Black Hole Entropy from Loop Quantum Gravity

Carlo Rovelli

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 7 June 1996)

We study the idea that the statistical entropy *governing thermal interactions of a black hole with its exterior* is determined by the microstates of the hole having distinct effects on the exterior, and over which a hole in a given macroscopic configuration thermally fluctuates. We argue that for a (macroscopically) Schwarzschild black hole this ensemble is formed by horizons with the same area. We compute the number of states in this ensemble from first principles using nonperturbative loop quantum gravity. We obtain a statistical entropy proportional to the area, as in the Bekenstein-Hawking formula. [S0031-9007(96)01421-4]

PACS numbers: 04.70.Dy, 04.60.Ds

In this Letter, we present a derivation of the Bekenstein-Hawking expression for the entropy [1] of a Schwarzschild black hole of surface area A

$$S = c(k/\hbar G)A \tag{1}$$

(c is a constant of the order of unity, G, k, and $\hbar$ the Newton, Boltzmann and Planck constants, and we put the speed of light equal to 1) via a statistical mechanical computation from a full theory of quantum gravity [2]. We use loop quantum gravity [3], and, in particular, we make use of the spectrum of the area operator, recently computed [4,5]. The loop approach to quantum gravity is now developed to the point where one may begin to try it within concrete physical problems, with the aim of getting some insights on the quantum physics of gravity, as well as testing the approach itself.

Our strategy is based on the idea that the entropy of the hole originates from the microstates of the horizon that correspond to a given macroscopic configuration. As far as we know this idea was first suggested in a seminal work by York [6]. York notices that the hole's radiance implies that the (macroscopic) event horizon is located slightly inside the quasistatic timelike limit surface, leaving a thin shell between the two, which he proposes to interpret as the region over which the microscopic horizon fluctuates. He interprets these fluctuations as zero point quantum fluctuations of the horizon's quasinormal modes, and, by identifying the thermal energy of these oscillations with the shell's ("irreducible") mass, he is able to recover Hawking's temperature. We take two essential ideas from York's work: that the source of the hole entropy is in the degrees of freedom associated with the fluctuations of the shape of the (microscopic) horizon, and that the quasilocal measure of mass energy governing energetic exchanges between the horizon and its surroundings can be taken as the Christodoulou-Ruffini [7] "irreducible mass" M_{CR}. Using this, our aim here is to replace York's perturbative semiclassical approach with a direct calculation within nonperturbative quantum gravity.

The relevance of horizon's surface degrees of freedom for the entropy has been recently explored from various perspectives [8]. (See also [9] for an attempt to use the "membrane paradigm" [10]: interactions of a black hole with its surroundings can be described in terms of a fictitious physical membrane located close to the horizon.) An approach strictly related to ours has been suggested in Ref. [11], where it is argued that a physical split of a gauge system gives rise to boundary degrees of freedom, since the boundary breaks the gauge group. Using this idea the Bekenstein-Hawking formula can be derived, by counting boundary states, in 3D gravity. In general relativity, the broken component of the gauge group includes diffeomorphisms that move the surface, and the boundary degrees of freedom can probably be viewed as fluctuations of the horizon.

In this Letter, we present a general discussion supporting York's idea that the hole's entropy is determined by an ensemble of microstates of the horizon, and that this ensemble is formed by the geometries of the horizon with the same M_{CR}. Then we compute the number of these geometries in the loop representation.

Consider a physical system containing a nonrotating and noncharged black hole (say, a collapsed star) as well as other physical components such as dust, gas, or radiation, which we denote collectively as "matter." We are interested in the statistical thermodynamics of such a system. Due to Einstein's equations, the microscopic time-dependent inhomogeneities of the matter distribution generate time-dependent "microscopic" inhomogeneities in the gravitational field as well. In most physical problems, one can safely disregard these minute ripples of the geometry. However, in a statistical-thermodynamical treatment, minute fluctuations should not be disregarded: they are the source of the thermal behavior.

Thus, we have *two descriptions* of a physical black hole interacting with surrounding matter at finite temperature: the macroscopic description, stationary and coarse grained, and the microscopic description, where individual thermal fluctuations are not disregarded. Macroscopically, the noncharged nonrotating hole is described by a Schwarzschild metric with mass m, and horizon

 0031-9007/96/77(16)/3288(4)$10.00

area $A = 16\pi G^2 m^2$. Thus, *in a thermal context, the Schwarzschild metric represents the coarse-grained description of a microscopically fluctuating geometry*. Microscopically, the gravitational field is described by some complicated time-dependent non-spherically-symmetric metric. We believe that taking time-dependent nonsymmetric microstates into account is essential for a statistical understanding of the thermal behavior of black holes: Searching a statistical derivation of black hole thermodynamics from properties of spherically symmetric metrics alone is like trying to derive the thermodynamics of an ideal gas in a spherical box from the spherically symmetric motions of the molecules.

Consider a microstate of the system. Foliate spacetime with a family of spacelike surfaces Σ_t, labeled by a time coordinate t. The intersection h_t between the surface Σ_t and the boundary of the past of future null infinity defines the instantaneous (microscopic) configuration of the event horizon at time t. Thus, h_t is a closed 2D surface immersed in Σ_t. For most times, this microscopic configuration of the event horizon is not spherically symmetric. Let us denote by g_t the intrinsic geometry of the horizon h_t. Let $\mathcal{M}$ be the space of all possible geometries of a 2D surface. As t changes, the (microscopic) geometry of the horizon changes. Thus, g_t wanders in $\mathcal{M}$ as t changes.

Consider a closed thermodynamical system S ideally split into two subsystems S_1 and S_2. Let us *ideally* isolate the subsystem S_1. Call its energy E. Its entropy $S(E)$, defined by the log of the number of microstates that have energy E, governs the thermal behavior of S_1 in its interactions with S_2 (microcanonical). Let us apply this idea to our system. We consider our system to be formed by two subsystems: the hole and the rest. We want to associate an entropy S to the hole, describing its thermal exchanges with the exterior. S must count the number of microstates over which the hole would fluctuate in an ideal situation in which no energy is exchanged between the hole and its surroundings. The precise specification of this ensemble of microstates is crucial, and we now discuss it in detail.

First, configurations of the hole itself, and not that of the surrounding geometry, should affect the hole's entropy. Next, the behavior of a system containing the hole is not affected by the hole's interior. The hole interior, indeed, might be in one out of an infinite number of states without distinguishable effects on the outside. For instance, it might (in principle) be given by a Kruskal-like spacetime, with another "universe" (say, spatially compact, apart for the hole) with billions of galaxies on the other side. This huge number of internal states does not affect the interaction of the hole with its surroundings and it is therefore irrelevant here. Thus, we are interested only in configurations of the hole *that have (microscopically) distinct effects on the exterior*. From the exterior, the hole's future behavior is fully determined by the geometry of its surface. Thus, the entropy governing the thermal interactions of the hole is determined by the state of the geometry on the hole's surface, namely, by g_t.

Next, we have to determine the ensemble of the microstates g_t over which the hole would fluctuate under the ideal hypothesis of no energy exchange. The usual microcanonical ensemble is determined by fixing energy. Here, we must look for a notion of energy associated to the horizon's surface, governing energetic exchanges with the exterior. Following York, we take the Christodoulou-Ruffini quasilocal irreducible mass $M_{\mathrm{CR}} = \sqrt{A/16\pi G^2}$ as the relevant energy in this context (here A is the area of h_t), and we define the ensemble as the set of g_t in $\mathcal{M}$ with the same M_{CR}, namely, with the same area. There is a number of reasons supporting the choice of this ensemble. First, M_{CR} is geometrically well defined, governs the hole's energy exchanges, and agrees with the macroscopic black hole energy. Second, the ensemble must contain reversible paths only. In the classical theory these conserve area (Hawking theorem [12]). Quantum theory allows classically forbidden energy exchanges with the exterior (Hawking radiance), but it is unlikely, we believe, that it would allow a nonreversible evolution of the horizon to become reversible without energy exchange with the exterior. Third, we may reason backward and let the thermodynamics indicate to us the correct ensemble (which is how classical ensembles were first found). In this context, it is perhaps worthwhile recalling that difficulties to rigorously justifying the choice of the ensemble *a priori* plague conventional thermodynamics anyway.

A point to take into account (missed in an earlier stage of this work [13]) is that distinct physical regions on the black hole surface are distinguishable from each other for an external observer (observing the microstate). Indeed, consider initial data for the Einstein equations given in an asymptotically flat space containing a horizon. Consider data corresponding to a nonspherical localized deformation of the horizon. Then the future evolution of the field—say the radiation at future infinity—is affected by the location of the deformation on the event horizon. This fact seems to contradict diffeomorphism invariance, but it does not: The group G of the diffeomorphisms of the horizon acts on the horizon geometry as well as on the external geometry; we must factor away its action once, not twice. [If a group G acts (freely) on a set A and on a set B, then $\frac{A\times B}{G}$ is isomorphic to $A/G \times B$, and not to $A/G \times B/G$, as one could naively expect [14].] Therefore location on the horizon *relative to the exterior geometry* is gauge invariant.

Summarizing, we are interested in counting the number $N(A)$ of states of the geometry g_t of a surface h_t of area A, where different physical regions of h_t are distinguished from each other by the external geometry. (Notice that fluctuations widely away from uniformity—spherical symmetry—are presumably negligible, as is common in statistical mechanics.) The above discussion indicates then that $S(A) = k \ln N(A)$ is the entropy governing the horizon's thermal interactions with its surroundings. The "number" $N(A)$ is meaningless in the classical theory. As the entropy of the electromagnetic field in a cavity is well defined only if we take quantum theory into account, similarly we expect that $N(A)$ will be well defined in quantum gravity. The problem is thus to count the number of (orthogonal) quantum states of the geometry of a two dimensional surface, having total area A, that is, the dimension of the A eigenspace of the area operator. The problem is well defined, and can be translated into a direct computation, provided that a quantum theory of geometry is available [15].

In loop quantum gravity, the quantum states of the gravitational field are represented by s knots [16]. See [5] for a detailed recent introduction. An s knot is an equivalence class under diffeomorphisms of graphs immersed in space, carrying colors on their edges [corresponding to irreducible representations of SU(2)], and colors on their vertices (corresponding to invariant couplings between such representations). The relation between s knots and classical geometries was explored in [17].

If a surface Σ is given, its geometry is determined by its intersections with the s knot. Intersections are of three types: (a) an edge crosses the surface, (b) a vertex lies on the surface, and (c) a finite part of the s knot lies on the surface. Intuitively, type (a) is the only "generic" case, and we should disregard states of type (b) and (c). Ashtekar has suggested an argument for neglecting type (b) and (c) intersections [18]: we wish to describe the geometry of a fluctuating surface Σ as observed from the exterior, and we expect the state of its geometry to be stable under infinitesimal deformations of Σ. We may thus consider the surface as the limit of a sequence of surfaces Σ_ϵ, and its state as the (Hilbert norm) limit of the states of Σ_ϵ. Clearly, states of type (b) and (c) cannot appear in this way, and therefore we have to restrict our computation to states having intersections of type (a) only [19].

Given a quantum state and a surface, let $i = 1, \ldots, n$ label type (a) intersections, and p_i be the color of the edge through i. Thus, the quantum geometry of the surface is characterized by an n-tuple of n colors $\vec{p} = (p_1, \ldots, p_n)$, where n is arbitrary. In particular, it was shown in [4] that the total area of the surface Σ is

$$A = \sum_{i=1,n} 8\pi\hbar G \sqrt{p_i(p_i + 2)}. \tag{2}$$

Since physical points of Σ are distinguished by the external geometry, two sets of intersections attached to the external geometry in different ways define two states with (microscopically) distinct effects on the outside. (The permutation group of the intersections is the subgroup of G acting nontrivially, and, as shown above, this is *not* gauge.) Therefore, the quantum geometry on the surface is determined by the *ordered* n-tuples of integers $\vec{p} = (p_1, \ldots, p_n)$. Our task is reduced to the task of counting the ordered n-tuples of integers $\vec{p}$ such that (2) holds. (The first suggestion that this number may determine the black hole entropy is in [20].) More precisely, we are interested in the number of microstates (n-tuples $\vec{p}$) such that the right hand side of (2) is between A and $A + dA$, where $A \gg dA \gg \hbar G$.

Let $M = A/8\pi\hbar G$, and let $N(M)$ be the number of ordered n-tuples $\vec{p}$, with arbitrary n, such that

$$\sum_{i=1,n} \sqrt{p_i(p_i + 2)} = M. \tag{3}$$

First, we overestimate $N(M)$ by approximating the left hand side of (3) dropping the $+2$ term under the square root. Thus, we want to compute the number $N_+(M)$ of ordered n-tuples such that

$$\sum_{i=1,n} p_i = M. \tag{4}$$

The problem is an exercise in combinatorics. It can be solved, for instance, by noticing that if $(p_1, \ldots, p_n)$ is a partition of M [that is, it solves (4)], then $(p_1, \ldots, p_n, 1)$ and $(p_1, \ldots, p_n + 1)$ are partitions of $M + 1$. Since all partitions of $M + 1$ can be obtained in this manner, we have $N_+(M + 1) = 2N_+(M)$. Therefore $N_+(M) = C2^M$, where C is a constant. For large M,

$$\ln N_+(M) = (\ln 2)M.$$

Next, we underestimate $N(M)$ by approximating (3) as $\sqrt{p_i(p_i + 2)} = \sqrt{(p_i + 1)^2 - 1} \approx (p_i + 1)$. Thus, we wish to compute the number $N_-(M)$ of ordered n-tuples such that

$$\sum_{i=1,n} (p_i + 1) = M.$$

Namely, we have to count the partitions of M in parts with two or more elements. This problem can be solved by noticing that if $(p_1, \ldots, p_n)$ is one such partition of M and $(q_1, \ldots, q_m)$ is one such partition of $M - 1$, then $(p_1, \ldots, p_n + 1)$ and $(q_1, \ldots, q_m, 2)$ are partitions of $M + 1$. All partitions of $M + 1$ in parts with two or more elements can be obtained in this manner; therefore $N_-(M + 1) = N_-(M) + N_-(M - 1)$. It follows that $N_-(M) = Da_+^M + Ea_-^M$, where D and E are constants and $a_\pm$ are the two roots of the equation $a_\pm^2 = a_\pm + 1$. For large M the term with the highest root dominates

$$\ln N_-(M) = (\ln a_+)M = \ln[(1 + \sqrt{5})/2]M.$$

By combining the information from the two estimates, we conclude that $\ln N(M) = dM$, where

$$\ln[(1 + \sqrt{5})/2] < d < \ln 2 \quad \text{or} \quad 0.48 < d < 0.69.$$

(For another derivation, see [20].) Since the integers M are equally spaced, our computation yields immediately the density of microstates. The number $N(A)$ of microstates with area A grows for large A as $\ln N(A) = dA/8\pi\hbar G$. This gives immediately the Bekenstein-Hawking formula (1), with the constant of proportionality

$c = d/8\pi \sim 1/16\pi$. This is roughly 4π times smaller than Hawking's value $c = 1/4$.

Several issues remain open. We have worked in the simplified setting of a hole interacting with a given geometry, instead of working within a fully generally covariant statistical mechanics [21]. Also, it would be nice to have a direct characterization of the event horizon in the quantum theory: this could perhaps be given as the boundary between the edges (and vertices) of (each component of) the weave, whose modification does or does not affect expectation values of observables at future null infinity [22]. This approach might clarify the issue of the type (b) and (c) intersections. Finally, the numerical discrepancy with the Hawking's value indicates that something is still poorly understood. Jacobson [23] has suggested that finite renormalization effects of the Newton constant might account for this discrepancy and has begun to explore how the presence of matter might affect it. In summary, we have argued that black hole entropy is determined by the dimensions of the eigenspaces of the area operator. Using loop quantum gravity, we have shown that this entropy is proportional to the area.

I thank Abhay Ashtekar, John Baez, Riccardo Capovilla, Ted Jacobson, Kirill Krasnov, Luis Lenher, Don Marolf, Slava Mukhanov, Ted Newman, Jorge Pullin, Lee Smolin, Daniel Sudarski, and Ranjeet Tate for important discussions.

[1] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973); S. W. Hawking, Nature (London) **248**, 30 (1974). For a review of black hole thermodynamics, see R. M. Wald, *Quantum Field Theory on Curved Spacetime and Black Hole Thermodynamics* (University of Chicago Press, Chicago, 1994).

[2] For other attempts to derive the BH formula, see G. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2752 (1977); W. H. Zurek and K. S. Thorne, Phys. Rev. Lett. **54**, 2171 (1985); L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1987); R. M. Wald, Phys. Rev. D **48**, R3427 (1993); M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993); J. D. Bekenstein, Report No. gr-qc/9409015; R. D. Sorkin, Report No. gr-qc/9508002; G. Lifschytz and M. Ortiz, Report No. hep-th/9510115; F. Larsen and F. Wilczek, Phys. Lett. B **375**, 37 (1996); A. Strominger and C. Vafa, Report No. hep-th/9601029; G. Horowitz and A. Strominger, Report No. hep-th/9602051.

[3] C. Rovelli and L. Smolin, Phys. Rev. Lett. **61**, 1155 (1988); Nucl. Phys. **B331**, 80 (1990). For various perspectives on loop quantum gravity, see C. Rovelli, Class. Quantum Gravity **8**, 1613 (1991); A. Ashtekar, in *Gravitation and Quantization, Les Houches 1992*, edited by B. Julia and J. Zinn-Justin (Elsevier, Paris, 1995); L. Smolin, in *Quantum Gravity and Cosmology*, edited by J. Perez-Mercader, J. Sola, and E. Verdaguer (World Scientific, Singapore, 1992).

[4] C. Rovelli and L. Smolin, Nucl. Phys. **B442**, 593 (1995); A. Ashtekar and J. Lewandowski, Report No. gr-qc/9602046; L. Lenher, S. Frittelli, and C. Rovelli, "The Complete Spectrum of the Area in Loop Quantum Gravity," Pitt University, 1996 (to be published); R. Loll, Phys. Rev. Lett. **75**, 3048 (1995); Nucl. Phys. **B460**, 143 (1996).

[5] R. DePietri and C. Rovelli, Phys. Rev. D **54**, 2664 (1996).

[6] J. W. York, Phys. Rev. D **28**, 2929 (1983).

[7] D. Christodoulou, Phys. Rev. Lett. **25**, 1596 (1970); D. Christodoulou and R. Ruffini, Phys. Rev. D **4**, 3552 (1971).

[8] G 't Hooft, Report No. gr-qc/9402037; L. Susskind, Report No. hep-th/9309145; L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **48**, 3743 (1993); C. Teitelboim, Report No. hep-th/9510180; A. Buonanno, M. Gattobigio, M. Maggiore, L. Pilo, and C. Ungarelli, Nucl. Phys. **B451**, 677 (1995).

[9] M. Maggiore, Nucl. Phys. **B429**, 205 (1994).

[10] K. Thorne, R. H. Price, and D. A. Macdonald, *Black Holes: The Membrane Paradigm* (Yale University Press, New Haven, 1986).

[11] D. Carlip, Phys. Rev. D **51**, 632 (1995); Report No. gr-qc/9606043; A. P. Balachandran, L. Chandar, and A. Momen, Nucl. Phys. **B461**, 581 (1996); Report No. gr-qc/9512047.

[12] S. Hawking, Commun. Math. Phys. **25**, 152 (1972).

[13] C. Rovelli, *Black Hole Entropy from Loops?*, seminar at the Center for Gravitation and Geometry, Penn State University, November 1995 (unpublished).

[14] I thank Don Marolf for clarifying this point.

[15] For an overview of current ideas on quantum gravity and quantum geometry, see C. Isham, Report No. gr-qc/9510063; special issue on *Quantum Geometry and Diffeomorphism Invariant Quantum Field Theory*, edited by C. Rovelli and L. Smolin [J. Math. Phys. (N.Y.) **36** (1995)].

[16] C. Rovelli and L. Smolin, Phys. Rev. D **52**, 5743 (1995).

[17] A. Ashtekar, C. Rovelli, and L. Smolin, Phys. Rev. Lett. **69**, 237 (1992).

[18] A. Ashtekar (private communication).

[19] Type (b) and (c) intersections seem to add a divergent contribution to S. R. DePietri and C. Rovelli have computed the finite contribution of type (b) intersections using a technique developed in [20]. It raises S by a factor of 4 approximately (unpublished).

[20] K. Krasnov, Report No. gr-qc/9603025; Report No. gr-qc/9605047.

[21] C. Rovelli, Class. Quantum Gravity **10**, 1549, 1567 (1993); A. Connes and C. Rovelli, Class. Quantum Gravity **11**, 2899 (1994).

[22] S. Frittelli, C. N. Kozameh, E. T. Newman, C. Rovelli, and R. S. Tate, Report No. gr-qc/9603061.

[23] T. Jacobson, seminar at the Schrödinger Institute, Vienna, July 1996 (unpublished).