omega = (\mathbf{h}, \chi)$ of the elliptic system of constraint equations on $\hat{\mathbb{S}}$. Then we construct local solutions $\mathbf{g}$ near each point of $\mathbb{S} \subset \hat{\mathfrak{M}}$ and patch these together to obtain the desired Cauchy development $(\mathfrak{M}, \mathbf{g})$. It seems possible that both of these steps might encounter some topological obstructions on $\mathbb{S}$, but the authors assert that this is not the case.

Finally there is the global development $(\tilde{\mathfrak{M}}, \tilde{\mathbf{g}})$ for a maximal time duration. By general arguments on the partially ordered set of local developments the authors show that $(\tilde{\mathfrak{M}}, \tilde{\mathbf{g}})$ is unique. This uniqueness rests on the fact that $\mathbb{S}$ is a global Cauchy surface in $\tilde{\mathfrak{M}}$. Otherwise the uniqueness fails. For example, consider $\mathbb{S}$ to be the Riemannian flat torus T^3 with $\chi = 0$. Then either $T^3 \times R$ or $T^3 \times S^1 = T^4$ could serve as the locally Minkowski manifold $\tilde{\mathfrak{M}}$.

The authors then discuss the stability of the solution $(\tilde{\mathfrak{M}}, \tilde{\mathbf{g}})$, that is, the solution depends continuously on the data $(\mathbb{S}, \omega)$. This continuity can be established in case of C^∞-data but is not wholly satisfying to the authors in the sense of W^r-data. This uncertainty gives rise to fears that perhaps some unexpected esoteric shock-wave phenomena might be undetected.

At the end of the chapter there is a very brief statement that the same types of results are valid for the Einstein equations with matter and energy present.

Chapter 8. Space-time singularities. The most important theories of this text concern the mathematical existence proofs of singularities in an inextendible Lorentz manifold $(\mathfrak{M}, \mathbf{g})$, and the interpretations for astrophysical and cosmological models. Since there is no intrinsic positive-definite metric on $\mathfrak{M}$ there arise difficulties in defining the concepts of completeness, or the boundary at infinity. Yet the existence of finite inextendible nonspacelike curves is taken as the basic condition defining singularities.

On the Lorentz manifold $(\mathfrak{M}, \mathbf{g})$ a timelike geodesic is complete if it can be extended in $\mathfrak{M}$ for all real values of the proper time. If a timelike geodesic in $\mathfrak{M}$ were incomplete, then an inertial observer could suddenly cease to exist for no evident physical reason. Even lightlike geodesics possess a distinguished affine parameter, although this is not easy to understand physically, and thus admit the property of completeness. The authors state: "We shall therefore adopt the view that timelike and null geodesic completeness are minimum conditions for space-time to be considered singularity-free. Therefore if a space-time is timelike or null geodesically incomplete, we shall say that it has a singularity."

But then they continue to protest:

"If one is going to say that there is a singularity in a space-time in which a freely falling observer comes to an untimely end, one should presumably do the same for an observer in a rocket ship. What one needs is some generalization of the concept of an affine parameter to all C^1-curves, geodesic or nongeodesic."

In order to enlarge the class of singularities to incorporate incomplete C^1-curves, we define arc-length along each such curve $\mathcal{C}$ by using a basis of tangent vectors along $\mathcal{C}$ obtained by parallel translation according to the Christoffel connection. More technically we note that the bundle of Lorentz orthogonal frames $\mathcal{O}(\mathfrak{M})$ is automatically a parallelizable manifold. Thus a Euclidean metric at any one point of $\mathcal{O}(\mathfrak{M})$ defines a global Riemann metric on the bundle $\mathcal{O}(\mathfrak{M})$–and any two such Riemann metrics are equivalent. We define $(\mathfrak{M}, \mathbf{g})$ to be *b*-complete just in case the bundle $\mathcal{O}(\mathfrak{M})$ is complete as a Riemannian manifold. It might be advantageous to limit these considerations to nonspacelike C^1-curves in $\mathfrak{M}$, but this concept seems to be rather inaccessible mathematically.

DEFINITION. A space-time $(\mathfrak{M}, \mathbf{g})$ is *singularity-free* if it is *b*-complete.

While such incomplete singularities are mathematically significant they seem to lack any physical cause. In cosmology one might demand some evidence of very strong gravitational fields causing the break-down of space-time, or of its description within General Relativity. In this spirit we define a *curvature singularity* if the scalar curvature (or some other intrinsic curvature of either the Lorentz manifold $\mathfrak{M}$ or the Riemann manifold $\mathcal{O}(\mathfrak{M})$) becomes unbounded along some incomplete curve in $\mathfrak{M}$.

In view of the results on conjugate points in Chapter 6 it is easy to prove that certain types of space-time are geodesically incomplete and so contain singularities. For instance, consider a slightly weakened form of Hawking's theorem:

THEOREM 4′. *Space-time* $(\mathfrak{M}, \mathbf{g})$ *is not timelike geodesically complete if*:

(1) $R_{ab}K^aK^b \geqslant 0$ *for every nonspacelike vector* K,

(2) *there exists a compact spacelike* 3-*surface* $\mathcal{S}$ *which is a global Cauchy surface for* $\mathfrak{M}$,

(3) *the unit normals to* $\mathcal{S}$ *are everywhere converging* (*or everywhere diverging*) *on* $\mathcal{S}$.

PROOF. Since $\theta(0) < -\varepsilon < 0$ everywhere on $\mathcal{S}$, we conclude that there is a conjugate point (with $s < 3/\varepsilon$) on each geodesic γ normal to $\mathcal{S}$. If $\mathfrak{M}$ were geodesically complete, take a point $p = \gamma(4/\varepsilon)$ and construct the longest timelike geodesic $\hat{\gamma}$ from $\mathcal{S}$ to p. Since $\hat{\gamma}$ is longest, it must be maximal and so admits no conjugate point between $\mathfrak{M}$ and p–which provides a contradiction.

The first such theorem on singularities was obtained by Penrose (1965). "It was designed to prove the occurrence of a singularity in a star which collapsed inside its own Schwarzschild radius. If the collapse were exactly spherical, the solution could be integrated explicitly and a singularity would always occur. However it is not obvious that this would be the case if there were irregularities or a small amount of angular momentum. Indeed in Newtonian theory the smallest amount of angular momentum could prevent the occurrence of infinite density and cause the star to re-expand. However Penrose showed that the situation was very different in General Relativity: once the star had passed inside the Schwarzschild surface (the surface $r = 2m$) it could not come out again. In fact the Schwarzschild surface is defined only for an exactly symmetric solution but the more general criterion used by Penrose is equivalent for such a solution and is applicable also to solutions without exact symmetry. It is that there should exist a *closed trapped surface* $\mathcal{T}$." By this is meant a C^2 closed spacelike two-surface (normally, S^2) such that the two families of null geodesics orthogonal to $\mathcal{T}$ are convergent at $\mathcal{T}$.

THEOREM 1. *Space-time* ($\mathfrak{M}$, $\mathbf{g}$) *cannot be null geodesically complete if*:
(1) $R_{ab}K^aK^b \geqslant 0$ *for all null vectors* K^a,
(2) *there is a noncompact Cauchy surface* $\mathcal{H}$ *in* $\mathfrak{M}$,
(3) *there is a closed trapped surface* $\mathcal{T}$ *in* $\mathfrak{M}$.

The method of proof is to show that the boundary of the future of $\mathcal{T}$ would be compact if $\mathfrak{M}$ were null geodesically complete. This is then shown to be incompatible with $\mathcal{H}$ being noncompact.

The results on the physically appealing concept of curvature singularity are rather sparse. An interesting proposition is 8.5.2.

If $p \in \mathfrak{M}$ *is a limit point of a b-incomplete curve* λ *and if at* p, $R_{ab}K^aK^b \neq 0$ *for all nonspacelike vectors* K^a, *then* λ *corresponds to a curvature singularity.*

In conclusion, the authors emphasize that these singularities are structurally stable in the sense that they persist even if the space-time tensors g^{ab} and T^{ab} are perturbed. Hence they do not depend on idealized symmetries, but could have geometric and physical importance in general situations.

Chapter 9. Gravitational collapse and black holes.

Chapter 10. The initial singularity of the universe. Since the final two chapters deal largely with the interpretations in astrophysics and cosmology of the geometry, I shall give a brief review of them together.

The basic idea is that through self-gravity a star might collapse to a radius less than its Schwarzschild radius, and so become a black hole, that is, a gravitational source that is invisible to the outside world. The Schwarzschild radius of the Sun is 3 Km; thus normal stars are a long way from their Schwarzschild radii. While most stars have a mass M of the same order of magnitude as $M_\odot$, the Solar mass, there seems to be a critical limit at mass $M_L = 1.5 M_\odot$.

"If the mass is less than M_L, the star can settle down eventually to a white dwarf state in which it is supported by degeneracy pressure of nonrelativistic electrons, or possibly to a neutron star state in which it is supported by neutron degeneracy pressure. However if the mass is slightly greater than M_L, there is no low temperature equilibrium state. Therefore the star must either pass within its Schwarzschild radius, or manage to eject sufficient matter that its mass is reduced to less than M_L.

"Ejection of matter has been observed in supernovae and planetary nebula, but the theory is not yet very well understood. . . . Present calculations in fact indicate that stars of mass $M > 5M_L$ would not be able to eject sufficient mass to avoid a relativistic collapse.

"To summarize, it seems that certainly some, and probably most, bodies of mass $> M_L$ will eventually collapse within their Schwarzschild radius and so give rise to a closed trapped surface."

The remainder of Chapter 9 is dedicated to giving a general definition of black holes and showing that they probably settle down finally to a Kerr solution of the Einstein field equations. The results are technical, difficult, and probably provisional. For instance, to define a black hole we must consider a space-time $(\mathfrak{M}, \mathbf{g})$ partitioned by a family $\mathcal{S}(\tau)$ of spacelike partial Cauchy surfaces. Assume $(\mathfrak{M}, \mathbf{g})$ is conformally imbedded in some $(\tilde{\mathfrak{M}}, \tilde{\mathbf{g}})$, forming there a manifold with boundary $\overline{\mathfrak{M}} = \mathfrak{M} \cup \partial\mathfrak{M}$. Here the boundary $\partial\mathfrak{M}$ consists of two null surfaces $\mathcal{I}^{+}$ and $\mathcal{I}^{-}$ which represent future and past null infinity respectively.

"We define a *black hole* on the surface $\mathcal{S}(\tau)$ to be a connected component of the set $\mathcal{B}(\tau) \equiv \mathcal{S}(\tau) - J^{-}(\mathcal{I}^{+}, \overline{\mathfrak{M}})$. In other words, it is a region of $\mathcal{S}(\tau)$ from which particles or photons cannot escape to $\mathcal{I}^{+}$."

Under various hypotheses on the causal past J^{-} and future J^{+} of $\mathcal{I}^{\pm}$ in $\overline{\mathfrak{M}}$, and the Cauchy developments $D^{\pm}(\mathcal{S})$, it can be shown that black holes are ever increasing. That is black holes can merge together, and new black holes can form as a result of further bodies collapsing; however black holes can never bifurcate.

In Chapter 10 the expansion of the universe is considered to be similar to the collapse of a star, except that the time sense is reversed. The astronomical evidence for this view takes into account the background black-body radiation of $2.7°\,K$, which is isotropic over the sky, and general information on the Hubble factor for the red-shift of galaxy spectra. These data, together with philosophical principles of the homogenity of the spatial universe, lead to a Robertson-Walker model of the expanding universe. The incompleteness results of Chapter 8 (say Theorem 4′) explain that the cosmos has expanded out of an initial singularity some time in our past.

After such a detailed exposition of the contents of the text, let me close this essay with only a few comments on the overall significance of the mathematical and physical theories expounded.

The most interesting mathematical aspect of the text is the systematic applications of the calculus of variations to global Lorentz geometry. Until recently the study of Lorentz manifolds consisted of a few very general results of differential topology, and several very particular results involving symmetry groups. Now research in Lorentz geometry can develop along patterns quite familiar from Riemann geometry, but with important and stimulating differences.

Also the text provides an important contribution to the global theory of hyperbolic partial differential systems. Just as the global theory of elliptic equations is treated on Riemannian manifolds, so hyperbolic equations belong on Lorentz manifolds. Further the singularity theorems of Chapter 8 may eventually become part of this theory together with a variety of other phenomena suggested by turbulence, shock-waves, bifurcation, and catastrophe theory. In other words the singularities already explored may be only the forerunners of a whole series of geometric constructions of even greater physical significance–even for cosmology.

Concerning the physical aspects of the text, I express admiration tempered with caution. The authors advance a theory of the birth and death of the universe; do they prove it? The reviewer, who is generally rather skeptical, remains a skeptic here. I believe that the observational data are too feeble, and the physical theory too tentative to support strong and conclusive beliefs.

"It seems to be a good principle that the prediction of a singularity by a physical theory indicates that the theory has broken down, i.e. it no longer provides a correct description of observations. The question is: when does General Relativity break down? One would expect it to break down anyway when quantum gravitational effects become important. . . . This would correspond to a density of 10^{94} gm cm^{-3}. However one might question whether a Lorentz manifold is an appropriate model for space-time on length scales of this order."

Cosmology is not a "hard experimental science" in the sense of aerodynamics, or even macro-economics. That is, the experimenter does not have access to an effective input control to the physical system. While there are plenty of observed data from millions of stars, they are not necessarily the data one might want if there were a choice. Unless we greatly modify our philosophy of scientific knowledge, cosmology must remain a speculation.

I recall once hearing a lecture on economics where the authority asserted something like, "since the coefficient of σ^2 is negative, we must expect the government to raise taxes", and I said to myself, "Hey, wait just a minute–let's multiply by -1". I had somewhat the same emotion when I read that the coefficient of σ^2 in Raychaudhuri's equation is negative, therefore we must all fall eventually into a black hole and this is our final fate. But even if this is the case, it might not be too dismal. Remember, Alice found a Wonderland.

LAWRENCE MARKUS

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 6, November 1976

Topics in stochastic processes, by Robert B. Ash and Melvin F. Gardner, Probability and Mathematical Statistics, vol. 27, Academic Press, New York, 1975, viii + 321 pp., \$34.50.

The theory of stochastic processes has mushroomed in the last twenty years; not only because of its intrinsic interest, but also because it is closely connected with so many different areas of mathematics. It feeds on analytic techniques from measure theory, Fourier analysis, semigroups of operators and spectral theory, potential theory, ergodic theory; and in turn it has applications to topology, functional inequalities, differential equations, information theory and prediction theory, and through the stochastic integral to several areas of mathematical physics. Thus stochastic processes is a good modern example of an area of mathematics which has been stimulated by its applications, while itself leading to extensive research in more established areas in order to develop the techniques needed.

The essential apparatus of mathematical building bricks is both extensive and deep. There is therefore no hope of writing a self contained text book of acceptable length which covers more than a small subset of the theory. The subset chosen by Ash and Gardner is made by selecting some special