PHYSICAL REVIEW D VOLUME 7, NUMBER 8 15 APRIL 1973

Black Holes and Entropy*

Jacob D. Bekenstein†
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08540
and Center for Relativity Theory, The University of Texas at Austin, Austin, Texas 78712‡
(Received 2 November 1972)

There are a number of similarities between black-hole physics and thermodynamics. Most striking is the similarity in the behaviors of black-hole area and of entropy: Both quantities tend to increase irreversibly. In this paper we make this similarity the basis of a thermodynamic approach to black-hole physics. After a brief review of the elements of the theory of information, we discuss black-hole physics from the point of view of information theory. We show that it is natural to introduce the concept of black-hole entropy as the measure of information about a black-hole interior which is inaccessible to an exterior observer. Considerations of simplicity and consistency, and dimensional arguments indicate that the black-hole entropy is equal to the ratio of the black-hole area to the square of the Planck length times a dimensionless constant of order unity. A different approach making use of the specific properties of Kerr black holes and of concepts from information theory leads to the same conclusion, and suggests a definite value for the constant. The physical content of the concept of black-hole entropy derives from the following generalized version of the second law: When common entropy goes down a black hole, the common entropy in the black-hole exterior plus the black-hole entropy never decreases. The validity of this version of the second law is supported by an argument from information theory as well as by several examples.

I. INTRODUCTION

A black hole[1] exhibits a remarkable tendency to increase its horizon surface area when undergoing any transformation. This was first noticed by Floyd and Penrose[2] in an example of the extraction of energy from a Kerr black hole by means of what has come to be known as a Penrose process.[3] They suggested that an increase in area might be a general feature of black-hole transformations. Independently, Christodoulou[4,5] had shown that no process whose ultimate outcome is the capture of a particle by a Kerr black hole can result in the decrease of a certain quantity which he named the irreducible mass of the black hole, M_{ir}. In fact, most processes result in an increase in M_{ir} with the exception of a very special class of limiting processes, called reversible processes, which leave M_{ir} unchanged. It turns out that M_{ir} is proportional to the square root of the black hole's area[5,6] [see (1)]. Thus Christodoulou's result implies that the area increases in most processes, and thus it supports the conjecture of Floyd and Penrose. Christodoulou's conclusion is also valid for charged Kerr black holes.[4,6]

By an approach radically different from Christodoulou's, Hawking[7] has given a general proof that the black-hole surface area cannot decrease in *any* process. For a system of several black holes Hawking's theorem implies that the area of each individual black hole cannot decrease, and moreover that when two black holes merge, the area of the resulting black hole (provided, of course, that one forms) cannot be smaller than the sum of initial areas.

It is clear that changes of a black hole generally take place in the direction of increasing area. This is reminiscent of the second law of thermodynamics which states that changes of a closed thermodynamic system take place in the direction of increasing entropy. The above comparison suggests that it might be useful to consider black-hole physics from a thermodynamic viewpoint: We already have the concept of energy in black-hole physics, and the above observation suggests that something like entropy may also play a role in it. Thus, one can hope to develop a thermodynamics for black holes – at least a rudimentary one. In this paper we show that it is possible to give a precise definition of black-hole entropy. Based on it we construct some elements of a thermodynamics for black holes.

There are some precedents to our considerations. The idea of making use of thermodynamic methods in black-hole physics appears to have been first considered by Greif.[8] He examined the possibility of defining the entropy of a black hole, but lacking many of the recent results in black-hole physics, he did not make a concrete proposal. More recently, Carter[9] has rederived the result of Christodoulou[4,5] that the irreducible mass of a Kerr black hole is unchanged in a reversible transformation by applying to the black hole the criterion for a thermodynamically reversible transformation of a rigidly rotating star.[10] Carter's example shows the possibilities inherent in the use of thermodynamic arguments in black-hole physics.

In this paper we attempt a unification of black-hole physics with thermodynamics. In Sec. II we point out a number of analogies between black-hole physics and thermodynamics, all of which bring out the parallelism between black-hole area and entropy. In Sec. III, after a short review of elements of the theory of information, we discuss some features of black-hole physics from the point of view of information theory. We take the area of a black hole as a measure of its entropy—entropy in the sense of inaccessibility of information about its internal configuration. We go further in Sec. IV and propose a specific expression for black-hole entropy in terms of black-hole area. Earlier[11,12] we had proposed this same expression on different grounds; here we find the value of a previously unknown constant by means of an argument based on information theory. In Sec. V we propose a generalization of the second law of thermodynamics applicable to black-hole physics: When some common entropy goes down a black hole, *the black-hole entropy plus the common entropy in the black-hole exterior never decreases*.[11,12]

In Secs. VI and VII we construct several examples which provide support for the generalized second law. In addition, we analyze in Sec. VII a thought experiment proposed by Geroch[13] in which, with the help of a black hole, heat is apparently converted entirely into work in violation of the second law. We show that, in fact, due to fundamental physical limitations the conversion efficiency is somewhat smaller than unity. Moreover, the efficiency is no greater than the maximum efficiency allowed by thermodynamics for the heat engine which is equivalent to the Geroch process, so that this process cannot be regarded as violating the second law.

II. ANALOGIES BETWEEN BLACK-HOLE PHYSICS AND THERMODYNAMICS

We have already mentioned the resemblance between the tendency of black-hole area to increase, and the tendency of entropy to increase. Changes of a black hole or of a system of black holes select a preferred direction in time: that in which the black-hole area increases. Likewise, changes of a closed thermodynamic system select a preferred direction in time: that in which the entropy increases. This parallelism between black-hole area and entropy goes even deeper.

Black-hole area turns out to be as intimately related to the degradation of energy as is entropy. In thermodynamics the statement "the entropy has increased" implies that a certain quantity of energy has been degraded, i.e., that it can no longer be transformed into work. Now, as Christodoulou has emphasized,[4,5] the irreducible mass M_{ir} of a Kerr black hole, which is related to the surface area A of the black hole by[14]

$$M_{ir} = (A/16\pi)^{1/2}\,, \tag{1}$$

represents energy which cannot be extracted by means of Penrose processes.[3] In this sense it is inert energy which cannot be transformed into work. Thus, an increase in A, and hence in M_{ir}, corresponds to the degradation (in the thermodynamic sense) of some of the energy of the black hole.

The irreducible mass of a Schwarzschild black hole is just equal to its total mass. Thus, no energy can be extracted from such a black hole by means of Penrose processes. However, the merger of two Schwarzschild black holes can yield energy in the form of gravitational waves.[7] The only restriction on the process is that the total black-hole area must not decrease as a result of the merger.[7] However, the sum of the irreducible masses of individual black holes may (in fact, does) decrease. We see that for a system of several black holes the degraded energy E_d is more appropriately given by

$$E_d = (\textstyle\sum A/16\pi)^{1/2} = (\sum M_{ir}{}^2)^{1/2} \tag{2}$$

than by $\sum M_{ir}$. According to this formula the degraded energy of a system of black holes is smaller than the sum of degraded energies of the black holes considered separately. Thus by combining Schwarzschild black holes which are already "dead," one can still obtain energy.[7] Analogously, by allowing two thermodynamic systems which are separately in equilibrium to interact, one can obtain work, whereas each system by itself could have done no work. From the above observations the parallelism between black-hole area and entropy is again evident.

We shall now construct the black-hole analog of the thermodynamic expression

$$dE = TdS - PdV\ . \tag{3}$$

For convenience we shall from now on write all our equations in terms of the "rationalized area" of a black hole α defined by

$$\alpha = A/4\pi\ . \tag{4}$$

Consider a Kerr black hole of mass M, charge Q, and angular momentum $\vec{L}$. (3-vectors here refer to components with respect to the Euclidean frame at infinity.) Its rationalized area is given by[5,7]

$$\alpha = r_+^2 + a^2$$

$$= 2Mr_+ - Q^2, \tag{5}$$

where

$$\vec{a} = \vec{L}/M, \tag{6}$$

$$r_\pm = M \pm (M^2 - Q^2 - a^2)^{1/2}. \tag{7}$$

Differentiating (5) and solving for dM we obtain

$$dM = \Theta d\alpha + \vec{\Omega} \cdot d\vec{L} + \Phi dQ, \tag{8}$$

where

$$\Theta \equiv \tfrac{1}{4}(r_+ - r_-)/\alpha, \tag{9a}$$

$$\vec{\Omega} \equiv \vec{a}/\alpha, \tag{9b}$$

$$\Phi \equiv Qr_+/\alpha. \tag{9c}$$

In (8) we have the black-hole analog of the thermodynamic expression (3): The terms $\vec{\Omega}\cdot d\vec{L}$ and ΦdQ clearly represent the work done *on* the black hole by an external agent who increases the black hole's angular momentum and charge by $d\vec{L}$ and dQ, respectively. Thus $\vec{\Omega}\cdot d\vec{L} + \Phi dQ$ is the analog of $-PdV$, the work done *on* a thermodynamic system. Comparing our expression for work with the expressions for work done on rotating[15] and charged[16] bodies, we see that $\vec{\Omega}$ and Φ play the roles of rotational angular frequency and electric potential of the black hole, respectively.[5,9] The α in (8) resembles the entropy S in (3) as we have noted before: For any change of the black hole $d\alpha \geq 0$,[5-7] while for any change of a closed thermodynamic system $dS \geq 0$. Moreover, it is clear from (7) and (9a) that Θ, the black-hole analog of temperature T, is non-negative just as T is. From the above observations the formal correspondence between (3) and (8) is evident.

All the analogies we have mentioned are suggestive of a connection between thermodynamics and black-hole physics in general, and between entropy and black-hole area in particular. But so far the analogies have been of a purely formal nature, primarily because entropy and area have different dimensions. We shall remedy this deficiency in Sec. IV by constructing out of black-hole area an expression for black-hole entropy with the correct dimensions. Preparatory to this we shall now look at black-hole physics from the point of view of the theory of information.

III. INFORMATION AND BLACK-HOLE ENTROPY

The connection between entropy and information is well known.[17,18] The entropy of a system measures one's uncertainty or lack of information about the actual internal configuration of the system. Suppose that all that is known about the internal configuration of a system is that it may be found in any of a number of states with probability p_n for the nth state. Then the entropy associated with the system is given by Shannon's formula[17,18]

$$S = -\sum_n p_n \ln p_n \ . \tag{10}$$

This formula is uniquely determined by a few very general requirements which are imposed in order that S have the properties expected of a measure of uncertainty.[17]

It should be noticed that the above entropy is dimensionless. This simply means that we choose to measure temperature in units of energy. Boltzmann's constant is then dimensionless.

Whenever new information about the system becomes available, it may be regarded as imposing some constraints on the probabilities p_n. For example, the information may be that several of the p_n are, in fact, zero. Such constraints on the p_n always result in a decrease in the entropy function.[18] This property is formalized by the relation[17,18]

$$\Delta I = -\Delta S, \tag{11}$$

where ΔI is the new information which corresponds to a decrease ΔS in one's uncertainty about the internal state of the system. Equation (11) is the basis for Brillouin's identification of information with negative entropy.[18] All the above comments apply to such divers systems as a quantity of gas in a box or a telegram. A familiar example of the relation between a gain of information and a decrease in entropy is the following. Some ideal gas in a container is compressed isothermally. It is well known that its entropy decreases. On the other hand, one's information about the internal configuration of the gas increases: After the compression the molecules of the gas are more localized, so that their positions are known with more accuracy than before the compression.

The second law of thermodynamics is easily understood in the context of information theory. The entropy of a thermodynamic system which is not in equilibrium increases because information about the internal configuration of the system is being lost during its evolution as a result of the washing out of the effects of the initial conditions. It is possible for an exterior agent to cause a decrease in the entropy of a system by first acquiring information about the internal configuration of the system. The classic example of this is that of Maxwell's demon.[18] But information is never free. In acquiring information ΔI about the system, the agent inevitably causes an increase in the entropy of the rest of the universe which exceeds ΔI.[18] Thus, even though the entropy of the system decreases in accordance with (11), the over-all entropy of the universe increases in the process.

The conventional unit of information is the "bit" which may be defined as the information available when the answer to a yes-or-no question is precisely known (zero entropy). According to the scheme (11) a bit is also numerically equal to the maximum entropy that can be associated with a yes-or-no question, i.e., the entropy when no information whatsoever is available about the answer. One easily finds that the entropy function (10) is maximized when $p_{\text{yes}} = p_{\text{no}} = \frac{1}{2}$. Thus, in our units, one bit is equal to ln2 of information.

Let us now return to our original subject, black holes. In the context of information a black hole is very much like a thermodynamic system. The entropy of a thermodynamic system in equilibrium measures the uncertainty as to which of all its internal configurations compatible with its macroscopic thermodynamic parameters (temperature, pressure, etc.) is actually realized. Now, just as a thermodynamic system in equilibrium can be completely described macroscopically by a few thermodynamic parameters, so a bare black hole in equilibrium (Kerr black hole) can be completely described (insofar as an exterior observer is concerned) by just three parameters: mass, charge, and angular momentum.[1] Black holes in equilibrium having the same set of three parameters may still have different "internal configurations." For example, one black hole may have been formed by the collapse of a normal star, a second by the collapse of a neutron star, a third by the collapse of a geon. These various alternatives may be regarded as different possible internal configurations of one and the same black hole characterized by their (common) mass, charge, and angular momentum. It is then natural to introduce the concept of black-hole entropy as the measure of the *inaccessibility* of information (to an exterior observer) as to which particular internal configuration of the black hole is actually realized in a given case.

At the outset it should be clear that the black-hole entropy we are speaking of is *not* the thermal entropy inside the black hole. In fact, our black-hole entropy refers to the equivalence class of all black holes which have the same mass, charge, and angular momentum, not to one particular black hole. What are we to take as a measure of this black-hole entropy? The discussion of Sec. II predisposes us to single out black-hole area. But to be more general we shall only assume that the entropy of a black hole, S_{bh}, is some *monotonically increasing* function of its rationalized area:

$$S_{\text{bh}} = f(\alpha). \tag{12}$$

Although our motivating discussion for the introduction of the concept of the black-hole entropy made use of the specific properties of stationary black holes, we shall take (12) to be valid for any black hole, including a dynamically evolving one, since the surface area is well defined for any black hole. This choice is supported by the following observations.

As mentioned earlier, the entropy of an evolving thermodynamic system increases due to the gradual loss of information which is a consequence of the washing out of the effects of the initial conditions. Now, as a black hole approaches equilibrium, the effects of the initial conditions are also washed out (the black hole loses its hair)[1]; only mass, charge, and angular momentum are left as determinants of the black hole at late times. We would thus expect that the loss of information about initial peculiarities of the hole will be reflected in a gradual increase in S_{bh}. And indeed Eq. (12) predicts just this; by Hawking's theorem S_{bh} increases monotonically as the black hole evolves. This agreement is evidence in favor of the choice (12).

We mentioned earlier that the possibility of causing a decrease in the entropy of a thermodynamic system goes hand in hand with the possibility of obtaining information about its internal configuration. By contrast, an exterior agent cannot acquire any information about the interior configuration of a black hole. The one-way membrane nature of the event horizon prevents him from doing so.[1] Therefore, we do not expect an exterior agent to be able to cause a decrease in the black hole's entropy. Equation (12) is in agreement with this expectation; by Hawking's theorem S_{bh} never decreases. Here we have a new piece of evidence in favor of the choice (12).

One possible choice for f in (12), $f(\alpha) \propto \alpha^{1/2}$, is untenable on various grounds. Consider two black holes which start off very distant from each other. Since they interact weakly we can take the total black-hole entropy to be the sum of the S_{bh} of each black hole. The black holes now fall together, merge, and form a black hole which settles down to equilibrium. In the process no information about the black-hole interior can become available; on the contrary, much information is lost as the final black hole "loses its hair." Thus, we expect the final black-hole entropy to exceed the initial one. By our assumption that $f(\alpha) \propto \alpha^{1/2}$, this implies that the irreducible mass [see (1)] of the final black hole exceeds the sum of irreducible masses of the initial black holes. Now suppose that all three black holes are Schwarzschild ($M = M_{ir}$). We are then confronted with the prediction that the final black-hole mass exceeds the initial one. But this is nonsense since the total black-hole mass can only decrease due to gravitational radiation losses. We thus see that the choice $f(\alpha) \propto \alpha^{1/2}$ is untenable.

The next simplest choice for f is

$$f(\alpha) = \gamma\alpha\,, \tag{13}$$

where γ is a constant. Repetition of the above argument for this new f leads to the conclusion that the final black-hole area must exceed the total initial black-hole area. But we know this to be true from Hawking's theorem.[7] Thus the choice (13) leads to no contradiction. Therefore, we shall adopt (13) for the moment; later on we shall exhibit some more positive evidence in its favor.

Comparison of (12) and (13) shows that γ must have the units of $(\text{length})^{-2}$. But there is no constant with such units in classical general relativity. If in desperation we appeal to quantum physics we find only one truly universal constant with the correct units[14]: $\hbar^{-1}$, that is, the reciprocal of the Planck length squared. (Compton wavelengths are not universal, but peculiar to this or that particle; they clearly have no bearing on the problem.) We are thus compelled to write (12) as

$$S_{bh} = \eta\hbar^{-1}\alpha\,, \tag{14}$$

where η is a dimensionless constant which we may expect to be of order unity. This expression was first proposed by us earlier[11,12] from a different point of view.

We need not be alarmed at the appearance of $\hbar$ in the expression for black-hole entropy. It is well known[15] that $\hbar$ also appears in the formulas for entropy of many thermodynamic systems that are conventionally regarded as classical, for example, the Boltzmann ideal gas. This is a reflection of the fact that entropy is, in a sense, a count of states of the system, and the underlying states of any system are always quantum in nature. It is thus not totally unexpected that $\hbar$ appears in (14). These observations also suggest that it would be somewhat pretentious to attempt to calculate the precise value of the constant $\eta\hbar^{-1}$ without a full understanding of the quantum reality which underlies a "classical" black hole. Since there is no hope at present of obtaining such an understanding, we bypass the issue, and in the next section we use a semiclassical argument to arrive at a value for $\eta\hbar^{-1}$ which should be quite close to the correct one.

IV. EXPRESSION FOR BLACK-HOLE ENTROPY

In our attempt to obtain a value for $\eta\hbar^{-1}$ we shall employ the following argument. We imagine that a particle goes down a Kerr black hole. As it disappears some information is lost with it. According to the discussion of Sec. III we expect the black-hole entropy, as the measure of inaccessible information, to reflect the loss of the information associated with the particle by increasing by an appropriate amount. How much information is lost together with the particle? The amount clearly depends on how much is known about the internal state of the particle, on the precise way in which the particle falls in, etc. But we can be sure that the absolute minimum of information lost is that contained in the answer to the question "does the particle exist or not?" To start with, the answer is known to be yes. But after the particle falls in, one has no information whatever about the answer. This is because from the point of view of this paper, one knows nothing about the physical conditions inside the black hole, and thus one cannot assess the likelihood of the particle continuing to exist or being destroyed. One must, therefore, admit to the loss of one bit of information (see Sec. III) at the very least.

Our plan, therefore, is to compute the minimum possible increase in the black hole's area which results from the disappearance of a particle down the black hole, then to compute the corresponding minimum possible increase of black-hole entropy by means of our original formula (12), and finally to identify this increase in entropy with the loss of one bit of information in accordance with the scheme (11). If our procedure is reasonable we should then recover the functional form of f given by (13), together with a definite value for γ.

There are many ways in which a particle can go down a black hole, all leading to varying increases in black-hole area. We are interested in that method for inserting the particle which results in the smallest increase. This method has already been discussed by Christodoulou[4-6] in connection with his introduction of the concept of irreducible mass. The essence of Christodoulou's method is that if a freely falling point particle is captured by a Kerr black hole from a turning point in its orbit, then the irreducible mass and, consequently, the area of the hole are left unchanged. For reasons that will become clear presently we wish to allow the particle to have a nonzero radius. As shown in Appendix A, Christodoulou's method can be generalized easily so as to allow for this, as well as for the possibility that the particle is brought to the horizon by some method other than by free fall. We find in Appendix A that the increase in area for the generalized Christodoulou process is no longer precisely zero. But interestingly enough, the minimum increase in rationalized area, $(\Delta\alpha)_{min}$, turns out to be independent of the parameters of the black hole. For a spherical particle of rest mass μ, and proper radius b,

$$(\Delta\alpha)_{min}=2\mu b\,. \tag{15}$$

For a point particle $(\Delta\alpha)_{min}=0$; this is Christodoulou's result.

Expression (15) gives the minimum possible increase in black-hole area that results if a given particle is added to a Kerr black hole. We can try to make $(\Delta\alpha)_{min}$ smaller by making b smaller. However, we must remember that b can be no smaller than the particle's Compton wavelength $\hbar\mu^{-1}$, or than its gravitational radius 2μ, whichever is the larger. The Compton wavelength is the larger for $\mu\leq 2^{-1/2}\hbar^{1/2}$, and the gravitational radius is the larger for $\mu>2^{-1/2}\hbar^{1/2}$ $(2^{-1/2}\hbar^{1/2}\approx 10^{-5}\,g)$. Thus, if $\mu\leq 2^{-1/2}\hbar^{1/2}$, then $2\mu b$ can be as small as $2\mu\hbar\mu^{-1}=2\hbar$. But if $\mu>2^{-1/2}\hbar^{1/2}$, then $2\mu b$ can be no smaller than $4\mu^2>2\hbar$. We conclude that quantum effects set a lower bound of $2\hbar$ on the increase of the rationalized area of a Kerr black hole when it captures a particle. Moreover, this limit can be reached only for a particle whose dimension is given by its Compton wavelength. Of course, only such an "elementary particle" can be regarded as having no internal structure. Therefore, the loss of information associated with the loss of such a particle should be minimum. And indeed we find that the increase in black-hole entropy is smallest for just such a particle. This supports our view that $2\hbar$ is the increase in rationalized area associated with the loss of one bit of information.

Following our program we shall equate the minimum increase in black-hole entropy, $(\Delta S_{bh})_{min}=2\hbar\, df/d\alpha$, with ln2, the entropy increase associated with the loss of one bit of information. Integration of the resulting equation gives $f(\alpha)=(\frac{1}{2}\ln 2)\hbar^{-1}\alpha$. Thus, we have arrived again at (13) by an alternate route, and have obtained the value of γ into the bargain. We now have

$$S_{bh}=(\tfrac{1}{2}\ln 2)\hbar^{-1}\alpha\,, \tag{16}$$

which is of the same form as (14). Our argument has determined the dependence of S_{bh} on α in a straightforward manner. However, our value $\eta=\frac{1}{2}\ln 2$ might presumably be challenged on the grounds that it follows from the assumption that the smallest possible radius of a particle is precisely equal to its Compton wavelength whereas the actual radius is not so sharply defined. Nevertheless, it should be clear that if η is not exactly $\frac{1}{2}\ln 2$, then it must be very close to this, probably within a factor of two. This slight uncertainty in the value of η is the price we pay for not giving our problem a full quantum treatment. However, in what follows we shall suppose that $\eta=\frac{1}{2}\ln 2$. Examples to be given later will show that this value leads to no contradictions.

How is the entropy of a system of several black holes defined? It is natural to define it as the sum of individual black-hole entropies. Then Hawking's theorem tells us that the total black-hole entropy of the system cannot decrease. But this is just what we would expect since the information lost down the black holes is unrecoverable. This observation lends support to our choice.

In conventional units (16) takes the form

$$S_{bh}=(\tfrac{1}{2}\ln 2/4\pi)kc^3\hbar^{-1}G^{-1}A$$

$$=(1.46\times 10^{48}\text{ erg }{}^\circ\text{K}^{-1}\text{ cm}^{-2})A\,, \tag{17}$$

where k is Boltzmann's constant. We see that the entropy of a black hole is enormous. For example, a black hole of one solar mass would have $S_{bh}\approx 10^{60}$ erg $^\circ$K^{-1}. By comparison the entropy of the sun is $S\approx 10^{42}$ erg $^\circ$K^{-1}; those of a white dwarf or a neutron star of one solar mass even smaller. The large numerical value of black-hole entropy serves to dramatize the highly irreversible character of the process of black-hole formation. We may define a characteristic temperature for a Kerr black hole by the relation $T_{bh}^{-1}=(\partial S_{bh}/\partial M)_{L,Q}$ which is the analog of the thermodynamic relation $T^{-1}=(\partial S/\partial E)_V$. By using (8) and (16) we find

$$T_{bh}=2\hbar(\ln 2)^{-1}\Theta$$

$$=(0.165\,{}^\circ\text{K cm})\,(r_+-r_-)\,(r_+^2+a^2)^{-1}\,, \tag{18}$$

where $r_\pm$ and a are to be given in centimeters. We introduce this T_{bh} in anticipation of our discussion of an example in Sec. VII. But we emphasize that one should not regard T_{bh} as *the* temperature of the black hole; such an identification can easily lead to all sorts of paradoxes, and is thus not useful.

V. THE GENERALIZED SECOND LAW

Suppose that a body containing some common entropy goes down a black hole. The entropy of the visible universe decreases in the process. It would seem that the second law of thermodynamics is transcended here in the sense that an exterior observer can never verify by direct measurement that the total entropy of the whole universe does not decrease in the process.[19] However, we know that the black-hole area "compensates" for the disappearance of the body by increasing irreversibly. It is thus natural to conjecture that the second law is not really transcended provided that it is expressed in a generalized form: *The common entropy in the black-hole exterior plus the black-hole entropy never decreases*. This statement means that we must regard black-hole entropy as a genuine contribution to the entropy content of the universe.

Support for the above version of the second law comes from the following argument. Suppose that a body carrying entropy S goes down a black hole (which may have existed previously or may be formed by the collapse of the body). The S is the uncertainty in one's knowledge of the internal configuration of the body. So long as the body was still outside the black hole, one had the option of removing this uncertainty by carrying out measurements and obtaining information up to the amount S. But once the body has fallen in, this option is lost; the information about the internal configuration of the body becomes truly inaccessible. We thus expect the black-hole entropy, as the measure of inaccessible information, to increase by an amount S. Actually, the increase in S_{bh} may be even larger because any information that was available about the body to start with will also be lost down the black hole. Therefore, if we denote by ΔS_c the change in common entropy in the black-hole exterior ($\Delta S_c \equiv -S$), then we expect that

$$\Delta S_{bh} + \Delta S_c = \Delta(S_{bh} + S_c) > 0 . \tag{19}$$

This is just the generalized second law which we proposed above: The generalized entropy $S_{bh} + S_c$ never decreases. Examples supporting this law will be given in Sec. VI-VII.

This is a good place to mention the question of fluctuations. We know that the common entropy of a closed thermodynamic system can decrease spontaneously as a result of statistical fluctuations, i.e., the second law, being a statistical law, is meaningful only if statistical fluctuations are small. Is black-hole entropy also subject to decreases of a statistical nature? Not classically – Hawking's theorem guarantees that. Quantum mechanically there are two ways by which the black-hole entropy can undergo statistical decreases. One of them depends on the quantum fluctuations of the metric of the black hole which one has reasons to expect.[1] Such fluctuations would be reflected in small random fluctuations in the area, and thus in the entropy of the black hole, and some of these fluctuations would be expected to be decreases in entropy. However, even if one regards a black hole as a purely classical object, it is still possible for its area and entropy to undergo small decreases when the black hole absorbs a single quantum under certain conditions.[20] However, the probability of such an event occurring in any given trial is very small. Therefore, the decrease in area and entropy is of a statistical nature, and is quite analogous to the decrease in entropy of a thermodynamic system due to statistical fluctuations. This discussion serves us warning that the law (19) is expected to hold only insofar as statistical fluctuations are negligible.

We noticed earlier (Sec. IV) the very large magnitude of black-hole entropy. In fact, one can say that the black-hole state is the maximum entropy state of a given amount of matter. The point is that in the gravitational collapse of a body into a black hole, the loss of information down the black hole is the maximum allowed by the laws of physics. Thus if the body collapses to form a Kerr black hole, all information about it is lost with the exception of mass, charge, and angular momentum.[1] These quantities are given in terms of Gaussian integrals,[1] and so information about them cannot be lost. But all other information about the body is eventually lost. Therefore, the resulting black hole must correspond to the maximum (generalized) entropy which can be associated with the given body.

VI. EXAMPLES OF THE GENERALIZED SECOND LAW AT WORK

In the examples which follow we endeavor to subject the generalized second law to the most stringent test possible in each case by maximizing the entropy going down the black hole with a given body while minimizing the associated increase in black-hole entropy.

A. Harmonic Oscillator

As a first example we take an harmonic oscillator composed of two particles of rest mass $\frac{1}{2}m$ each connected by a nearly massless spring of spring constant K. We imagine the oscillator to be enclosed in a spherical box and to be maintained at temperature T. We assume for simplicity that conditions are such that the oscillator is nonrelativistic ($T \ll m$). Let ω be the vibrational frequency of the oscillator. Then the (normalized) probability that the oscillator is in its nth quantum state is given by the canonical distribution

$$p_n = (1 - e^{-x})^{-1} e^{-nx} \ , \quad x = \hbar\omega/T \ . \tag{20}$$

The entropy of the oscillator as computed from (10) is

$$S = x(e^x - 1)^{-1} - \ln(1 - e^{-x}) \ , \tag{21}$$

and the mean vibrational energy,

$$\langle E \rangle \equiv \sum p_n (n + \tfrac{1}{2})\hbar\omega ,$$

is

$$\langle E \rangle = [(e^x - 1)^{-1} + \tfrac{1}{2}] \hbar\omega \ . \tag{22}$$

We remark that the thermal distribution (20) maximizes the entropy of the oscillator for given $\langle E \rangle$, and is thus ideally suited to our plan for subjecting the generalized second law to the most stringent test possible.

Suppose that the box goes down a Kerr black hole. The corresponding increase in black-hole entropy cannot be smaller than the lowest limit derived by the method of Appendix A. From (15) and (16) we have $\Delta S_{bh} \geq \mu b \hbar^{-1} \ln 2$, where b is the outer radius of the box and μ is its total rest mass. Clearly b must be at least as large as half of the mean value $\langle y \rangle$ of the separation of the two masses y. And $\langle y \rangle$ in turn must clearly be larger than Δy, the root mean square of the thermal oscillation of y $[(\Delta y)^2 \equiv \langle\langle (y - \langle y \rangle)\rangle\rangle]$, so that y will always be positive. Now according to the (quantum) virial theorem $\frac{1}{2}\langle E \rangle$ is equal to the mean potential energy of the oscillator $\frac{1}{2}K(\Delta y)^2$. Since the reduced mass of the oscillator is $\frac{1}{4}m$, we have $K = \frac{1}{4}m\omega^2$. We thus find from all the above that

$$b > \langle E \rangle^{1/2} m^{-1/2} \omega^{-1} \ .$$

Remembering that $\mu > m + \langle E \rangle$ (because the box itself must have some mass) we obtain

$$\Delta S_{bh} > \langle E \rangle^{1/2} m^{-1/2} (\hbar\omega)^{-1} (m + \langle E \rangle) \ln 2 \ . \tag{23}$$

We assume that the entropy given by (21) is the only contribution to the entropy in the box. This amounts to neglecting the contribution of the black body radiation in the box, etc., a sensible procedure if T is not very high. Then $\Delta S_c = -S$ and we have

$$\Delta(S_{bh} + S_c) > \xi^{-1/2}(1 + \xi)[\tfrac{1}{2} + (e^x - 1)^{-1}] \ln 2 - x(e^x - 1)^{-1} + \ln(1 - e^{-x}) \ , \tag{24}$$

where we have introduced the notation $\xi \equiv m\langle E \rangle^{-1}$ and used Eqs. (21) and (22) for S and $\langle E \rangle$. We now show that $\Delta(S_{bh} + S_c) > 0$ as required by the generalized second law. The expression in (24) regarded as a function of x for given ξ has a single minimum at

$$x = x_m \equiv \xi^{-1/2}(1 + \xi) \ln 2$$

which has the value

$$\tfrac{1}{2}x_m + \ln[1 - \exp(-x_m)] \ .$$

Our assumption that the oscillator is nonrelativistic means that $\xi \gg 1$, and hence that $x_m \gg 2\ln 2$. Under these conditions the minimum is positive (in fact, it is positive for $\xi \gtrsim 1$). It follows immediately that $\Delta(S_{bh} + S_c)$ is positive for all x and all ξ which are compatible with the requirement of a nonrelativistic oscillator. The generalized second law is obeyed over the entire regime for which our treatment is valid.

B. Beam of Light

As a second example we consider a beam of light which is aimed at a Kerr black hole. This example is particularly interesting because it shall bring us face to face with the issue of fluctuations as a limitation on the applicability of the second law. We shall restrict our attention only to those cases for which geometrical optics is a valid approximation. We shall thus represent the path of the beam by a null geodesic in the Kerr background.

We shall take it that the beam is thermalized at a certain temperature T. This implies that its entropy is a maximum for given energy. The entropy is easily calculated; in fact, the entropy and energy for each mode in the beam are given by the same expressions (21) and (22) which apply to a harmonic oscillator, except that one must omit the zero-point energy term $\frac{1}{2}\hbar\omega$. The total entropy S and mean energy $\langle E \rangle$ of the beam are obtained by integrating these expressions weighed by the conventional density of states

$$\rho = 2\omega^2 (2\pi)^{-3} V d\Omega \tag{25}$$

over all ω. In (25) V is the volume of the beam and $d\Omega$ is the solid angle it subtends. Integrating by parts the expression for S, one easily obtains the relation

$$S = \frac{4}{3}\langle E\rangle T^{-1}, \tag{26}$$

which, not surprisingly, is identical to that for radiation inside a black-body cavity of temperature T.[15] [In Ref. 12 (26) was given with an incorrect numerical factor.]

As the beam nears the black hole, it is deflected by the gravitational field. Insofar as its effects on electromagnetic radiation are concerned, a stationary gravitational field can be mocked up by an appropriate nonabsorbing refractive medium in flat spacetime.[21] But the propagation of a beam of light through such a medium is a reversible process.[22] We infer from this that the entropy of the beam will remain unchanged as the beam nears the black hole. Thus the entropy change of the visible universe when the beam goes down the hole is just

$$\Delta S_c = -\frac{4}{3}\langle E\rangle\ T^{-1}. \tag{27}$$

What is the increase in black-hole entropy associated with the process? From (8) we see that the increase in α is minimized when the angular momentum that the hole gains from the beam is maximized for given $\langle E\rangle$. Now, the gain in angular momentum is limited because the beam will not be captured if it carries too much angular momentum. In Appendix B we take this into account in calculating (in the geometrical optics limit) the minimum possible increase in α for given $\langle E\rangle$ of the beam. We find that $\Delta\alpha \geq \beta M\langle E\rangle$ where β ranges from 8 for the case of a Schwarzschild hole to $4(1-\sqrt{3}/2)$ for the case of an extreme Kerr hole, this last value being the smallest possible β. From (16) it follows that

$$\Delta S_{\rm bh} \geq (\tfrac{1}{2}\beta \ln 2) M \hbar^{-1}\langle E\rangle. \tag{28}$$

Our assumption (Appendix B) that geometrical optics is always applicable means that the bulk of wavelengths in the beam are much shorter than the characteristic dimension of the hole $\approx M$. Thus, if ω_c is some characteristic frequency in the beam, then we require that $\omega_c \gg M^{-1}$. From the form of the Planck spectrum (22) we see that $\hbar\omega_c \approx T$; therefore (27) tells us that

$$|\Delta S_c| \ll \frac{4}{3} M\hbar^{-1}\langle E\rangle\ . \tag{29}$$

Comparison of (28) and (29) shows that a violation of the generalized second law (19) cannot arise in the regime under consideration.

In the above discussion the condition that geometrical optics be applicable prevented us from taking T to be arbitrarily small. As a result it turned out to be impossible for $|\Delta S_c|$ to exceed $\Delta S_{\rm bh}$, and so a violation of the second law was ruled out. But there is a way to circumvent the restriction on T. One simply selects the temperature T (arbitrarily) to be as small as one pleases, and arranges for all frequencies $\omega<\omega_c'$ to be filtered out of the beam. Here $\omega_c' \gg M^{-1}$ is a definite frequency unrelated to T. It should be clear that geometrical optics will be a valid approximation for this case also, so that we may take over the result (28). But the result (27) must be modified since we are here dealing with a truncated frequency spectrum. We are mostly interested in the regime $T\ll\hbar\,\omega_c'$. Then for all frequencies in the beam $x=\hbar\omega/T\gg 1$. It follows from (21) and (22) that for each mode the entropy to energy ratio is T^{-1} ($S\approx xe^{-x}$, $\langle E\rangle \approx \hbar\omega e^{-x}$). Therefore instead of (27) we have

$$\Delta S_c = -\langle E\rangle\ T^{-1}. \tag{30}$$

It now appears that if

$$T< T_c \equiv \hbar(\tfrac{1}{2}\beta M \ln 2)^{-1},$$

then $\Delta S_{\rm bh}+\Delta S_c$ will be negative in contradiction with the generalized second law.

The resolution of the above paradox is that in the regime $T< T_c$ statistical fluctuations are already dominant so that our entire picture of the process is invalid. To verify the importance of fluctuations we calculate the mean number of quanta N in the beam by integrating the mean number of quanta per mode, $(e^x-1)^{-1}$, weighed by the density of states (25) over all $\omega>\omega_c'$. For $T=T_c$ we get (recall that $\hbar\omega_c'/T_c \gg 1$ by our assumptions)

$$N\approx \frac{V}{M^3}\,\frac{d\Omega}{4\pi}\ \delta^{-3}(\delta M\,\omega_c')^2 \exp(-\delta M\,\omega_c')\ , \tag{31}$$

where $\delta\equiv\frac{1}{2}\beta\ln 2$ ($0.2\leq\delta\leq 2.8$). It is clear that for any beam aimed at the black hole $d\Omega/4\pi\ll 1$. Recalling that $M\omega_c'\gg 1$ by assumption, we see from (31) that each quantum occupies a mean volume much larger than M^3. But the cross section of the beam must be smaller than $\sim M^2$ if the beam is to go down the black hole. Thus the mean separation between quanta is much larger than M, the characteristic dimension of the black hole. In case $T< T_c$ the above effect is even more accentuated.

We conclude that in the regime $T\lesssim T_c$ for which the second law (19) appears to break down, our description of the process as a continuous beam going down the black hole is invalidated by the large fluctuations in the concentration of energy in the beam (or equivalently, the large fluctuations in the energy of each section of the beam). In this regime $\langle E\rangle$ is no longer a good measure of the actual energy. It appears, therefore, that statistical fluctuations are responsible for the breakdown of the second law in the context in which we have applied it here. But we can demonstrate that the law has not lost all its meaning by adopting a point of view more suitable to the circumstances at hand than the one used above.

We take the point of view that quanta are going down the black hole one at a time, rather than in a continuous stream. Thus we must check the validity of the generalized second law for the infall of each quantum. The analysis of Appendix B still leads to formula (28) for the increase in black-hole entropy except that $\langle E\rangle$ is replaced by $\hbar\omega$, the energy of the quantum. To compute the common entropy going down the black hole we reason as follows. From our point of view a quantum of definite frequency is going down the black hole. Thus we are no longer dealing with the probability distribution (20); instead we shall ascribe probability $\frac{1}{2}$ to each of the two possible polarizations of the quantum. Then according to (10) the entropy associated with the quantum is ln2. Therefore,

$$\Delta S_{\mathrm{bh}}+\Delta S_c \geq (\tfrac{1}{2}\beta \ln 2)M\omega - \ln 2. \tag{32}$$

Since $\frac{1}{2}\beta \geq 0.268$, and since we are assuming that $M\omega \gg 1$, we see that $\Delta S_{\mathrm{bh}}+\Delta S_c$ is in fact positive: The generalized second law is upheld for the infall of each quantum.

VII. A PERPETUAL MOTION MACHINE USING A BLACK HOLE?

Geroch[13] has described a procedure using a black hole which appears to violate the second law of thermodynamics by converting heat into work with unit efficiency. He envisages a box filled with black-body radiation which is slowly lowered by means of a string from far away down to the horizon of a black hole, at which point its energy as measured from infinity vanishes. Therefore, if the box's rest mass is μ, then the agent lowering the string obtains work equal to μ out of the process. The box is then allowed to emit into the black-hole radiation of (proper) energy $\Delta\mu$. Finally, the agent retrieves the box; since its rest mass is now $\mu-\Delta\mu$, he must do work $\mu-\Delta\mu$ to accomplish this. Therefore, in the whole process the agent obtains net work $\Delta\mu$ at the expense of heat $\Delta\mu$– conversion with unit efficiency. We shall now show that, in fact, due to fundamental physical limitations, the efficiency of the Geroch process is slightly smaller than unity, so that no violation of the second law is entailed here.

The box under consideration must have a nonzero radius (see below). Because of this its energy as measured from infinity is never quite zero when it is as close to the horizon as it can possibly be. We shall assume that the box is in the shape of a sphere of radius b. Then according to the analysis of Appendix C the minimum value of the energy is

$$E=2\mu b\Theta , \tag{33}$$

where Θ is defined by (9a). It follows that in lowering the box from infinity to the horizon, the agent obtains only work $\mu(1-2b\Theta)$ rather than μ. After the box has radiated into the black hole, its rest mass becomes $\mu-\Delta\mu$ and according to (33) its energy at the horizon is just $2(\mu-\Delta\mu)b\Theta$. Thus the agent must do work $(\mu-\Delta\mu)(1-2b\Theta)$ to retrieve the box to infinity where its energy is $\mu-\Delta\mu$. Therefore, in the over-all process the agent obtains net work $\Delta\mu(1-2b\Theta)$ in exchange for the expenditure of heat $\Delta\mu$. The efficiency of conversion is

$$\epsilon = 1-2b\Theta , \tag{34}$$

which is smaller than unity. In practical situations $b \ll r_+$ so that $b\Theta \ll 1$ and the efficiency can be quite near to unity. But the departure of ϵ from unity, albeit small, serves to resolve the problem raised by Geroch's example: There is no violation of the Kelvin statement of the second law.[23]

We must now explain why b cannot be arbitrarily small. Physically the reason is that the box must be large enough for the wavelengths characteristic of radiation of some temperature T to fit into it. More formally we can argue as follows. The frequency of the photon ground state associated with the box, ω_0, cannot exceed that frequency ω_p at which the Planck photon-number spectrum

$$\propto \omega^2[\exp(\hbar\omega/T)-1]^{-1}$$

peaks. Otherwise the frequencies of all photon states would lie in the exponential tail of the spectrum, the occupation number of each state would be small, and the resulting large fluctuations would make the concept of temperature meaningless. We have the conventional relation $\omega_0 b' = \pi$, where b' is the interior radius of the box ($b'<b$), and we easily find that $\hbar\omega_p < 2T$. Therefore, $\omega_0 < \omega_p$ implies that $b > \pi\hbar/2T$. It is thus clear that there is a lower limit for b.

We may write the efficiency (34) in a more transparent form by recalling that $\Theta = \frac{1}{2}T_{\mathrm{bh}} \ln 2/\hbar$ [see (18)], where T_{bh} is the characteristic temperature associated with the black hole. Since $b > \pi\hbar/2T$ we find that

$$\epsilon < 1 - T_{bh}/T \ . \tag{35}$$

We now recall that the efficiency of a heat engine operating between two reservoirs, one at temperature T and the second at temperature $T_{bh} < T$, is restricted by $\epsilon \leq 1 - T_{bh}/T$. We thus see that the Geroch process is no more efficient than its "equivalent reversible heat engine." This observation makes it evident again that Geroch's process is not in violation of the second law of thermodynamics. Finally, we wish to remark that since our primary formula (33) is valid only when the box is small compared to the black hole ($b\Theta \ll 1$), we can vouch for the validity of (35) only when $T_{bh} \ll T$. However, due to the smallness of T_{bh} this condition will be satisfied in all cases of practical interest.

We now verify that the Geroch process is in accord with the generalized second law (19). We mentioned earlier that the agent obtains work $\Delta\mu(1-2b\Theta)$ for a decrease $\Delta\mu$ in the rest mass of the box. This means that the black hole's mass must increase by $2\Delta\mu b\Theta$ in the complete process. According to (8) and (16) the corresponding increase in black hole entropy is $\Delta S_{bh} = \Delta\mu b \hbar^{-1} \ln 2$ (angular momentum is not added to the hole; see Appendix C). But since $b > \pi\hbar/2T$ we have that

$$\Delta S_{bh} > (\tfrac{1}{2}\pi \ln 2)\Delta\mu/T \ . \tag{36}$$

On the other hand, the decrease in entropy of the box is clearly $\Delta\mu/T$ (heat/temperature). Thus

$$\Delta S_c = -\Delta\mu/T \ . \tag{37}$$

From (36) and (37) it follows that $\Delta(S_{bh} + S_c) > 0$ as required by the generalized second law.

ACKNOWLEDGMENTS

The author is grateful to Professor J. A. Wheeler for his interest in this work, his encouragement and his penetrating criticism. He also thanks Professor Karel Kuchař and Professor Brandon Carter, and Dr. Bahram Mashhoon for valuable suggestions and comments.

APPENDIX A

Here we shall calculate the minimum possible increase in black-hole area which must result when a spherical particle of rest mass μ and proper radius b is captured by a Kerr black hole. We are interested in the increase in area ascribable to the particle itself, as contrasted with any increase incidental to the process of bringing the particle to black-hole horizon. For example, there is some *circumstantial* evidence for believing that when the particle is lowered into the black hole by a string, there occurs an increase in black-hole area even as the particle is being lowered.[11] Furthermore, the area will experience an additional increase due to the gravitational waves radiated into the black hole by the string as it relaxes when the particle is dropped.[11] Similarly, if the particle falls freely to the horizon it emits gravitational waves into the hole even before it falls in; the amount of radiation may even be significant.[24] This radiation will also result in an increase in area. Here we shall ignore all these incidental effects and concentrate on the increase in area caused by the particle all by itself.

We assume that the particle is neutral so that it follows a geodesic of the Kerr geometry when falling freely. We shall employ Boyer-Lindquist coordinates for the charged Kerr metric[25]

$$ds^2 = g_{tt}\,dt^2 + 2g_{t\phi}dt\,d\phi + g_{\phi\phi}d\phi^2 + g_{rr}\,dr^2 + g_{\theta\theta}d\theta^2. \tag{A1}$$

For later reference we give g_{rr}:

$$g_{rr} = (r^2 + a^2 \cos^2\theta)\Delta^{-1} \ , \tag{A2}$$

where

$$\Delta \equiv r^2 - 2Mr + a^2 + Q^2 \ . \tag{A3}$$

The event horizon is located at $r = r_+$ where $r_\pm$ are defined by (7). We have

$$\Delta = (r - r_-)(r - r_+) \ . \tag{A4}$$

First integrals for geodesic motion in the Kerr background have been given by Carter.[25] Christodoulou[5] uses the first integral

$$E^2[r^4 + a^2(r^2 + 2Mr - Q^2)] - 2E(2Mr - Q^2)ap_\phi - (r^2 - 2Mr + Q^2)p_\phi{}^2 - (\mu^2 r^2 + q)\Delta = (p_r\Delta)^2 \tag{A5}$$

as a starting point of his analysis. In (A5) $E = -p_t$ is the conserved energy, p_ϕ is the conserved component of angular momentum in the direction of the axis of symmetry, q is Carter's fourth constant of the motion,[25] μ is the rest mass of the particle, and p_r is its covariant radial momentum.

Following Christodoulou we solve (A5) for E:

$$E = Bap_\phi + \{[B^2a^2 + A^{-1}(r^2 - 2Mr + Q^2)]p_\phi^2 + A^{-1}[(\mu^2r^2 + q)\Delta + (p_r\Delta)^2]\}^{1/2}, \tag{A6}$$

where

$$A \equiv r^4 + a^2(r^2 + 2Mr - Q^2), \tag{A7}$$

$$B \equiv (2Mr - Q^2)A^{-1}. \tag{A8}$$

At the event horizon $\Delta = 0$ [see (A4)] so that there

$$A = A_+ = (r_+^2 + a^2)^2, \quad B = B_+ = (r_+^2 + a^2)^{-1}. \tag{A9}$$

Furthermore, at the horizon $Ba = \Omega$ [see (9b)], and the coefficients of p_ϕ^2 and $\mu^2r^2 + q$ in (A6) vanish. However,

$$p_r\Delta = (r^2 + a^2\cos^2\theta)p^r$$

does not vanish at the horizon in general. If the particle's orbit intersects the horizon, then we have from (A6) that

$$E = \Omega p_\phi + A_+^{-1/2}|p_r\Delta|_+.$$

As a result of the capture, the black hole's mass increases by E and its component of angular momentum in the direction of the symmetry axis increases by p_ϕ. Therefore, according to (8) the black hole's rationalized area will increase by $\Theta^{-1}A_+^{-1/2}|p_r\Delta|_+$. As pointed out by Christodoulou this increase vanishes *only if* the particle is captured from a turning point in its orbit in which case $|p_r\Delta|_+ = 0$. In this case we have

$$E = \Omega p_\phi. \tag{A10}$$

The above analysis shows that it is possible for a black hole to capture a *point* particle without increasing its area. How is this conclusion changed if the particle has a nonzero proper radius b? First we note that regardless of the manner in which the particle arrives at the horizon (being lowered by a string, splitting off from a second particle which then escapes, etc.), it must clearly acquire its parameters E, p_ϕ, and q while every part of it is still outside the horizon, i.e., while it is not yet part of the black hole. Moreover, as the particle is captured, it must already be detached from whatever system brought it to the horizon, so that it may be regarded as falling freely. Therefore, Eq. (A6) should always describe the motion of the particle's center of mass at the moment of capture.

It should be clear that to generalize Christodoulou's result to the present case one should evaluate (A6) not at $r = r_+$, but at $r = r_+ + \delta$, where δ is determined by

$$\int_{r_+}^{r_+ + \delta} (g_{rr})^{1/2}\,dr = b$$

($r = r_+ + \delta$ is a point a proper distance b outside the horizon). Using (A2) we find

$$b = 2\delta^{1/2}(r_+^2 + a^2\cos^2\theta)^{1/2}(r_+ - r_-)^{-1/2}. \tag{A11}$$

To obtain this we have assumed that $r_+ - r_- \gg \delta$ (black hole not nearly extreme). Expanding the argument of the square root in (A6) in powers of δ, replacing δ by its value given by (A11), and keeping only terms to O(b) we get

$$E = \Omega p_\phi + [(r_+^2 - a^2)(r_+^2 + a^2)^{-1}p_\phi^2 + \mu^2r_+^2 + q]^{1/2} \times \tfrac{1}{2}b(r_+ - r_-)(r_+^2 + a^2)^{-1}(r_+^2 + a^2\cos^2\theta)^{-1/2} \tag{A12}$$

Here we have already assumed that the particle reaches a turning point as it is captured since we know that this minimizes the increase in black-hole area. Equation (A12) is the generalization to $O(b)$ of the Christodoulou condition (A10).

What is q in (A12)? We can obtain a lower bound for it as follows. From the requirement that the θ momentum p_θ be real it follows that [25]

$$q \geq \cos^2\theta[a^2(\mu^2 - E^2) + p_\phi^2/\sin^2\theta]; \tag{A13}$$

the equality holds when $p_\theta = 0$ at the point in question. If we replace E in (A13) by Ωp_ϕ [see (A12)] we obtain

$$q \geq \cos^2\theta[a^2\mu^2 + p_\phi^2(1/\sin^2\theta - a^2\Omega^2)],$$

which is correct to zeroth order in b. We know that $1/\sin^2\theta \geq 1$; it is easily shown that $a^2\Omega^2 \leq \frac{1}{4}$ for a charged Kerr black hole. Therefore $q \geq a^2\mu^2\cos^2\theta$. Substituting this into (A12) we find

$$E \geq \Omega p_\phi + \tfrac{1}{2}\mu b(r_+ - r_-)(r_+^2 + a^2)^{-1} \tag{A14}$$

which is correct to $O(b)$. By retracing our steps we see that the equality sign in (A14) corresponds to the case $p_\phi = p_\theta = p^r = 0$ at the point of capture. The increase in black-hole area, computed by means of (8), (9a), and (A14), is

$$\Delta\alpha \geq 2\mu b\ . \tag{A15}$$

This gives the fundamental lower bound on the increase in black-hole area. We note that it is independent of M, Q, and L.

APPENDIX B

Here we shall calculate the minimum possible increase in black-hole area which must result when a light beam of energy $E > 0$ coming from infinity is captured by a Kerr black hole. If the black hole is nonrotating the increase is simply obtained by setting $dM = E$ in (8):

$$\Delta\alpha = 8ME \quad \text{for} \quad a = 0. \tag{B1}$$

If the black hole is rotating, $\Delta\alpha$ can be minimized by maximizing the angular momentum p_ϕ which is brought in by the beam [see (8)]. To accomplish this we consider the effective potential V for the motion of a massless particle in a Kerr background.

This V is just the value of E given by (A6) regarded as a function of r for $\mu = 0$ and $p_r = 0$ (E equals V at a turning point). This potential starts off at a value Ωp_ϕ at $r = r_+$ (see Appendix A), increases with r, reaches a maximum, and then falls off to zero as $r\to\infty$. For the beam to be captured by the hole it is necessary that p_ϕ be small enough for the peak of the potential barrier to be smaller than E of the beam. The optimum case we seek corresponds to the peak being just equal to E so that p_ϕ has its largest possible value.

It is clear that we must take q in (A6) as small as possible in order to have the lowest possible potential peak for given p_ϕ. Let us first take $q < 0$. Then according to Carter[25] there are solutions to the geodesic equation only if $|p_\phi| < aE$. From (8) it follows that

$$\Delta\alpha = \Theta^{-1}(E - \Omega p_\phi)$$

$$> \Theta^{-1}E(1 - \Omega a).$$

But since $\Omega a \leq \frac{1}{2}$ and $\Theta^{-1} \geq 8M$ it follows that

$$\Delta\alpha > 4ME \quad \text{for} \quad q < 0. \tag{B2}$$

Next we take $q = 0$. Two cases are possible[25]: Either $|p_\phi| < aE$ as above so that (B2) is again applicable, or else the orbit is purely equatorial. In the second case one may calculate the peak of the barrier numerically and then find the optimum increase in α. It turns out that $(\Delta\alpha)_{\min}$ decreases monotonically with a for fixed M. The limit of $(\Delta\alpha)_{\min}$ as $a \to M$ may be computed analytically because in this limit one can find the height of the potential analytically with sufficient accuracy. One finds

$$(\Delta\alpha)_{\min} \to 4\,(1 - \tfrac{1}{2}\sqrt{3}\,)ME \quad \text{as} \quad a \to M. \tag{B3}$$

It is clear that for $q > 0$ the potential peak will be higher and the increase in area will be larger than the one given by (B3). Thus we find that the minimum increase in area results when the beam is captured by an extreme Kerr black hole from a purely equatorial orbit.

APPENDIX C

Here we compute the value of the energy (as measured from infinity) of a particle of rest mass μ and proper radius b which is hanging from a string just outside the horizon of a Kerr black hole. It is clear that the particle will not be moving in the r or θ directions; hence $p^r = p_\theta = 0$ for it. We cannot claim that the particle does not move in the ϕ direction. In fact, since it will be within the ergosphere in general, it cannot avoid moving in the ϕ direction.[1] Our intuitive notion that the particle is "not moving" must be applied only in a locally nonrotating (Bardeen) frame.[26] The particle is at rest in such a frame if for it

$$\frac{d\phi}{dt} = -\frac{g_{t\phi}}{g_{\phi\phi}}.$$

It follows that

$$p_\phi = \mu\left(g_{t\phi}\frac{dt}{d\tau} + g_{\phi\phi}\frac{d\phi}{d\tau}\right) = 0.$$

If the particle were to be dropped, it would clearly keep its energy E and it would still have $p_\phi = p^r = p_\theta = 0$, at least momentarily. We may thus compute E for the particle hanging in the string at a proper distance b from the horizon by setting $p_\phi = 0$ in (A12). For q we take the value given by (A13) with the equality sign ($p_\theta = 0$) $p_\phi = 0$, and $E = 0$ [since for $p_\phi = 0$, E is of $O(b)$]. Thus

$$q = \mu^2 a^2 \cos^2\theta$$

and

$$E = \tfrac{1}{2}\,\mu\, b(r_+ - r_-)\,(r_+{}^2 + a^2)^{-1}$$

$$= 2\mu b\Theta\ . \tag{C1}$$

*Based in part on the author's Ph.D. thesis, Princeton University, 1972; work supported in part by the National Science Foundation Grants No. GP 30799X to Princeton University and No. GP 32039 to the University of Texas at Austin.

†National Science Foundation Predoctoral Fellow when this work was initiated.

‡Present address.

[1]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[2]R. Penrose and R. M. Floyd, Nature 229, 177 (1971).

[3]R. Penrose, Riv. Nuovo Cimento 1, 252 (1969).

[4]D. Christodoulou, Phys. Rev. Letters 25, 1596 (1970).

[5]D.Christodoulou, Ph.D. thesis, Princeton University, 1971 (unpublished).

[6]D. Christodoulou and R. Ruffini, Phys. Rev. D 4, 3552 (1971).

[7]S. W. Hawking, Phys. Rev. Letters 26, 1344 (1971); contribution to *Black Holes*, edited by B. DeWitt and C. DeWitt (Gordon and Breach, New York, 1973).

[8]J. M. Greif, Junior thesis, Princeton University, 1969 (unpublished).

[9]B. Carter, Nature 238, 71 (1972).

[10]Ya. B. Zel'dovich and I. D. Novikov, *Stars and Relativity* (University of Chicago Press, Chicago, 1971), p. 268.

[11]J. D. Bekenstein, Ph.D. thesis, Princeton University, 1972 (unpublished).

[12]J. D. Bekenstein, Lett. Nuovo Cimento 4, 737 (1972).

[13]R. Geroch, Colloquium at Princeton University, December 1971.

[14]We use units with $G = c = 1$ unless otherwise specified.

[15]L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley, Reading, Mass., 1969).

[16]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Addison-Wesley, Reading, Mass., 1960), p. 5.

[17]The mathematical definition of information was first given by C. E. Shannon; the relevant papers are reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communications* (University of Illinois Press, Urbana, 1949). An elementary introduction to information theory is given by M. Tribus and E. C. McIrvine, Sci. Amer. 225 (No. 3), 179 (1971). The derivation of statistical mechanics from information theory was first carried out by E. T. Jaynes, Phys. Rev. 106, 620 (1957); 108, 171 (1957).

[18]The relation between information theory and thermodynamics is discussed in detail by L. Brillouin, *Science and Information Theory* (Academic, New York, 1956), especially Chaps. 1, 9, 12–14.

[19]This was first pointed out to the author by J. A. Wheeler (private communication).

[20]J. D. Bekenstein, Phys. Rev. D 7, 949 (1973).

[21]See for example, A. M. Volkov, A. A. Izmest'ev, and G. V. Skrotskii, Zh. Eksp. Teor. Fiz. 59, 1254 (1970) [Sov. Phys. JETP 32, 686 (1971)].

[22]See E. Hisdal, Phys. Norv. 5, 1 (1971), and references cited therein.

[23]In Ref. 12 we gave an alternate resolution of the paradox posed by Geroch based on the apparent tendency of the black-hole area to increase as the box is being lowered.[11] The present approach is preferable in that it is independent of the validity of the interpretation given in Ref. 11, and in that it fits very well into the thermodynamic approach of this paper as will be evident presently.

[24]See M. Davis, R. Ruffini, and J. Tiomno, Phys. Rev. D 5, 2932 (1972) for the radiation of a particle falling radially into a Schwarzschild black hole. It is not clear whether the large amount of radiation found to go down the black hole is a device of the linearized approximation, or whether the effect will persist for other types of orbits or for Kerr black holes. Therefore we do not base any of our arguments here on this effect as we did in Refs. 11 and 12.

[25]B. Carter, Phys. Rev. 174, 1559 (1968).

[26]J. M. Bardeen, Astrophys. J. 161, 103 (1970).