General Relativity and Gravitation, Vol. 13, No. 12, 1981

Entropy of Self-Gravitating Radiation

RAFAEL D. SORKIN,[1] ROBERT M. WALD, and ZHANG ZHEN JIU[2]

Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637

Received January 6, 1981

Abstract

We examine the entropy of self-gravitating radiation confined to a spherical box of radius R in the context of general relativity. We expect that configurations (i.e., initial data) which extremize total entropy will be spherically symmetric, time symmetric distributions of radiation in local thermodynamic equilibrium. Assuming this is the case, we prove that extrema of S coincide precisely with static equilibrium configurations of the radiation fluid. Furthermore, dynamically stable equilibrium configurations are shown to coincide with local maxima of S. The equilibrium configurations and their entropies are calculated and their properties are discussed. However, it is shown that entropies higher than these local extrema can be achieved and, indeed, arbitrarily high entropies can be attained by configurations inside of or outside but arbitrarily near their own Schwarzschild radius. However, if we limit consideration to configurations which are outside their own Schwarzschild radius by at least one radiation wavelength, then the entropy is bounded and we find $S_{\max} \lesssim MR$, where M is the total mass. This supports the validity for self-gravitating systems of the Bekenstein upper limit on the entropy to energy ratio of material bodies.

§(1): *Introduction*

Recently, a number of interesting conclusions and speculations have been obtained by considering gedanken experiments involving a self-gravitating gas confined by a box. Studies of the processes of black hole formation and evaporation in such a system have led to insights into the nature of quantum gravitational dynamics (see, e.g., [1], [2]). However, in previous calculations [3] of the entropy contributions of the "ordinary matter" in the box (as opposed to

[1] Present address: Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742.
[2] Present address: Department of Physics, Central China Teachers' College, Wuhan, Peoples' Republic of China.

0001-7701/81/1200-1127$03.00/0

the entropy of the black hole), the self-gravitating nature of this matter has been neglected. The purpose of this paper is to study this effect.

Our interest in calculating the entropy of self-gravitating radiation confined to a box stems mainly from recent work of Bekenstein [4]. He has conjectured that for an arbitrary object of "size" R, the entropy to energy ratio is bounded by the black hole value,

$$S/E \leqq 2\pi R \tag{1}$$

He has supported this conjecture with some calculations of S/E for systems in flat space-time. Our main aim is to determine if equation (1) holds for self-gravitating systems.

We begin our investigation in Section 2 by studying the extrema of total entropy of self-gravitating radiation confined by a box of radius R in the context of general relativity. Assuming that such extrema are spherically symmetric and time symmetric, we prove that the extrema coincide with static equilibrium configurations. These equilibrium configurations are studied in Section 3. However, in Section 4 we show that these local extrema are not global maxima of entropy. For example, configurations of radiation which are well within their own Schwarzschild radius can exhibit arbitrarily high entropies, in violation of equation (1). Configurations which are outside of—but arbitrarily close to —their own Schwarzschild radius can also achieve arbitrarily high entropies. However, the configurations which are within their own Schwarzschild radius can be produced classically only by starting with a white hole. If one rules out the existence of white holes [5] these configurations cannot be produced. Furthermore, it is not plausible that configurations arbitrarily close to their own Schwarzschild radius can be achieved on account of the finite size of the constituents of the radiation. A reasonable limit appears to be that the configuration lies outside of its own Schwarzschild radius by a (proper) distance D at least as large as the (proper) wavelength λ of the ambient radiation,

$$D \gtrsim \lambda \tag{2}$$

As shown in Section 4, this limits the maximum possible entropy S to the Bekenstein formula and gives $S_{\max} \approx S_{\mathrm{bh}}$. Thus, if our limitation, equation (2), is correct, support is given to the Bekenstein limit for self-gravitating systems.

§(2): *Extrema of Total Entropy*

Our aim is to find the maximum entropy configurations of self-gravitating matter in general relativity having total energy M and confined by a spherical box of radius R. We shall neglect all quantum effects except, of course, for the basic quantum discreteness of radiation which gives it a well defined, finite entropy. In particular, we neglect the Casimir (i.e., zero-point) energy of the matter. In the cases we consider, this should be unimportant when the number of species of radiation is small, as we will assume. (Note, however, that the argument in [4] relies crucially on the magnitude and sign of the Casimir energy.) We shall treat gravity via classical general relativity and assign no entropy to the gravitational field although graviton contributions to entropy may be included by treating gravitons as a species of radiation.

In general relativity, an "instant of time" is described by a spacelike hypersurface, Σ. The geometry and extrinsic curvature, K_{ab}, of Σ are related to the energy density μ and momentum density J^a measured by an observer moving orthogonal to Σ by the Einstein constraint equations,

$$D^a(K_{ab} - Kh_{ab}) = 8\pi J_b \tag{3}$$

$$^{(3)}R = 16\pi\mu + K^{ab}K_{ab} - K^2 \tag{4}$$

where D_a is the derivative operator associated with the metric h_{ab} on the hypersurface, and $^{(3)}R$ is its scalar curvature. The total mass M of the configuration at the instant of time represented by Σ is determined entirely by the geometry of Σ. On the other hand, the total entropy, S, of the configuration depends on the composition and distribution of the constituents of matter. Thus, our first task in maximizing the total entropy is to maximize the entropy locally at fixed μ and J_a. We assume that this is done by putting all the matter in the form of radiation in a thermal (black body) distribution.

The stress energy tensor of thermal radiation is that of a perfect fluid with equation of state $P = \rho/3$,

$$T_{ab} = \rho u_a u_b + \tfrac{1}{3}\rho(g_{ab} + u_a u_b) \tag{5}$$

where u^a is the 4-velocity of the local rest frame of the radiation. By the standard formulas for black body radiation, the rest frame energy density, ρ, and entropy density, s, are given in terms of the locally measured temperature, T, by

$$\rho = bT^4 \tag{6}$$

$$s = \tfrac{4}{3}bT^3 \tag{7}$$

where b is a constant of order unity in Planck units $G = c = \hbar = k = 1$, assuming that the number of species of radiation is of order unity.[3] Equations (6) and (7) allow us to express the entropy density in terms of the energy density,

$$s = \alpha\rho^{3/4} \tag{8}$$

where $\alpha = \frac{4}{3}b^{1/4}$.

We expect that the configurations which extremize total entropy in a spher-

[3] b is proportional to the number of species, n, of radiation. (More precisely, b is proportional to $8n_B + 7n_F$, where n_B and n_F are, respectively, the number of helicity states of bosons and fermions.) Hence, α in equation (8) is proportional to $n^{1/4}$. Therefore, according to our formulas, by allowing $n \to \infty$ we can get arbitrarily large entropy density to energy density ratios and can violate the Bekenstein limit equation (1). However, for very large n the Casimir energy may become appreciable and invalidate our formulas. Nevertheless for large n we still could construct states not confined by box walls (and, hence, with no Casimir energy) which initially are confined to a region of size R but have S/E larger than the Bekenstein limit. By using massive particles, the rate of spreading of such a system can be made small.

ical box of radius R will be spherically symmetric. (Indeed, in general we could not even define the notion of a "spherical box" for nonspherical solutions.) Thus, we shall restrict attention only to spherically symmetric configurations. It should be emphasized, however, that we have not proven that there are no non-spherical extrema of entropy, and, indeed, we would not expect a proof of this statement (or a counterexample to it) to be easier to obtain than a proof of (or counterexample to) the conjecture that all static fluid stars in general relativity must be spherical.

The extrema of S for spherically symmetric perfect fluid stars have been studied by Cocke [6], who concluded that they correspond to equilibrium configurations. However, Cocke *assumed* that the space-time metric was static and used the Einstein constraint equations as well as the radial–radial component of Einstein's equation (though it was not necessary to use this additional equation [7]) to derive the equation of hydrostatic equilibrium by varying S. We wish to *derive* the fact that extrema of S correspond to static configurations and thus cannot assume *a priori* that the space-time metric is static. Thus, the assumptions made in our derivation below are weaker than Cocke's in that staticity of the metric is not assumed. On the other hand, Cocke treated an arbitrary perfect fluid while we consider only radiation. In addition, there are a number of other relatively minor differences between our discussion below and Cocke's, e.g., he imposes boundary conditions relevant to a star, whereas we impose boundary conditions relevant to a gas confined by a box.

We begin our analysis of spherically symmetric, radiation-fluid extrema of S by arguing that all such configurations must describe a "moment of time symmetry," or must at least represent a possibly non-time-symmetric moment in a space-time which possesses a moment of time symmetry. By a "moment of time symmetry" we mean that the extrinsic curvature, K_{ab}, of the hypersurface representing that instant of time vanishes. By the Einstein constraint equations, this also implies that the radiation-fluid 4-velocity must be orthogonal to the hypersurface. Physically, we expect extrema of entropy to be configurations of time symmetry because mass motion of the fluid should cost energy but contribute no entropy. For a non-time-symmetric configuration, it should be possible to convert the energy of mass motion implied by nonvanishing K_{ab} into thermal energy, thereby producing a nearby configuration of higher entropy. The following argument substantiates this physical expectation.

Given a solution of the constraint equations (3), (4) we can use the Einstein evolution equations with a $P = \rho/3$ fluid stress tensor, equation (5), to produce a solution of Einstein's equation

$$G_{ab} = 8\pi T_{ab} \tag{9}$$

Conservation of stress energy, $\nabla^a T_{ab} = 0$, implies that for this solution the entropy density 4-vector,

$$s^a = su^a \tag{10}$$

[where s is given by equation (8)] is conserved,

$$\nabla_a s^a = 0 \tag{11}$$

This means that the total entropy

$$S = \int s^a n_a \, d\Sigma \tag{12}$$

is independent of hypersurface Σ, where n^a denotes the normal to Σ. Since the space-time under consideration is asymptotically flat we expect–although this has not been proven–that there will exist an asymptotically flat, maximal (i.e., $K = 0$) hypersurface. On such a hypersurface, the "kinetic term" $K^{ab}K_{ab} - K^2 = K^{ab}K_{ab}$ in equation (4) is positive definite. Thus, if $K_{ab} \neq 0$, we can modify the solution by keeping h_{ab} fixed, scaling down K_{ab} and J_a by a constant factor, and increasing μ to compensate for the decrease in $K^{ab}K_{ab}$ in equation (4). By increasing μ and decreasing J_a, we increase the entropy density $s^a n_a$ of the radiation. Hence, the resulting solution will have the same total mass (since the metric h_{ab} is unchanged) but higher total entropy. By evolving these new data, we can obtain a configuration of higher entropy nearby our original configuration. Thus, for an extremum of entropy, we need $K_{ab} = 0$ on a $K = 0$ hypersurface. This shows that–modulo our ability to find a $K = 0$ hypersurface in the evolved spacetime–a necessary condition for a configuration to be an extremum of entropy is that it belongs to a space-time which has a time symmetric hypersurface ($K_{ab} = 0$).

If the evolved space-time of our extremum configuration has a $K = 0$ hypersurface as we have assumed, then in a neighborhood of this hypersurface it must be foliated by other $K = 0$ hypersurfaces [8]. Our argument above requires for an extremum of S that $K_{ab} = 0$ on each such hypersurface, which implies that the space-time is static. However, we shall now prove this fact directly by solving for all spherically symmetric, time-symmetric configurations which extremize S at fixed M with respect to all variations which preserve spherical symmetry and time symmetry. By doing so, we shall prove that the necessary *and* sufficient condition for an extremum is that it be data for a static, equilibrium configuration of a $P = \rho/3$ radiation-fluid.

For time symmetric data, the initial value constraint equations become simply

$$^{(3)}R = 16\pi\rho \tag{13}$$

We can write our spherically symmetric spatial metric, h_{ab}, in the form

$$ds^2 = g_{rr}\, dr^2 + r^2\, d\Omega^2 \tag{14}$$

In terms of g_{rr}, we have

$$^{(3)}R = \frac{2}{r^2}\frac{d}{dr}\left[r(1 - g_{rr}^{-1})\right] \tag{15}$$

Equation (13) has the unique solution regular at the origin,

$$g_{rr} = \left[1 - \frac{2m(r)}{r}\right]^{-1} \tag{16}$$

where

$$m(r) = 4\pi \int_0^r \rho(r')\,(r')^2\, dr' \tag{17}$$

The total mass M of the solution is determined by the asymptotic behavior of g_{rr} as $r \to \infty$. We have

$$M = \lim_{r \to \infty} m(r) = m(R) \tag{18}$$

where we neglect the energy density of the confining box and assume that no matter lies outside the box. On the other hand, the total entropy is given by

$$S = \int s^a n_a \, d\Sigma$$

$$= 4\pi\alpha \int_0^R \rho^{3/4} \left[1 - \frac{2m(r)}{r}\right]^{-1/2} r^2\, dr \tag{19}$$

Thus, our task is to extremize the integral

$$I = \int_0^R \left(\frac{1}{r^2}\frac{dm}{dr}\right)^{3/4} \left[1 - \frac{2m(r)}{r}\right]^{-1/2} r^2\, dr \tag{20}$$

with respect to all variations $\delta m(r)$ which vanish at the end points: $\delta m(R) = 0$ to keep M fixed; $\delta m(0) = 0$ for regularity at the origin. Hence, the condition for $m(r)$ to be an extremum is simply that it satisfy the Euler–Lagrange equations,

$$\frac{d}{dr}\left(\frac{\partial L}{\partial m'}\right) - \frac{\partial L}{\partial m} = 0 \tag{21}$$

for the Lagrangian,

$$L = (m')^{3/4} \left(1 - \frac{2m}{r}\right)^{-1/2} r^{1/2} \tag{22}$$

where $m' = dm/dr$.

Equation (21) yields

$$-\tfrac{3}{16} m'' r^2 + \tfrac{3}{8} m'' m r + \tfrac{3}{8} m' r - \tfrac{1}{4} (m')^2\, r - \tfrac{3}{2} m' m = 0 \tag{23}$$

This equation can be put in more recognizable form by substituting $m' = 4\pi r^2 \rho$ everywhere. We obtain

$$\frac{d}{dr}(\rho/3) = -\frac{(\rho + \rho/3)\,[m(r) + 4\pi r^3 \rho/3]}{r[r - 2m(r)]} \tag{24}$$

Equation (24) is precisely the general relativistic Tolman-Oppenheimer-Volkoff equation of hydrostatic equilibrium for a perfect fluid with $P = \rho/3$. Thus, the requirement that S be an extremum yields precisely the condition that the radiation fluid be in static equilibrium.

Thus, we have shown that the static, spherically symmetric equilibrium configurations of the radiation extremize S with respect to variations which preserve spherical symmetry and time symmetry. We now point out that the restriction to spherically symmetric and time symmetric variations is not necessary; the equilibrium configurations extremize S with respect to *all* variations which keep M fixed. Namely, if we found a nonspherical first order variation for which $\delta M = 0$ but $\delta S \neq 0$, we could average this variation over the rotation group. By symmetry, each of the rotated perturbations would have the same $\delta M = 0$ and $\delta S \neq 0$, so by averaging over the rotation group we would produce a spherical variation with $\delta M = 0$ and $\delta S \neq 0$. Thus, the absence of such spherical variations also implies the absence of nonspherical variations. Similarly, if a non-time-symmetric variation changed S at fixed M, a time symmetric variation with the same $\delta \rho$ and δh_{ab} but with $\delta J_a = \delta K_{ab} = 0$ would also satisfy the varied constraint equations, since the variation of equation (3) would be trivially satisfied while the variation of equation (4) would remain satisfied since δK_{ab} does not enter that equation because $K_{ab} = 0$ in the background geometry. On the other hand, the variations of S, equation (12), and M, equation (18) depend only on $\delta \rho$ and δh_{ab}, so the time symmetric variation would yield the same $\delta S \neq 0$ with $\delta M = 0$. Thus, the absence of such time symmetric variations also implies the absence of such non-time-symmetric variations.

In summary, we have proven that all spherically symmetric, static equilibrium configurations of a $P = \rho/3$ radiation fluid are extrema of total entropy. Conversely, if we assume that all extrema are spherically symmetric and that all fluid space-times of the type considered contain a maximal hypersurface, then these equilibrium configurations are the only extrema of S.

It is worth noting that in ordinary (i.e., non-general-relativistic) dynamics, there is good reason to expect a connection between equilibrium configurations and extrema of S, specifically that local maxima of S should represent stable equilibrium configurations. This is because in ordinary dynamics the entropy S of a configuration is expected to measure the logarithm of the fraction of time the system spends throughout its dynamical history "looking like" the given configuration [9]. Hence, if the entropy is low, the system should change its macroscopic appearance to a state of higher entropy on a relatively short time scale. Since the macroscopic appearance of a system in a state of maximum entropy should not change over long time scales, such systems should appear to be in stable equilibrium and thus should represent stable equilibrium solutions of the macroscopic equations used to describe the system. It is interesting that this argument cannot be made *globally* in general relativistic dynamics for the simple reason that there is no preferred time slicing of space-time and thus the concept of the "fraction of time" spent by a system in a given state does not have any obvious meaning. Indeed, this is one of the main difficulties involved in understanding the nature of black hole thermodynamics. However, in the case of matter configurations, the above argument can be made *locally:* the entropy of each local element of matter should not decrease with time, i.e., $\nabla_a s^a \geqq 0$. But this implies the total matter entropy, $S = \int s^a n_a \, d\Sigma$, should not decrease with time, and thus a configuration of maximum S should be in stable equilibrium. In the case of perfect fluid radiation, the macroscopic equations actually preserve S, as already noted above, equation (11). Thus, local maxima of S must be dynamically stable because there is no nearby state to which they can evolve via the macroscopic equations. Above, we showed that extrema of S correspond to equilibrium configurations. In the next section, we shall show that local maxima of S correspond precisely to locally stable equilibrium configurations.

§(3): *Spherical Equilibrium Configurations*

In this section we study in more detail the equilibrium states of our system, which comprise a two-dimensional set parameterized for example by total mass M and box radius R. According to the above discussion a spherically symmetric equilibrium state—or equivalently a spherical extremum of the entropy S at fixed M—corresponds to a solution of equation (23) for which $m(0) = 0$, $m(R) = M$. [Regularity at the origin demands further that at $r = 0$, $m(r) \sim 4\pi/3 \, r^3 \rho(0) \sim r^3$; elementary flatness only that $m/r \to 0$; see (14), (16), (17).] Because (23) possesses a scale invariance we can simplify it by introducing the dimensionless variables

$$\mu := m(r)/r \tag{25}$$

$$q := dm/dr = 4\pi r^2 \rho \tag{26}$$

$$z := \ln r \tag{27}$$

in terms of which (23) becomes

$$\frac{dq}{dz} = \frac{2q(1 - 4\mu - \frac{2}{3}q)}{1 - 2\mu} \tag{28a}$$

and this together with the identity

$$\frac{d\mu}{dz} = q - \mu \tag{28b}$$

yields a pair of first-order equations equivalent to the second-order equation (23).

To discover the "qualitative" character of (28) it helps to plot the solutions as curves in the μ-q plane (Figure 1). Eliminating dz from (28) gives the slope at each point of this plane:

$$(1 - 2\mu)(q - \mu)\,dq = 2q(1 - 4\mu - \tfrac{2}{3}q)\,d\mu \tag{29}$$

Thus a solution curve will be vertical where it crosses the line $\mu = q$ (α in the figure) and horizontal where it crosses the line $4\mu + \frac{2}{3}q = 1$ (β in the figure). The curve of interest to us begins at the origin with slope 3 [since $q/\mu \sim 4\pi r^2 \rho/(m/r) \sim 3$; alternately it can be verified directly that in the neighborhood of $\mu = q = 0$ the general solution of (29) is $(3\mu - q)^2\, q \approx$ const, whence $q \sim 3\mu$ for the unique solution passing through the origin]. By (28) both dq/dz and $d\mu/dz$ are positive, and must remain so until the solution curve crosses β, at which point the numerator of (28a) changes sign and the curve turns downward, but continues in a rightward direction (with z remaining a good parameter). Similar reasoning shows that the curve proceeds to spiral around the point $\mu = q = \frac{3}{14}$ (0 in the figure). Near this point one can expand (28) in terms of $x = \mu - \frac{3}{14}$, $y =$

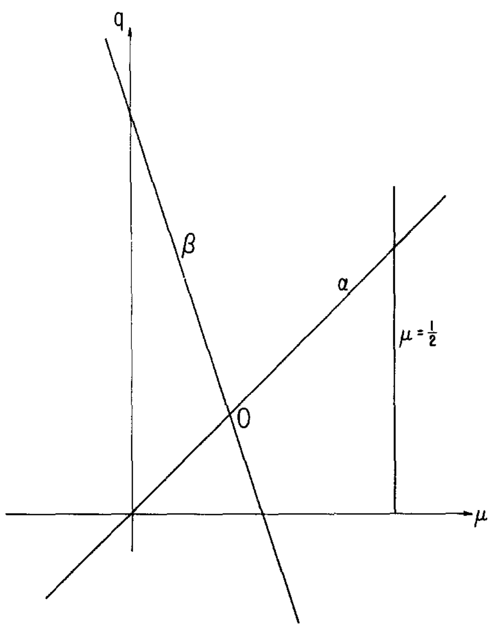

Fig. 1. The μ–q plane. Solutions intersect the line α ($\mu = q$) vertically and the line β ($4\mu + 2q/3 = 1$) horizontally in such a way as to spiral around the point O ($\mu = q = 3/14$).

$q - \frac{3}{14}$, obtaining

$$\frac{d}{dz}\begin{pmatrix} x \\ y \end{pmatrix} \approx \begin{pmatrix} -1 & 1 \\ -3 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

which is a linear system with eigenvalues $\nu_\pm = -\frac{3}{4} \pm i\sqrt{47}/4$. It follows that x and y behave asymptotically like linear combinations of the real and imaginary parts of $e^{\nu_\pm z}$, i.e., of $r^{-3/4}\sin(\frac{1}{4}\sqrt{47}\ln r)$ and $r^{-3/4}\cos(\frac{1}{4}\sqrt{47}\ln r)$ [10]. In particular r increases monotonically to ∞ as the solution curve approaches the singular solution $\mu = q = \frac{3}{14}$. Figure 2 depicts a numerically obtained approximation to the solution curve C of interest to us. Notice that $\mu = m(r)/r$ remains always $\leqslant 0.25$ for equilibria regular at $r = 0$. Thus radiation cannot attain equilibrium if confined to an R less than about twice its own Schwarzschild radius.

Consider now a fixed radius R and the 1-parameter sequence of spherical equilibria confined to a box of this radius. Each such equilibrium corresponds to an initial segment of the curve C in Figure 2, beginning with $r = 0$ at $\mu = q = 0$ and terminating with $r = R$ at a point $\mu = M/R$, $q = 4\pi R^2 \rho(R)$. Conversely, because of the scale invariance of equations (28) each such initial segment of C (equivalently the end point of such a segment) corresponds to a unique equilibrium with radius R, obtained from any homologous equilibrium solution by the appropriate choice of λ in

$$\begin{aligned} r &\longrightarrow \lambda r \\ m &\longrightarrow \lambda m \\ \rho &\longrightarrow \lambda^{-2}\rho \end{aligned} \tag{30}$$

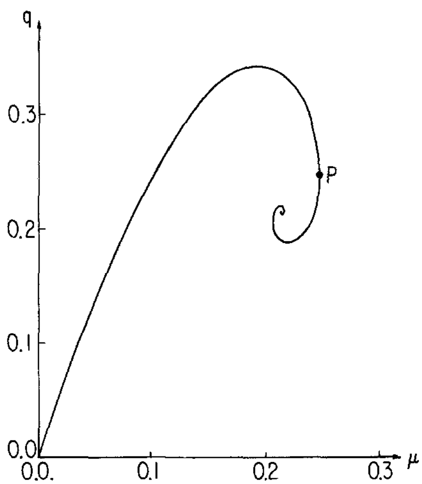

Fig. 2. The curve, C, corresponding to equilibria regular at the origin. P is the turning point. The exterior Schwarzschild metrics correspond to the piece of the μ axis between 0 and 1/2.

In discussing the stability of our solutions it is the interpretation of the points of C as representing distinct equilibria at fixed R which will be appropriate, rather than our original interpretation according to which distinct points of C represent distinct values of r in a single equilibrium solution. With this in mind let us settle also on a fixed parametization of C, which we can do by employing that solution ζ, of

$$d\zeta = d\mu/(q - \mu)$$

[cf. (28b)] which vanishes at the point of maximum μ (point P in Figure 2).

In the present notation, our earlier expression (19) for the entropy can be written as

$$S = a \int_0^R q^{3/4} r^{1/2} (1 - 2\mu)^{-1/2} \, dr \tag{31}$$

where $a = (4\pi)^{1/4} \alpha$. In the first place it is clear from this that under the rescaling (30) $S \sim \lambda^{3/2}$ so that, for constant ζ (i.e., for a homologous family of solutions) $S \sim R^{3/2} \sim M^{3/2}$. This means that for $M \sim R$, S will be far less than the entropy ($\sim M^2$) of the corresponding black hole. However, it is not necessary to talk in order-of-magnitude terms since we can actually evaluate S exactly in terms of μ and q.

To that end—and also because it will be needed for our subsequent analysis of stability—let us introduce a parameter β conjugate to M with respect to S. For equilibrium configurations we define β by

$$dS = \beta \, dM \tag{32}$$

where the variations are carried out at fixed R, and the definition is consistent because at equilibrium $dS = 0$ when $dM = 0$, as we have seen. Putting $q = dm/dr$ in (31) and repeating the variation performed in Section 2, but without setting $\delta m(R) = 0$, produces

$$\delta S = \int_0^R \delta S/\delta m \, \delta m \, dr + \beta \delta m(R)$$

where $\delta S/\delta m = 0$ for equilibria [i.e., eq. (23)] and we find

$$\beta = \tfrac{3}{4} a r^{1/2} q^{-1/4} (1 - 2\mu)^{-1/2} \big|_{r=R}$$

If $T = T(R)$ is the local temperature at $r = R$, then, by (26), (6), and the definitions $a = (4\pi)^{1/4} \alpha$ and $\alpha = \frac{4}{3} b^{1/4}$,

$$\beta = T^{-1} (1 - 2M/R)^{-1/2}$$

Thus β^{-1} is the surface temperature $T(R)$ corrected by a "red-shift factor" $(1 - 2M/R)^{1/2}$ relating the surface $r = R$ to spatial infinity.[4]

Correcting the work $p\,dV$ (which is performed at radius R) by this same factor we can express the first law of thermodynamics as

$$dM = \beta^{-1}\, dS - (1 - 2\mu)^{1/2}\, p\, dV \tag{33}$$

where $p = \frac{1}{3}\rho(R)$, $\mu = \mu(R)$ and where $dV = 4\pi R^2\, dR(1 - 2\mu)^{-1/2}$ is the change in volume due to an alteration dR in the radius of the confining box. Recall now that $M = \mu R$ and that [by (31) and (30)] we can write $S = S(M, R) = \sigma R^{3/2}$, where $\sigma = \sigma(\zeta(R))$ depends only on the similarity class of the configuration in question. For a variation in which this class does not change, (33) becomes

$$\mu\, dR = (1 - 2\mu)^{1/2}\, T\sigma d(R^{3/2}) - (\rho/3)(4\pi R^2\, dR)$$

which can be solved for σ (since dR drops out) with the result

$$S = \sigma R^{3/2} = \frac{a}{6}\, \frac{q + 3\mu}{q^{1/4}(1 - 2\mu)^{1/2}}\, r^{3/2}\bigg|_{r=R} \tag{34}$$

[This formula can be confirmed directly by comparison with (31): defining $f(r)$ so that (34) reads $S = af(r)|_{r=R}$, one verifies easily that for solutions $\mu(r)$, $q(r)$ of (28), $df(r)/dr$ coincides with the integrand of (31).] Note that for M, R at all large (compared to Planck scales) $S/(MR) = S/(\mu R^2) = \sigma/\mu\, R^{-1/2} \ll 1$ when q and μ lie on C; hence the Bekenstein limit is satisfied by a large margin.

To conclude this section we discuss briefly the stability of our equilibrium configurations with respect to *spherical* perturbations. Let us call an equilibrium *dynamically* stable when small perturbations about it remain small for all time and *thermodynamically* stable when it is a strict local maximum of the entropy at fixed total energy M. At the end of Section 2 we argued that thermodynamic stability should imply dynamical stability. Conversely, it is plausible that an unstable equilibrium will show at least *secular* instability to any dissipative mechanism which would allow it to reach the nearby configurations of higher entropy; but in general this is all that can be said. In the present case, however, we will establish directly both the above implication and its converse: thermodynamic and dynamical stability coincide [6].

The equation of motion for a perturbation, δm, of $m(r)$ about its equilibrium value [with $\delta m(0) = 0 = \delta m(R)$] is

$$k(r)^2\, \delta \ddot{m} + L \cdot \delta m = 0 \tag{35}$$

[4] Properly it is $e^{\nu} = |g_{00}|^{1/2}$ and not $e^{-\lambda} = (1 - 2m/r)^{1/2}$ which is the red-shift factor. However, one can think of $e^{-\lambda}|_{r=R}$ as the red shift between $r = R$ and $r = \infty$ which *would* apply if a Schwarzschild metric (for which $\nu = -\lambda$) prevailed for $R \leqslant r < \infty$. Such a factor appears because $m(r)$ is not a "local energy" but an energy "as registered at ∞."

where a superior dot is a time derivative,

$$k(r)^2 = \tfrac{3}{16}\, e^{-\nu} r^{1/2} q^{-5/4} (1 - 2\mu)^{-3/2} \tag{36}$$

and for any function $\psi(r)$

$$L \cdot \psi = (A\,\psi')' + C\psi \tag{37}$$

with

$$\begin{aligned} A(r) &= \tfrac{1}{16}\, r^{1/2} q^{-5/4} (1 - 2\mu)^{-1/2} \\ C(r) &= \tfrac{1}{12}\, r^{-3/2} q^{-1/4} (1 - 2\mu)^{-5/2}\ (3 + 2q) \end{aligned} \tag{38}$$

[This can be derived by straightforward perturbation of the $\binom{\mu}{\nu} = \binom{0}{0}, \binom{1}{1}$, and $\binom{1}{0}$ components of $G^{\mu}_{\nu} = 8\pi T^{\mu}_{\nu}$ and of the $\nu = 1$ component of $\nabla_{\mu} T^{\mu}{}_{\nu} = 0$: For the metric one uses $ds^2 = -e^{\nu} dt^2 + (1 - 2m/r)^{-1}\, dr^2 + r^2\, d\Omega^2$ and perturbs ν and m about their equilibrium values. For $T^{\mu}{}_{\nu}$ one uses (5) with $u^2 = u^3 = 0$ and perturbs ρ and μ^1 while maintaining the relations $\rho = 3p$, $u^{\alpha}u_{\alpha} = -1$. Finally one eliminates $\delta\rho$, $\delta\nu$, and u^1 to obtain an equation for δm alone, whose equivalence with (35) follows immediately by virtue of equations (28) for the equilibrium values of q, μ.] Clearly (35) (with $\delta m \equiv \psi$) is the Euler-Lagrange equation for the Lagrangian

$$\mathcal{L}(\psi) = \int_0^R k(r)^2\, \dot{\psi}^2\, dr - V(\psi) \tag{39}$$

$$V(\psi) = \int [A(r)\, \psi'(r)^2 - C(r)\, \psi(r)^2]\, dr \tag{40}$$

and the boundary conditions $\psi(0) = \psi(R) = 0$. (Notice that variations of ψ correspond to *second* variations of m.) Since the "kinetic term" $\int k^2\, \dot{\psi}^2\, dr$ is positive definite, energy arguments based on the analogy with a ball rolling in a potential V lead to the conclusion that dynamical stability is equivalent to the positivity of $V(\psi)$ [11].

The condition for thermodynamic stability is that the second variation, $\delta^2 S$, be negative definite in δm, for $\delta m(R) = 0$.[5] The second variation of (31) is

$$\begin{aligned} \delta^2 S = 3a \int \Big[&\tfrac{-1}{16}\, r^{1/2} q^{-5/4} (1 - 2\mu)^{-1/2} (\delta m')^2 \\ &+ \tfrac{1}{2}\, r^{-1/2} q^{-1/4} (1 - 2\mu)^{-3/2}\, \delta m \delta m' + r^{-3/2} q^{3/4} (1 - 2\mu)^{-5/2}\, \delta m^2 \Big]\, dr \end{aligned}$$

[5] Strictly speaking weak positivity ($\delta^2 S \geqslant 0$) is *necessary* for stability while strict positivity ($\delta^2 S > 0$ if $\delta m \neq 0$) is sufficient. When $\delta^2 S$ is positive but not strictly positive, stability will be determined by higher variations of S. Analogous remarks apply for dynamical stability, which, in particular, cannot always be decided by reference to the first variation of the equations of motion.

Integrating the second term by parts and assuming $\delta m \to 0$ at $r = R$ and (sufficiently rapidly) at $r = 0$ yields, with use of (28),

$$\delta^2 S = 3a \int [-A(r)(\delta m')^2 + C(r)(\delta m)^2]\, dr$$

$$= -3aV(\delta m) \tag{41}$$

We see that $\delta^2 S$ is negative iff V is positive, so that, as claimed, a single condition governs both dynamical and thermodynamical stability. Notice, incidentally that, because $\delta S = 0$ in equilibrium we can also express our conclusion by saying that the entropy S serves as a "potential energy" in the Lagrangian (39).

Where along the curve C of Figure 2 will $\delta^2 S$ be positive definite? On physical grounds a very dilute black-body gas in a box is obviously stable against gravitational collapse; hence $\delta^2 S$ is positive near $\mu = q = 0$ (i.e., for $\zeta \to -\infty$).[6] The standard theory of stellar stability then tells us [12]—or at least suggests strongly—that the equilibria along C continue to locally maximize S until the first maximum of M, or equivalently of μ (recall we have fixed R), whereupon instability sets in iff a plot of the equilibria in the β-M plane turns in a *clockwise* direction (with μ plotted upward and β to the right) and continues until the first *counter*-clockwise turning point.[7] Given all this, consider the turning point $\zeta = 0$, in the μ-q plane. At this point $d\mu/d\zeta = 0$, whence $d\beta/d\zeta$ and $d(q^{-1})/d\zeta$ have the same sign since $\beta = (1 - 2\mu)^{-1/2}\, T^{-1}$ and $q \propto T^4$. Thus the condition $d\beta/d\zeta\; d^2\mu/d\zeta^2 < 0$ for clockwise turning in the β-μ plane is equivalent to the condition $dq/d\zeta\; d^2\mu/d\zeta^2 > 0$ for clockwise turning in the μ-q plane; and $\zeta = 0$ is in fact a point of instability onset. The largest value of M/R attainable by a stable configuration of self-gravitating radiation in a spherical box is thus the same as the largest value attainable by any equilibrium: $\mu_{\max} \approx 0.246$. For such a configuration, the entropy as given by (34) with $\mu = q \approx 0.246$ is

$$S_{\max} \approx 0.327 a R^{3/2}$$

[6] Analytically stability means that, for ζ sufficiently small, the A term in (40) always dominates the C term. This in turn can be shown to follow (for perturbations of finite central pressure) from the inequality

$$\int_0^1 \psi^2 x^{-2}\, dx \leqslant \frac{4}{9} \int_0^1 (\psi')^2 x^{-2}\, dx$$

which holds for all ψ vanishing at $x = 0$ and vanishing faster than $x^{3/2}$ at $x = 0$. (To prove it substitute $\psi = u x^{3/2}$.)

[7] Some of these assertions are more definitely reliable than others when, as here, the equilibria belong to an infinite-dimensional manifold of configurations. In particular the proof that S fails to maximize just beyond a clockwise turning point can be carried out in any Banach manifold [13]. Also in our case instability can never *disappear* as one moves along C; for any configuration for $\zeta_2 > \zeta_1$ contains a homologous replica of the ζ_1 configuration and thus must be unstable if the latter is.

Finally we remark that solution curves other than C are easily computed and may have physical interest. Although they do not describe configurations regular at $r = 0$, they have portions $r_{\min} < r < R$ which can be "joined on" to other forms of matter for $r < r_{\min}$ to produce everywhere-regular static solutions of Einstein's equations. (c.f. [14].) *A priori*, such solutions might provide much higher entropies than any *equilibrium* configuration of pure radiation and the same M and R. This does not seem to be true if one requires the auxiliary matter to obey an energy condition, but even if it were, we would expect (Section 2) to find pure-radiation configurations of still higher entropy. In fact pure radiation—but nonequilibrium—configurations exist with arbitrarily great entropy, and it is to the consideration of such configurations that we now turn.

§(4): *Maximum Entropy Configurations*

In the previous section we found all the equilibrium configurations of a radiation fluid of mass M in a spherical box of radius R. Under the assumptions stated in Section 2, these yield all the local extrema—in particular, all the local maxima—of S. However, it is certainly possible that S can take on values higher than given by its local maxima. To see that this is indeed the case and that arbitrarily large values of S can be achieved, consider the time symmetric initial data whereby a "Friedmannian" region of uniform density ρ is joined to a Schwarzschild exterior of mass M, as illustrated in Figure 3.

The initial value constraint, equation (13), requires that, in the Friedmann region,

$$\rho = \frac{3}{8\pi a^2} \tag{42}$$

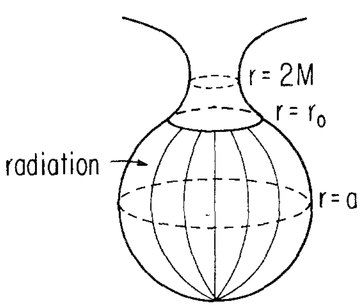

Fig. 3. A Friedmann 3-sphere of radiation joined onto the Schwarzschild exterior of mass M at a moment of time symmetry. By letting $a \to \infty$, we can make the total entropy arbitrarily large.

where a is the "radius" of the Friedmann 3-sphere. Matching the geometry across the surface $r = r_0$ requires

$$\tfrac{4}{3}\pi r_0^3 \rho = M \tag{43}$$

For the configuration shown in Figure 3 we also must have

$$2M < r_0 < a \tag{44}$$

a condition which follows from equations (42) and (43) provided $a > 2M$. The total entropy is

$$S = \alpha \rho^{3/4} 2\pi a^3 (\pi - \theta_0 + \tfrac{1}{2} \sin 2\theta_0) \tag{45}$$

where $\theta_0 = \sin^{-1} (r_0/a)$. Thus, if for fixed M we let $a \to \infty$, let $\rho \to 0$ via equation (42) and vary r_0 according to equation (43), we find that

$$S \sim a^{3/2} \longrightarrow \infty \tag{46}$$

Thus, arbitrarily high entropies can be achieved for any given value of M (and any $R > 2M$) for configurations of the type shown in Figure 3.

However, a configuration of this type cannot be built classically without starting from a white hole. If we evolve the configuration of Figure 3 backward into the past–using only energy inequalities on the matter, but making no further assumptions about its properties–we are inevitably led to a white hole, just as evolution into the future inevitably leads to a black hole. This can be seen most easily in our example by examining the domain of dependence of the "Schwarzschild part" of the initial data surface. Since $T_{ab} = 0$ on this part of the surface, conservation of T_{ab} and the energy condition $T_{ab} v^a v^b \geqq 0$ for all timelike v^a imply that $T_{ab} = 0$ throughout its domain of dependence [15]. The vacuum Einstein evolution equations then imply that the space-time must be Schwarzschild in the entire domain of dependence of the "Schwarzschild part" of the initial data surfaçe. This is sufficient to prove the existence of an initial white hole and white hole singularity. More generally, the existence of a past trapped surface on the initial data surface, together with the energy condition $(T_{ab} - \frac{1}{2} T g_{ab}) v^a v^b \geqq 0$ for all timelike v^a, imply the existence of a singularity [15] to the past of the initial surface. This means that starting from nonsingular initial data with no past trapped surfaces we cannot construct a configuration of the type illustrated in Figure 3. There is reason to believe that white holes cannot exist [5]. If so, then the configuration of Figure 3 cannot exist either. The question remains as to what is the maximum entropy configuration that can be achieved starting from physically reasonable initial conditions, i.e., nonsingular initial data with no past trapped surfaces.

As in Section 2, we shall continue to restrict consideration to spherically symmetric, time symmetric configurations, as these should maximize the entropy. For such configurations there is a simple condition which is equivalent to the nonexistence of past (or future) marginally trapped surfaces, namely,

$$r > 2m(r) \tag{47}$$

i.e., at each r the configuration must be outside its own Schwarzschild radius. How much entropy can a configuration satisfying equation (47) achieve? If there is no limit as to how close to equality one can get in equation (47), again arbitrarily large total entropies can be attained. This is illustrated in Figure 4. Putting matter in the $r \approx 2m$ "throat" of such a configuration costs essentially no energy but yields a finite entropy. By "stretching out" this throat we can achieve arbitrarily large entropy to energy ratios. This can be seen more clearly as follows. We define $\epsilon(r) > 0$ by

$$\epsilon = r - 2m(r) \tag{48}$$

The total entropy, S, is

$$\begin{aligned} S &= \alpha \int_0^R \rho^{3/4} \frac{r^{1/2}}{\epsilon^{1/2}} r^2 \, dr \\ &= \alpha \left(\frac{1}{8\pi}\right)^{3/4} \int_0^R \left(1 - \frac{d\epsilon}{dr}\right)^{3/4} \frac{1}{r^{3/2}} \frac{r^{1/2}}{\epsilon^{1/2}} r^2 \, dr \\ &= \alpha \left(\frac{1}{8\pi}\right)^{3/4} \int_0^R \left(1 - \frac{d\epsilon}{dr}\right)^{3/4} \frac{r}{\epsilon^{1/2}} \, dr \end{aligned} \tag{49}$$

On the other hand, the total mass is,

$$M = m(R) = \tfrac{1}{2} \left[R - \epsilon(R)\right] \tag{50}$$

Thus, by letting $\epsilon \to 0$ (keeping $d\epsilon/dr \ll 1$) we can make $S \to \infty$ while keeping M finite.

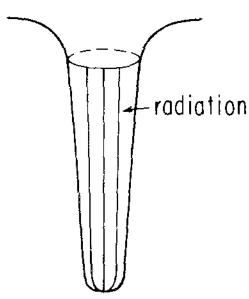

Fig. 4. A radiation configuration which is just barely outside its own Schwarzschild radius, i.e., $r = 2m(r) + \epsilon(r)$ with $\epsilon \ll r$. By letting $\epsilon \to 0$ we can make the total entropy arbitrarily large.

However, it does not seem reasonable that one can produce configurations which are arbitrarily close to being at their own Schwarzschild radius. Such configurations could arise (without initial white holes) only with the aid of a second system (an "external agent"), e.g., with the use of strings, boxes, etc., which assembles the radiation into the required configuration and then releases it. However, the thermal radiation of energy density ρ is composed of constituents of finite size λ given by

$$\lambda \sim \rho^{-1/4} \tag{51}$$

It does not seem plausible that one could *assemble* a configuration which is closer to its own Schwarzschild radius than the size of its constituents. This suggests that it will be physically possible to produce only configurations which satisfy

$$D(r) \gtrsim \lambda(r) \tag{52}$$

where $D(r)$ denotes the proper distance of the radiation from its Schwarzschild radius and $\lambda(r)$ denotes a typical proper wavelength of the radiation at that radius. In addition, we note that the quantum fluctuations in the energy density of radiation in a region of size λ are of the same order as the energy density itself. Hence, for configurations which violate equation (52), the validity of our classical analysis, which ignores fluctuations, is questionable. Thus, even if equation (52) did not provide a cutoff for physically producible configurations, it might well provide a cutoff for the validity of our classical description of them and thus for the validity of our calculation of their entropy.

Since without the limitation of equation (52), we obtain $S \to \infty$ as $\epsilon \to 0$, we expect that with the limitation of equation (52) the highest entropies will be achieved in the strong field regime $\epsilon \ll r$. (In the opposite limit of weak gravitational fields, the Bekenstein analysis applies.) For $\epsilon \ll r$, the relation between ϵ and D is

$$\begin{aligned} D &= g_{rr}^{1/2}\,\epsilon \\ &= r^{1/2}\,\epsilon^{1/2} \end{aligned} \tag{53}$$

since $g_{rr} = (1 - 2m/r)^{-1} = r/\epsilon$. Using equations (51) and (53) we find that equation (52) becomes

$$\epsilon^{-1/2} \lesssim r^{1/2}\,\rho^{1/4} \tag{54}$$

Substituting the inequality in the first line of equation (49), we obtain the following bound for S:

$$\begin{aligned} S &\lesssim \int_0^R \rho r^3\, dr \\ &\leqq R \int_0^R \rho r^2\, dr \sim RM \end{aligned} \tag{55}$$

Thus, in the strong field regime $\epsilon \ll r$ we find the limit

$$S/M \lesssim R \tag{56}$$

in agreement with the Bekenstein limit. Furthermore, this limit can be achieved. For example, a configuration with $\rho = (8\pi r^2)^{-1}$ for $r > 1$ (i.e., $r > l_P$, where $l_P \approx 10^{-33}$ cm is the Planck length) and $\rho = 0$ for $r < 1$ will have $\epsilon = 1$ and will satisfy equation (52). According to equation (49) its entropy will be

$$S \sim \int_0^R r\,dr = R^2/2$$

$$\sim MR \tag{57}$$

thus achieving the limit, equation (56).

Thus, if our condition, equation (52) on the achievability of configurations is correct, we have found that the maximum possible entropy of a precollapse configuration is bounded by the Bekenstein limit. Furthermore, since $\epsilon \ll r$ we find[8]

$$S_{\max} \sim MR \sim M^2 \sim S_{\text{bh}} \tag{58}$$

where S_{bh} denotes the entropy of a Schwarzschild black hole of mass M.

Now, $S_{\max}$ measures the number of internal states, N, of precollapse radiation configurations,

$$N \sim \exp(S_{\max}) \tag{59}$$

The equality of $S_{\max}$ with S_{bh} appears to indicate that S_{bh} may indeed measure the "number of internal states," N_{bh}, of a black hole,[9] as has been suggested by many authors. However, we have not been able to obtain a satisfactory argument relating N to N_{bh}. It is not true that the number of precollapse configurations must equal the number of black hole internal configurations (even assuming that the latter notion proves to be well defined). "Conservation of states" would require that the number of black hole states at a given "time" equal the

[8] Since these precollapse configurations have entropy comparable to that of a black hole and since, at least in ordinary (i.e., non-general-relativistic) dynamics, the entropy of a configuration measures the "fraction of time" the system spends in the same macroscopic state, it might appear that these precollapse configurations should be present as frequently as black holes. However, as discussed above, presumably an assembling agent is needed to produce the precollapse configuration without starting from a white hole. This assembling agent must be in a very special state (possibly involving a high degree of correlation with the radiation) in order to have produced this configuration, since if we "run the clock backwards" this agent must "catch" all the radiation before it collapses to a black hole. Thus, the entropy of the assembling agent should be very low, and the total entropy of the agent plus precollapse configuration should be much lower than that of an agent plus black hole configuration, where no special state of the agent is required.

[9] It should be stressed that the true meaing of the "number of internal black hole states" in general relativistic dynamics is unclear at present, just as the notion of the "fraction of time" a system spends in a given configuration does not have an obvious meaning.

number of states at an earlier time that evolve to the black hole at the given time. However, these states at an earlier time may include many black hole configurations as well as the precollapse configurations. Furthermore, since evolution of pure states to density matrices may occur, "conservation of states" might not apply; one could have a larger number of final states due to the "branching out" of the state vectors. These statements indicate that the number, N, of precollapse configurations still should be a lower bound for N_{bh}. However, even this conclusion is not obvious because, as mentioned above, we do not expect to be able to produce these high-entropy precollapse configurations without the assistance of a second "assembling agent" system. Therefore, the "conservation of states" (or, more properly, the "nondecrease of states") argument should be applied only to the joint black-hole–assembling-agent system. It is possible that when this is done, the formation of the black hole will result in the destruction of correlations between the assembling system and the precollapse configuration. If this happens, there will be a large effective increase in the "number of states" without the need for a large number of black hole internal states, and the argument giving N as a lower bound for N_{bh} fails. Thus, we have not been able to demonstrate a relationship between N and N_{bh}. Nevertheless, the suggestion remains that $N_{bh} \sim N \sim \exp(S_{bh})$.

Acknowledgments

We wish to thank Don Page for an enlightening phone conversation leading to equation (34). This research was supported in part by NSF grant PHY-78-24275 to the University of Chicago.

References

1. Penrose, R. (1979). In *General Relativity: An Einstein Centennary Survey,* eds. S. W. Hawking and W. Israel, Cambridge University Press; Page, D. (1981). *Gen. Rel. Grav.*, **13**, 1117.
2. Wald, R. M. (1980). *Phys. Rev. D*, **21**, 2742.
3. Gibbons, G. W., and Perry, M. J. (1978). *Proc. R. Soc. London Ser. A*, **358**, 467.
4. Bekenstein, J. (1981). *Phys. Rev.*, **D23**, 287.
5. Wald, R. M., and Ramaswamy, S. (1980). *Phys. Rev. D*, **21**, 2736.
6. Cocke, W. J. (1965). *Ann. Inst. Henri Poincaré*, **2**, 283.
7. Katz, J. (1972). *J. Phys. A*, **5**, 781.
8. Cantor, M., Fischer, A., Marsden, J., O'Murchadha, N., and York, J. (1976). *Commun. Math. Phys.*, **49**, 187.
9. Wald, R. M. (1979). *Phys. Rev. D*, **20**, 1271.
10. Chandrasekhar, S. (1972). In *General Relativity*, ed. L. O. Raifeartaigh, Clarendon Press, Oxford.
11. Friedman, J. L., and Schutz, B. F. (1975). *Astrophys. J.*, **200**, 204; Schutz, B. F. (1972). *Astrophys. J. Suppl.*, **24** (208), 319.
12. Poincaré, H. (1885). *Acta Math.*, **7**, 259; Lynden-Bell, D., and Wood, R. (1968). *Mon. Not. R. Astron. Soc.*, **138**, 495; Ipser, J. R., and Horwitz, G. (1979). *Astrophys. J.*, **232**, 863; Katz, J. (1978). *Mon. Not. R. Astron. Soc.*, **183**, 765; Thorne, K. S. (1967). In *High Energy Astrophysics*, eds. C. DeWitt, E. Schatzman, and P. Véron, Gordon and Breach, New York.
13. Sorkin, R. D. (1981). *Astrophys. J.*, **249**, 254.
14. Bondi, H. (1964). *Proc. R. Soc. London, Ser. A*, **282**, 303.
15. Hawking, S. W., and Ellis, G. R. F. (1973). *The Large Scale Structure of Spacetime*, Cambridge University Press.