

Generalized second law of thermodynamics in black-hole physics*

Jacob D. Bekenstein

Center for Relativity Theory, The University of Texas at Austin, Austin, Texas 78712

(Received 17 September 1973)

In previous work we introduced the concept of black-hole entropy, which we identified with the surface area of the black hole in question expressed in units of the Planck length squared. We suggested that the appropriate generalization of the second law for a region containing a black hole is that the black-hole entropy plus the common entropy in the black-hole exterior never decreases. Here we establish the validity of this law for the infall of an entropy-bearing system into a much larger and more massive generic stationary black hole. To do this we determine a general lower bound for the increase in black-hole entropy, and an upper bound for the entropy of the system, while allowing for quantum effects at each stage. In passing we show that the generalized second law is a statistical law which becomes overwhelmingly probable in the limit of a macroscopic system. We also consider briefly more general situations. Finally, we give two simple examples of predictions made by the generalized second law for black-hole formation processes.

I. INTRODUCTION

Black-hole physics mirrors thermodynamics in many respects. For example, Christodoulou¹ has shown that the efficiencies of processes for extracting energy from a black hole are limited by their irreversibility; the most efficient processes are the "reversible" ones. This result has a clear thermodynamic ring to it which has been well brought out by Carter.² Further analogies between black-hole physics and thermodynamics have been noted by the present author,^{3,4} by Carter,⁵ and by Bardeen, Carter, and Hawking.⁶ The formal black-hole analog of the ordinary second law of thermodynamics (OSL), "the entropy of a closed system never decreases," is Hawking's theorem,⁷ "the surface area of a black hole never decreases."

We have conjectured^{3,8,4} that this last analogy is more than formal, namely, that the area of a black hole expressed in suitable units may be regarded as the black-hole entropy—entropy in the sense of information about the black-hole interior inaccessible to exterior observers. Logically related is a second conjecture,^{3,8,4} the generalized second law of thermodynamics (GSL): *The sum of the black-hole entropy and the common (ordinary) entropy in the black-hole exterior never decreases.* Arguments and examples supporting these conjectures are set forth in detail in Ref. 4 (henceforth called paper I) together with reasons for taking the black-hole entropy as

$$S_{\text{bh}} = \left(\frac{1}{2} \ln 2\right) \hbar^{-1} \alpha. \quad (1)$$

Here α is the rationalized black-hole area (event-horizon surface area divided by 4π), and \hbar is Planck's constant. Units with $G=c$ =Boltzmann's

constant $k=1$ are used so that $\hbar=2.6 \times 10^{-66}$ cm².

The GSL is the only guise in which the second law has a chance of retaining both its validity *and* its usefulness for regions of the universe near black holes. The OSL is inapplicable just to the exterior of the holes since they are sinks of entropy. And if applied to both interiors and exterior, the OSL is not useful since there is no way for exterior observers to measure the interior ordinary entropy and so to apply the law. We say that the OSL is transcended.^{3,4} Not so the GSL, since black-hole entropy can be determined without delving into black-hole interiors. Thus by generalizing the concept "entropy" to include black-hole entropy, one can hope to make the second law valid *and* useful in the presence of black holes. One also hopes that the GSL will provide the sort of rapid and effortless solutions to certain problems that the OSL is noted for in more ordinary contexts. But first it is clearly imperative to establish the general validity of the GSL beyond reasonable doubt; some progress in this direction is reported here.

Considering first the case of a system entering a far larger and more massive generic stationary black hole, we show in several steps that the associated increase in black-hole entropy, ΔS_{bh} , is at least as large as the maximum entropy S_{max} that the system can carry down the hole, as required by the GSL. In Sec. II we find a general lower bound for ΔS_{bh} by treating the system as a small perturbation on the hole, while making allowance for the system's nonzero dimensions which, in the final analysis, are dictated by quantum effects. The result agrees with a less general one obtained in paper I. In Sec. III we obtain an upper bound on S_{max} for a macroscopic system by thermodynamic

and quantum considerations, and show that it is in general smaller than the corresponding ΔS_{bh} . In Sec. IV we compute S_{max} for a simple microscopic system by quantum and statistical arguments, and show that in the mean $\Delta S_{\text{bh}} \geq S_{\text{max}}$. Statistical violations of the GSL are possible, but they become highly improbable as the system grows large. In Sec. V we consider situations more complicated than those studied earlier, and we elucidate the physical factors relevant to the operation of the GSL by means of an example. We also consider the prospects for a general proof of the GSL. Finally, in Sec. VI we show with two examples how one can make predictions with the GSL.

II. INCREASE IN BLACK-HOLE ENTROPY

We require a lower bound on the ΔS_{bh} associated with the absorption of a generic system by a much larger and more massive stationary black hole. Before we begin, it is worthwhile pointing out the inadequacy of treating the system as a point particle in our context. If we did this, we could argue, with Christodoulou,¹ Israel,⁹ and Bardeen *et al.*,⁶ that the increase in black-hole area can be made arbitrarily small (Christodoulou's "reversible" process). The argument is simplest for a Schwarzschild hole: Imagine the system deposited at rest (in stationary coordinates) arbitrarily near the horizon. Its energy (as measured from infinity) is red-shifted away so that in absorbing it the hole experiences no change in mass or area. Indeed, Israel, and Bardeen *et al.*, have inferred from such arguments that ΔS_{bh} can be made negligible compared to the entropy of the system, so that the GSL cannot work. However, as already recognized by Christodoulou, a reversible process is an ideal whose perfect attainment is prevented by the atomicity of matter. This may be unimportant in other contexts, but when the subject is entropy, whose very meaning springs from the atomicity, it cannot be ignored. Specifically, if one allows for the nonzero dimensions of the system required by the quantum nature of matter, then the center of mass of the system cannot be deposited at rest nearer to the horizon than the typical "radius" of the system, and the increase in area will no longer vanish. We shall now obtain a very general lower bound for this increase and show that it suffices to make the GSL work. An alternative but less general discussion leading to the same results has appeared in paper I.

We first recall some facts about black holes.^{7,10} A generic black hole is a region of spacetime enclosed in an event horizon (one-way membrane), which is a null hypersurface generated by a congruence of null geodesics characterized locally by

a (real) convergence ρ and shear σ . The proper 2-area δA of an element of horizon changes according to

$$\frac{d\delta A}{dv} = -2\rho\delta A, \quad (2)$$

where v is the affine parameter of a typical local generator. In turn ρ satisfies

$$\frac{d\rho}{dv} = \rho^2 + |\sigma|^2 + 4\pi T_{\beta\gamma} l^\beta l^\gamma, \quad (3)$$

where $T_{\beta\gamma}$ is the stress-energy tensor of the matter at the horizon, and $l^\beta = dx^\beta/dv$ is the (null) tangent vector to the local generator (as well as the outgoing normal to the horizon). We assume the weak energy condition⁷: $T_{\beta\gamma} l^\beta l^\gamma \geq 0$.

If we calculate $d^2\delta A/dv^2$ from (2), eliminate first derivatives with (2) and (3), integrate (over area) for given v , and then over v from v to $v = \infty$, we get

$$\frac{dA}{dv} = 2 \int_v^\infty dv' \int_H (4\pi T_{\beta\gamma} l^\beta l^\gamma + |\sigma|^2 - \rho^2) \delta A(v'). \quad (4)$$

Here H is a spacelike section of the horizon for fixed v and $A(v) = \int_H \delta A(v)$ is the total area of this section; we have assumed that $dA/dv \rightarrow 0$ as $v \rightarrow \infty$. Thus far the formalism can describe both a system collapsing to form a black hole and one falling into a pre-existing hole. To make the problem tractable we specialize to the latter case and consider only a stationary black hole (possibly more general than Kerr) and a system which makes only a small perturbation on it, i.e., one of small mass and moderate density. We may then regard $T_{\beta\gamma}$ as a first-order quantity. Now both σ and ρ vanish for a perfectly stationary hole.^{7,10} Thus for our perturbed hole both σ and ρ will be first-order quantities, so that $|\sigma|^2 - \rho^2$ may be neglected in (4) to first order.

The system first makes contact with the horizon at some v , say $v=0$. We obtain a *lower bound* for the total increase in area by integrating (4) from $v=0$ to $v=\infty$ (Ref. 11):

$$\Delta\alpha \geq 2 \int_0^\infty dv \int_v^\infty dv' \int_H T_{\beta\gamma} l^\beta l^\gamma \delta A(v'). \quad (5)$$

Here $\alpha = A/4\pi$ is the rationalized area. Let us now reexpress (5) in terms of the rest mass and dimensions of the system.

We assume that the system is small in size compared to the hole. Then we may work in a locally inertial frame (not necessarily the rest frame) whose origin lies within the system as the latter crosses the horizon. Let the frame be equipped

with an orthonormal tetrad λ_a^ν such that $\lambda_a^\nu \lambda_b^\nu = \eta_{ab}$ (Minkowski metric).¹² The λ_0^ν , the time axis, is the 4-velocity of the frame; λ_i^ν for $i=1, 2, 3$ is the space triad. For a generator which intersects the system we may write (for the interval encompassed in the frame)

$$l^\nu = K(\lambda_0^\nu + n^i \lambda_i^\nu), \quad (6)$$

where K and n^i are four appropriate parameters. Since $l^\nu l_\nu = 0$, $n^i n^i = 1$. Now in our particular frame $l^0 = K$, $l^i = Kn^i$, and the Christoffel symbols vanish. Then the geodesic equation for a generator shows that K and n^i are constants along each generator. Due to the smallness of the system, they will also be constant from generator to generator inside the system to within fractional errors of the order of the size of the system divided by that of the hole.

We shall choose λ_0^ν to be orthogonal to one of the (constant v) H surfaces at an event within the system. To the same approximation used above, λ_0^ν will be orthogonal to the H surface at all points within the system (if H is not pathological) so that H will be a constant-time surface. Now on a given generator two H surfaces differing by dv are separated by the interval $dx^\nu = l^\nu dv$. By the constancy of K , the time difference between them, $dt = dx^0 = K dv$, is constant within the system. Hence, all relevant H surfaces are constant-time surfaces within the system (and λ_0^ν is orthogonal to all) in our approximation. Since l^ν is the null outgoing normal to the H 's, it is clear from (6) that the n^i are the components of the unit spacelike (outgoing) normal to these surfaces. Thus the normal distance dx between H surfaces differing by dv is also Kdv , and the surfaces are parallel within the system.

We now introduce $T_{ab} = T_{\beta\gamma} \lambda_a^\beta \lambda_b^\gamma$, the ordinary components of the stress-energy tensor in our frame. In terms of these

$$T_{\beta\gamma} l^\beta l^\gamma = K^2 (T_{00} + 2T_{0i} n^i + T_{ij} n^i n^j), \quad (7)$$

which we substitute in (5). Consider first $\int_H T_{0i} n^i \delta A$. Since $T_{ab} = 0$ outside the system, the integral is unchanged if extended to a closed constant-time 2-surface enclosing that part of the system *outside* the horizon, and coinciding with H inside the system. On H , n^i is the *inward* normal to this surface; we redefine n^i to mean this same normal for the rest of the surface, and δA to mean the corresponding area element. Then by Gauss's theorem the integral is just $-\int_V T_{0i,i} dV$, where V denotes the volume of the part of the system outside the horizon at the given time. Conservation of energy as expressed in our frame implies that $T_{0i,i} = \partial T_{00} / \partial t$. Thus

$$\int_H T_{0i} n^i \delta A = -\frac{\partial}{\partial t} \int_V T_{00} dV - \int_H T_{00} \delta A, \quad (8)$$

where we allow for the facts that V decreases in time, while $dt = dx$ for the interval between H surfaces.

Now consider $\int_H T_{ij} n^i n^j \delta A$. It is consistent with our earlier remarks to regard the n^i as constants everywhere inside the system at a given time. Using $T_{ij,j} = \partial T_{i0} / \partial t$ and repeating the above procedure we get

$$\int_H T_{ij} n^i n^j \delta A = -\frac{\partial}{\partial t} \int_V T_{i0} n^i dV - \int_H T_{i0} n^i \delta A. \quad (9)$$

But

$$\int_V T_{i0} n^i dV = \int_x^\infty dx' \int_\Sigma T_{i0} n^i dA, \quad (10)$$

where dA is the area element on a 2-surface $\Sigma(x')$ parallel and simultaneous to H , x' measures distance normal to such surfaces starting from $x'=0$ at the innermost point of the system, and the integral over x' goes from the H surface (at $x'=x$) to the outermost point of the system. As before, we transform the surface integral in (10) to one over the volume V' of the system lying outward of $\Sigma(x')$. Since $dx = dt$ we have

$$\int_V T_{i0} n^i dV = -\frac{\partial}{\partial t} \int_x^\infty dx' \int_{V'} T_{00} dV' - \int_V T_{00} dV. \quad (11)$$

Substituting (8) and (9) into the integral of (7) and eliminating all reference to T_{0i} by means of (8) and (11) we get

$$\int_H T_{\beta\gamma} l^\beta l^\gamma \delta A = K^2 \frac{\partial^2}{\partial t^2} \int_x^\infty dx' \int_{V'} T_{00} dV'. \quad (12)$$

To integrate over v as in (5) we replace Kdv by dt in accordance with our earlier remark, and the integrals become trivial. After an integration by parts we get

$$\Delta\alpha \geq 2 \int_V x T_{00} dV, \quad (13)$$

where the integral is evaluated at $v=0$, i.e., at the moment when the system first makes contact with the horizon; $x=0$ labels points of the system right at the horizon. In making use of an inertial frame in the calculations, we have assumed implicitly that the self-gravity of the system is negligible. Then T_{00} represents the energy density of the system, and according to (13) $\Delta\alpha$ is at least

twice the product of the system's relativistic energy \mathcal{E} and the normal distance \bar{x} from its center of mass (c.m.) to the horizon at $v=0$ ($t=0$), all quantities measured in the inertial frame.

The \mathcal{E} exceeds the system's rest mass μ by a Lorentz factor $\gamma \geq 1$ corresponding to the relative motion between the c.m. and the frame. We can make this motion have no component in the x direction (normal to the H surfaces) by applying to our frame an appropriate Lorentz boost in that direction. This does not affect the constant-time character of the surfaces, or the argument leading to (13). Because there is no Lorentz contraction of the system in the x direction in the new inertial frame, $\bar{x} = b$ where b is the normal distance from the c.m. to the horizon at the instant of contact as measured in the rest frame of the system. Thus we get our final result,

$$\Delta\alpha \geq 2\mu b. \quad (14)$$

Equation (14) agrees with the less general result (A15) obtained in paper I by a different approach. Here the result does not depend on specific properties of Kerr holes (applying also to holes surrounded by massive disks, etc.), on the shape of the system, or on its motion (it may be subject to forces). We have already shown in paper I that the lower bound $\Delta\alpha = 2\mu b$ is actually attainable.

Combining (1) and (14) we get

$$\Delta S \geq \mu b \hbar^{-1} \ln 2. \quad (15)$$

For a macroscopic system b , which is essentially an effective radius, must be enormous compared to the Compton lengths of the constituent particles, which themselves are large compared to \hbar/μ . Therefore $\Delta S_{\text{bh}} \gg 1$. Thus, while the minimum increase in black-hole area [Eq. (14)] can be quite minute (10^{-19} cm² for the typical automobile), the corresponding increase in black-hole entropy [Eq. (15)] is always large in natural ($k=1$) units ($\approx 10^{46}$ for the automobile). Failure to allow for this point on the part of Israel,⁹ and Bardeen *et al.*,⁶ vitiates their final conclusion. We now go on to compare ΔS_{bh} to the entropy carried by the system down the hole.

III. CHANGE IN GENERALIZED ENTROPY: THERMODYNAMIC ANALYSIS

We call the sum of S_{bh} and the common entropy in the black-hole exterior the generalized entropy S_g . For the process in question $\Delta S_g = \Delta S_{\text{bh}} - S$, where S is the common entropy carried by the system down the hole. The following demonstration that $\Delta S_g > 0$ applies to a closed macroscopic system describable by ordinary thermodynamics. We first

define a characteristic temperature $T_b = \hbar (b \ln 2)^{-1}$. If our system is odd-shaped, we take for b the smallest "radius" of the system. Since for a macroscopic system b is large compared to the Compton length of any of its constituent nuclei or elementary particles, $T_b/m \ll 1$, where m is the rest mass of any such particle. Hence for absolute temperatures $T \leq T_b$, the thermal motions within the system are nonrelativistic. (The case of massless particles is treated later.)

We may express the entropy of the system at absolute temperature T as

$$S = S_0 + \int_0^T C(T') T'^{-1} dT', \quad (16)$$

where S_0 is its ground-state entropy (at $T=0$), and $C(T) > 0$ is its heat capacity at constant volume and analogous parameters. Likewise, the rest-mass energy μ is just

$$\mu = E_0 + \int_0^T C(T') dT', \quad (17)$$

where E_0 is the ground-state energy (including rest masses).¹³ From Eqs. (15)–(17) we have

$$\Delta S_g \geq E_0/T_b - S_0 + \int_0^T C(T')(T_b^{-1} - T'^{-1}) dT'. \quad (18)$$

The last integral reaches its absolute minimum at $T = T_b$; thus for any T ,

$$\Delta S_g \geq E_0/T_b - S_0 + \int_0^{T_b} C(T)(T_b^{-1} - T^{-1}) dT. \quad (19)$$

This new integral involves only the range $0 \leq T \leq T_b$ for which the thermal motions are nonrelativistic.

We now show that S_0 is negligible compared to E_0/T_b . In general $S_0 = \ln g$, where g is the degree of degeneracy of the ground state; in practice one deals with (nuclear or particle) spin degeneracy. Suppose we have to do with N_n "nuclei" of spin $\frac{1}{2}\hbar$; then $g = N_n^2$ and $S_0 = N_n \ln 2$. If m_n is the mass of a nucleus, then $E_0 \geq N_n m_n$ since other particles, i.e., electrons, may be present.¹⁴ As mentioned earlier $m_n/T_b \gg 1$; thus $S_0 \ll E_0/T_b$. If we try to circumvent this condition by considering nuclei of large spin $s\hbar$ so that $S_0 = N_n \ln(2s+1)$, we are soon thwarted. For $s \gg 1$ the spin is quasiclassical. We may then apply the result of Møller¹⁵ that a spinning particle of mass m_n and spin $s\hbar$ has an effective radius no smaller than $s\hbar/m_n$. By implication $b \gg s\hbar/m_n$, $m_n/T_b \gg s$, and so $E_0/T_b \gg N_n s$. For large s , $s > \ln(2s+1)$ so that again $S_0 \ll E_0/T_b$. Thus we neglect S_0 in (19).

Now consider the integral in (19). For moderately high T , $C(T)$ varies slowly over each range

of T corresponding to a given phase,¹⁶ and is of the order of the number of effective degrees of freedom of the system at that T . This last number will never exceed the number of elementary particles in the system by more than a factor of order unity. For low T various degrees of freedom become ineffective or are "frozen out" (a strictly quantum effect), and $C(T)$ drops, ultimately vanishing as $T \rightarrow 0$ in harmony with the third law of thermodynamics.¹⁷ It is consistent with all known results to assume that $C(T)$ vanishes at least as rapidly as T . It is then possible to define an "envelope" for C :

$$\mathfrak{C}(T) = \begin{cases} C_m T/T_c, & \text{for } T < T_c \\ C_m, & \text{for } T \geq T_c. \end{cases} \quad (20)$$

Here C_m is the largest value of C in $0 \leq T \leq T_b$, and T_c is the largest parameter for which $\mathfrak{C}(T) \geq C(T)$ in that range. Depending on the value of T_b two basic cases are possible, one with $T_c = T_b$, the second with $0 < T_c < T_b$. Both are illustrated in Fig. 1. Now the integrand in (19) is negative; hence by replacing $C(T)$ by $\mathfrak{C}(T)$ we obtain a lower bound for it. Thus

$$\Delta S_g \geq E_0/T_b - C_m \left[\frac{1}{2} T_c/T_b + \ln(T_b/T_c) \right]. \quad (21)$$

First we consider the case with $T_c = T_b$. Let \bar{m} be the mean mass of an elementary particle defined by $E_0 = N\bar{m}$, where N is the total number of such particles. As before $\bar{m}/T_b \gg 1$ so that $E_0/T_b \gg N$. We remarked that C/N is bounded from above by a number of order unity; likewise for C_m/N . Thus we see from (21) that $\Delta S_g > 0$ since $E_0/T_b \gg \frac{1}{2} C_m$. The same argument will clearly go through if T_b/T_c exceeds 1 by a not very large factor (say up to 10; $\ln 10 \approx 2.3$ only). This case is relevant when the graph of $C(T)$ is concave downward and T_b lies below the leveling-off point T_0 .

Now consider the other cases under the label $T_c < T_b$. Just as above we see that $D \equiv \frac{1}{2} C_m T_c/T_b \ll E_0/T_b$, so we neglect D in (21). We now obtain an upper bound for T_b/T_c . Examples such as that in Fig. 1 show that T_c is either near the point T_0 from which $C(T)$ takes its last plunge to zero, or else is much larger. The final drop in $C(T)$ reflects the freezing out of the last group of degrees of freedom. It is well known that a degree of freedom is frozen out when T becomes comparable to or smaller than the energy spacing of its lower quantum states. Hence T_0 is no smaller than the energy spacing of the lower levels of the degrees of freedom in question. Let us estimate this spacing.

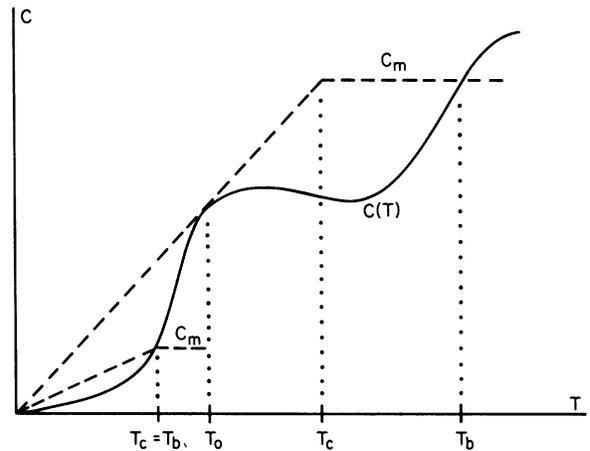


FIG. 1. Typical heat capacity $C(T)$ (solid line) and its "envelope" $\mathfrak{C}(T)$ (dashed lines) for two choices of T_b ; T_0 is the point from which $C(T)$ starts its final drop to zero.

We may regard a degree of freedom as some motion involving an effective mass m_e and confined to a region of dimensions a . Order-of-magnitude estimates via WKB or uncertainty-principle arguments indicate that the energy spacing of the lower levels is of order $\hbar^2(m_e a^2)^{-1}$. Moreover, comparison with precise results shows that this estimate is correct for such varied degrees of freedom as translational motion, rotation and vibration of molecules, electronic motion, etc. It is clear that for our system $a \leq b$ so we may take $T_0 \geq \hbar^2(m_e b^2)^{-1}$. Since $T_c \geq T_0$, $T_b/T_c \leq (\ln 2)^{-2} m_e/T_b$. As usual $m_e/T_b \gg 1$ so that $C_m \ln(T_b/T_c) \leq (C_m/N) N \ln(m_e/T_b)$. This last term is seen to be small compared to $E_0/T_b = N\bar{m}/T_b$ if, as in ordinary systems, the effective masses m_e of the last effective degrees of freedom do not exceed \bar{m} by more than a few orders of magnitude. Then according to (21) $\Delta S_g > 0$.

What if m_e/\bar{m} is sufficiently large to upset the last part in the previous argument? Consider a series of systems with fixed E_0 , \bar{m} , and b , but with successively larger m_e . It is clear that the number of degrees of freedom in question, which roughly equals $C(T_0)$, must scale like m_e^{-1} if m_e gets really large. Then Fig. 1 shows that T_c/T_0 will scale as m_e . Since $T_0 \propto m_e^{-1}$, T_c will be essentially unchanged as m_e is scaled up. It follows that increasing m_e cannot affect any term in (21) or our final conclusion that $\Delta S_g > 0$.

Finally, we consider a system with $E_0 = 0$; the prototype is black-body radiation for which $S_0 = 0$ as well. Equation (18) seems to say that for $0 < T \leq T_b$, $\Delta S_g < 0$ in violation of the GSL. However, thermodynamic methods are actually inapplicable

in this range due to the prevalence of quantum fluctuations.⁴ Specifically, fluctuations are negligible and ordinary thermodynamics applicable to a system only if $T \gg \hbar/\tau$, where τ is a characteristic time for change if the system is disturbed from equilibrium.¹⁷ For black-body radiation confined to dimensions b , $\tau \approx b$ and we need $T \gg T_b$ for thermodynamics to apply. If indeed $T \gg T_b$, ΔS_g is positive as is evident from (18) with $C(T) \propto T^3$. By contrast, if $T \lesssim T_b$, the entropy is ill defined¹⁷ and the GSL is made meaningless by the fluctuations. To conclude, we have shown that for a system describable by thermodynamics, $\Delta S_g > 0$ in agreement with the GSL.

IV. CHANGE IN GENERALIZED ENTROPY: STATISTICAL ANALYSIS

Thus far our arguments have not depended critically on the choice $\frac{1}{2} \ln 2$ for the numerical coefficient in formula (1) for S_{bh} . Any similar but not much smaller number would have served just as well. We now show that the above choice is unique in allowing "generalized reversible" processes for which $\Delta S_g = 0$ as well as those with $\Delta S_g > 0$. For a system describable by thermodynamics $\Delta S_g = 0$ is not a real possibility: We made it clear that the positive term in ΔS_g far outweighs the negative ones. Hence we must consider microscopic systems, i.e., systems composed of a small number of quanta. We regard such a system as a small wave packet. The results of Sec. II will still apply if we think of $T_{\beta\gamma}$ as the expectation value of the corresponding normal-ordered quantum field operator.¹⁸

The lower bound for the ΔS_{bh} associated with the absorption of the system by the hole is given by (15). It should be attainable—at least in the case of a Kerr hole—as shown in paper I. We first suppose that the packet contains N identical bosons of mass m , all in the same quantum state. For some packets $b = \hbar/m$ should be a real possibility (it would correspond to a sort of minimum-uncertainty packet); otherwise $b > \hbar/m$. If we neglect interactions, $\mu = Nm$ so that according to (15)

$$\Delta S_{bh} \geq N \ln 2. \quad (22)$$

Since m does not appear in (22), this relation should remain valid even for the case $m = 0$ to which our arguments are not strictly applicable. According to (22), ΔS_{bh} is not smaller than the equivalent of one "bit" per quantum. It was this feature which prompted us in paper I to choose the coefficient in (1) as we did.

Once the state of the quanta is agreed on, the packet is uniquely determined by N . We endow it

with entropy by regarding N as a random variable subject to some probability distribution $\{p_N\}$. Clearly $p_0 = 0$ because no quanta corresponds to no system. The statistical mean of any function $f(N)$ is just $\langle f \rangle = \sum_N p_N f(N)$; the entropy associated with the distribution is just (see paper I)

$$S = - \sum_N p_N \ln p_N. \quad (23)$$

We are especially interested in that normalized distribution which maximizes S for given $\langle N \rangle$. It is easily found by the method of Lagrange multipliers to be

$$p_N = (e^\beta - 1) e^{-\beta N}, \quad N \geq 1 \quad (24)$$

with $\beta = \ln \langle N \rangle - \ln \langle N - 1 \rangle$. From (23) it follows that

$$S_{\max} = \langle N \rangle \ln \langle N \rangle - \langle N - 1 \rangle \ln \langle N - 1 \rangle. \quad (25)$$

One easily finds now that the quantity $\langle N \rangle \ln 2 - S_{\max}$ is non-negative and vanishes only for $\langle N \rangle = 2$.

We conclude that $\langle \Delta S_g \rangle \equiv \langle \Delta S_{bh} \rangle - S \geq 0$ with the equality possible only if the lower bound in (15) is attained, if $b = \hbar/m$, and if the distribution (24) is realized with $\langle N \rangle = 2$. Note that any choice other than $\frac{1}{2} \ln 2$ in (1) would either result in $\langle \Delta S_g \rangle$ being negative sometimes, or would exclude the possibility $\langle \Delta S_g \rangle = 0$ altogether. Thus $\frac{1}{2} \ln 2$ is the "best" choice. The result $\langle \Delta S_g \rangle \geq 0$ means that the GSL is always satisfied in the mean (weak form of GSL). It need not always be satisfied in particular instances. For example, if the distribution (24) is realized with $\langle N \rangle = 2$ ($S_{\max} = 2 \ln 2$), the strong form of the GSL, $\Delta S_{bh} - S \geq 0$, can break down in an instance for which in fact $N = 1$, since $\Delta S_{bh} < 2 \ln 2$ is then possible. Thus the GSL is a statistical law susceptible to violations due to large fluctuations. In this it is no different from the OSL. In the macroscopic limit of large $\langle N \rangle$ the strong form of the GSL becomes a statistical certainty. This is seen as follows. For large $\langle N \rangle$, $S_{\max} \approx \ln \langle N \rangle$. Thus the probability p_v that $\Delta S_{bh} - S_{\max} < 0$ is no greater than the sum of the p_N from $N = 1$ to the next larger integer to $(\ln \langle N \rangle) / \ln 2$. One easily finds that $p_v \leq (\ln 2)^{-1} (\ln \langle N \rangle) / \langle N \rangle$. Thus a violation of the strong form of the GSL becomes improbable for large $\langle N \rangle$. This again has a familiar ring to it.

When the packet contains different types of quanta in different states, the lower bound (22) for ΔS_{bh} will not be attainable in general since $b > \hbar/m$ for the smallest m while μ contains all the various m . In order for the system to be specified by the various occupation numbers, the available one-quantum states must be agreed on and the information conveyed with the system. The simplest way to do this is to always have at least one quantum

occupying each available state, i.e., $p_0=0$ for each state. The entropy is maximized, first by taking the distributions for various states as statistically independent, and then by choosing (24) for each (boson) quantum state. The previous analysis can be applied and leads to the conclusion that $\langle \Delta S_g \rangle > 0$ and that ΔS_g is almost always positive for large occupation numbers. In this way we again demonstrate the validity of the GSL independently of Sec. III.

V. MORE GENERAL SITUATIONS

One may be interested in situations more general than those discussed above. For example, one might ask whether the GSL holds for the case of some tenuous matter accreting onto a stationary black hole from all directions. One might think that the law would hold if each small element of matter causes an increase in black-hole area at least equivalent to its own entropy. But this condition may be neither necessary nor meaningful. First, it might be impossible to identify the increase in area ascribable to a given element; however, this identification is possible in the case of steady-state accretion. Second, the fluid description of the matter assumed above may be inappropriate: In some regime the smallest representative elements may be larger than the hole. And third, entropy may be generated in the black-hole exterior as an organic part of the whole process; this extra entropy cannot be ignored. We illustrate some of these points with an example.³

Consider a Kerr hole of mass M immersed in a black-body radiation "bath" within a cavity of temperature T . We assume that the cavity does not rotate so that the radiation has zero angular momentum, and that the process of accretion has reached a steady state. Then [see paper I, Eqs. (8), (16), and (18)] to radiation of energy E (as measured from infinity) going down the hole there corresponds $\Delta S_{bh} = E/T_{bh}$, where T_{bh} is a certain characteristic temperature. For given M , $T_{bh} \leq \frac{1}{4} \hbar (M \ln 2)^{-1}$. Now far from the hole the entropy-to-energy ratio of the radiation will be the usual one, $\frac{4}{3} T^{-1}$.¹⁷ Since the accretion onto the hole takes place in steady state, it is plain that together with energy E , entropy $S = \frac{4}{3} ET^{-1}$ flows in. Thus for $T \geq \frac{4}{3} T_{bh}$, $\Delta S_{bh} - S \geq 0$ and the GSL is valid for each stage of the inflow.

One should not conclude that for $T < \frac{4}{3} T_{bh}$ the GSL is violated because in this regime a new feature must be taken into account. Indeed, for $T \leq \frac{4}{3} T_{bh}$, the typical wavelength in the radiation, \hbar/T , is already larger than the size of the hole $\leq 2M$. Thus the radiation no longer flows in as a fluid would, but rather most wavelengths in it must tunnel into

the hole individually. It is clear that the rate of tunnelling is sensitive to the wavelength: The shorter the wavelength, the higher the rate. In steady state, the spectrum of the radiation must harmonize with this rate, and so will no longer be of Planck type as for $T \gg T_{bh}$ (it would be of Planck type only if the rate were wavelength-independent). This means that the radiation cannot be in thermodynamic equilibrium with the cavity, for if it were the spectrum would be of Planck type. We may conclude that irreversible processes will go on in the cavity material (and perhaps in the radiation also), which will generate additional entropy. This extra entropy must be allowed for in applying the GSL which, in the spirit of the second law, need hold only for a closed system: black hole, radiation, and cavity. To prove that the extra entropy suffices to make the GSL work in the regime in question would be a difficult problem in nonequilibrium statistical mechanics which we shall not pursue. But it is clear that the various physical factors conspire in favor of the GSL.¹⁹

One is also interested in whether the GSL holds for the process of black-hole formation. Since one must here give up the assumption of a stationary hole, this case is certainly the most difficult to discuss. What sort of framework would be required to construct a proof of the GSL for this, or for that matter, for any situation more complicated than those treated in Secs. II–IV? We have repeatedly seen that quantum effects play a prominent role in the operation of the law: The GSL is at its roots a quantum law. We have also seen that the GSL is a statistical law. It seems, therefore, that such a proof could be obtained only in the context of quantum-statistical mechanics in curved spacetime. Since this discipline does not yet exist, we shall simply have to assume, on the strength of the arguments of paper I, and in the absence of evidence to the contrary, that the GSL is valid generally.

VI. PREDICTING WITH THE GSL

One may make predictions with the GSL; two simple examples follow. Consider first a cold nonrotating assembly of N nucleons. If its mass exceeds a critical mass of the order of $2M_\odot$, it will collapse spontaneously, presumably to form a Schwarzschild black hole. However, in principle smaller masses can also be induced to collapse. One way to accomplish this²⁰ would be to artificially compress the central portion of the system into a "seed" black hole which would then gobble up the rest of the nucleons. The energy expended in "seeding" can be made negligible. The initial common entropy of the system is at least as large as

the spin-degeneracy entropy of the spin- $\frac{1}{2}$ nuclei; thus $S \geq N \ln 2$ (see Sec. III). If the mass of the final Schwarzschild hole is M , then by (1)

$$S_{\text{bh}} = 2M^2 \hbar^{-1} \ln 2. \quad (26)$$

The GSL predicts that $S_{\text{bh}} - S \geq 0$, or $M^2 \geq \frac{1}{2} N \hbar$. Because of the binding energy of the initial system and the possible mass loss to gravitational waves, we must have $M < Nm$, where m is the nucleon rest mass. These relations imply that $N > \frac{1}{2} \hbar/m^2$ and $2M > \hbar/m$.

Thus according to the GSL, induced collapse to a black hole is possible only if no fewer than $\frac{1}{2} \hbar/m^2 \approx 10^{38}$ nucleons participate. This corresponds to $M > 10^{-19} M_{\odot}$ (10^{14} g) and a gravitational radius exceeding the Compton length of the nucleon. This is reasonable result, for if $2M < \hbar/M$, the "seed" hole would necessarily be smaller than the Compton length, and the nucleons would have great difficulty in falling in.²¹ Collapse could then take place only via quantum tunnelling, a very improbable process.²¹ If such collapse did take place nevertheless, it could be classed as a statistical violation of the GSL.²²

As a second example we consider a spherical thermal geon, that is, a spherical body of thermal radiation holding itself together by its own gravitational field.²³ The local temperature within it will not be strictly constant due to gravitational effects, but we shall replace it by an effective constant value T . The total radiation energy E is not a very precise concept when strong gravitational fields are present, but we shall use it nevertheless. Now the total entropy of the radiation should be approximately $S = \frac{4}{3} E T^{-1}$ (same as for

ordinary thermal radiation). If our geon collapses to a Schwarzschild hole, its mass M will be conserved. The S_{bh} will again be given by (26), and the GSL now predicts that

$$M^2 \geq \frac{2}{3} \hbar E (T \ln 2)^{-1} = 0.96 \hbar E T^{-1}. \quad (27)$$

The geon's effective radius R must clearly exceed the characteristic wavelength in the radiation $\approx \hbar/T$. Hence (27) is seen to be a necessary condition for M^2/R , the geon's gravitational binding energy, to exceed E . That is, (27) is a necessary condition for the geon to be able to collapse rather than explode.

We have seen how the GSL makes very reasonable (if not astounding) predictions for processes of black-hole formation. This success supports the view that it must be generally valid for such processes.

Note added in proof. The argument leading to Eq. (14) is perfectly valid if the system's intrinsic angular momentum \vec{s} is zero. If $\vec{s} \neq 0$ the position of the c.m. as computed in the inertial frame is at a distance $y = |\vec{v} \times \vec{s}| \mu^{-1}$ from that computed in the rest frame, where \vec{v} is the relative velocity between the frames [C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), p. 161]. However, for macroscopic systems $y \ll b$, so in this case at least the effect of \vec{s} is negligible. To see why $y \ll b$ consider first macroscopic rotation, which necessarily involves velocities much smaller than unity. Then $|\vec{s}| \ll \mu b$, so that $y \ll b$. If instead of macroscopic rotation we have all N spins aligned, $|\vec{s}| \approx \frac{1}{2} N \hbar$. But $b \gg \hbar/m \approx N \hbar/\mu$, so that again $y \ll b$.

*Work supported in part by National Science Foundation under Grant No. GP-32039.

¹D. Christodoulou, *Phys. Rev. Lett.* **25**, 1596 (1970); Ph.D. thesis, Princeton University, 1971 (unpublished).

²B. Carter, *Nat. Phys. Sci.* **238**, 71 (1972).

³J. D. Bekenstein, Ph.D. thesis, Princeton University, 1972 (unpublished).

⁴J. D. Bekenstein, *Phys. Rev. D* **7**, 2333 (1973).

⁵In *Black Holes*, edited by B. S. DeWitt and C. M. DeWitt (Gordon and Breach, New York, 1973).

⁶J. M. Bardeen, B. Carter, and S. W. Hawking, *Commun. Math. Phys.* **31**, 161 (1973).

⁷S. W. Hawking, *Phys. Rev. Lett.* **26**, 1344 (1971); *Commun. Math. Phys.* **25**, 152 (1972), and Ref. 5.

⁸J. D. Bekenstein, *Nuovo Cimento Lett.* **4**, 737 (1972).

⁹W. Israel, *Nuovo Cimento Lett.* **6**, 267 (1973).

¹⁰S. W. Hawking and J. B. Hartle, *Commun. Math. Phys.* **27**, 283 (1972).

¹¹We ignore the increase for $v < 0$ which need not vanish.

¹²Generic tetrad indices are denoted by a or b , space

tetrad indices by i or j , and time one by 0.

¹³Our implicit assumption of thermodynamic equilibrium is no restriction: For nonequilibrium S is smaller than at equilibrium. Thus if $\Delta S_g > 0$ for the equilibrium system, then automatically $\Delta S_g > 0$ for the nonequilibrium one.

¹⁴Ordinary binding energies are negligible; the possible exception is gravitational binding energy which we already assumed to be negligible.

¹⁵C. Møller, *Commun. Dublin Inst. Adv. Stud. A* **5** (1949).

¹⁶The specific heat at constant pressure is singular at phase transitions, but that at constant volume is not.

¹⁷L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley, Reading, Mass., 1969), 2nd edition, pp. 66, 344, 165.

¹⁸By calculating in an inertial frame we hope to bypass the thorny problem of quantization in curved spacetime.

¹⁹Another example involving an apparent violation of the second law was discussed in Ref. 8. The resolution given there was rightly criticized by L. Basano and

A. Morro, *Nuovo Cimento Lett.* **6**, 193 (1973), and by Israel (Ref. 9). However, we have independently given in paper I a different resolution to which their objections are inapplicable.

²⁰Ya. B. Zel'dovich and I. D. Novikov, *Stars and Relativity* (Univ. of Chicago Press, Chicago, 1971), p. 295.

²¹The existence of this "Compton barrier" has been demonstrated in detail by J. A. Wheeler, in *Cortona Symposium on Weak Interactions*, edited by L. Radicati

(Accademia Nazionale dei Lincei, Rome, 1971).

²²The existence of a nucleon number below which, classically, induced collapse will not occur has been inferred in a different way by B. K. Harrison, K. S. Thorne, M. Wakano, and J. A. Wheeler, *Gravitation Theory and Gravitational Collapse* (Univ. of Chicago Press, Chicago, 1965), p. 82.

²³J. A. Wheeler and E. A. Power, *Rev. Mod. Phys.* **29**, 480 (1957).