Black Hole Entropy and Higher-Curvature Interactions

Ted Jacobson[(a)]
Institute for Theoretical Physics, University of California, Santa Barbara, California 93106
and Department of Physics, University of Maryland, College Park, Maryland 20742-4111

Robert C. Myers[(b)]
Department of Physics, McGill University, Montreal, Quebec, Canada H3A 2T8
(Received 5 April 1993)

A general formula for the entropy of stationary black holes in Lovelock higher-curvature gravity theories is obtained by integrating the first law of black hole mechanics, which is derived by Hamiltonian methods. The entropy is not simply one quarter of the surface area of the horizon, but also includes a sum of intrinsic curvature invariants integrated over a cross section of the horizon.

PACS numbers: 04.50.+h, 05.70.Ce, 97.60.Lf

The study of black hole thermodynamics is motivated primarily by the hope of learning something about the nature of quantum gravity. Probing the limits of validity of the four laws of black hole thermodynamics [1] may provide one source of insight in this context. For instance, one would like to know to what extent these laws are valid when the back-reaction of quantum field energy is taken into account. Back-reaction leads to the consideration of higher-curvature interactions, which arise from quantum field renormalization [2]. This observation motivates the investigation of black hole thermodynamics in higher-curvature theories of gravity. Another motivation comes from string theory, where the low energy effective field theory of gravity contains higher-curvature terms [3].

In this Letter, we focus on a special class of higher-curvature theories, called Lovelock gravity [4]. These theories are the most general second-order gravity theories in higher dimensional spacetimes. Higher dimensional spacetimes are of interest in several candidate frameworks for unifying gravity with other interactions. Moreover in higher dimensions with Lovelock theory, one can explore the effect of higher-curvature terms in black hole thermodynamics without having to deal with complications that arise in true higher derivative theories. The Lagrangian density for Lovelock gravity in a spacetime of dimension D may be written $\mathcal{L}=\sum_{m=0}^{[D/2]} c_m\mathcal{L}_m$ where

$$\mathcal{L}_m(g)=\frac{1}{2^m}\sqrt{-g}\;\delta^{a_1b_1\cdots a_mb_m}_{c_1d_1\cdots c_md_m}\;R_{a_1b_1}{}^{c_1d_1}\cdots R_{a_mb_m}{}^{c_md_m} \tag{1}$$

and $\mathcal{L}_0=\sqrt{-g}$ [5], and in the sum, $[D/2]$ indicates the integer part of $D/2$. The δ symbol is a totally antisymmetric product of $2m$ Kronecker deltas, normalized to take values 0 and ± 1. Note $\mathcal{L}_1=\sqrt{-g}R$ by itself yields the Einstein Lagrangian. In general, $\mathcal{L}_m$ is the Euler class for a $2m$-dimensional manifold. Because of the antisymmetrization, no derivative appears at higher than second order in the equations of motion. Static, spherically symmetric black hole solutions have been found for Lovelock gravity [6, 7], and given these solutions, one can examine to what extent the laws of black hole thermodynamics apply [8].

Using the Euclidean approach [9], it is clear many of the essential features survive. Given a specific stationary solution, a Euclidean section is obtained by analytic continuation. Regularity of this section requires identifying the Killing time coordinate with a period $\beta=2\pi/\kappa$ (κ defined below). This section is regarded as a background in a periodic Euclidean path integral, which is interpreted as giving a thermodynamic partition function Z with inverse temperature β. In a semiclassical approximation, the Euclidean action I, regarded as a function of β and the extensive parameters (such as the angular momentum, J), is identified with $-\ln Z$. One defines the internal energy and entropy as

$$U\equiv\frac{\partial I}{\partial\beta},\quad S\equiv\beta\frac{\partial I}{\partial\beta}-I\,, \tag{2}$$

which then automatically satisfy the first law

$$T\delta S=\delta U-\left(T\frac{\partial I}{\partial J}\right)\delta J.$$

This result reproduces the first law of black hole mechanics if $U=M$, the black hole mass, and $(T\,\partial I/\partial J)=\Omega$, the angular velocity of the horizon. The latter relations appear to hold in the present context (i.e., higher curvatures and/or higher dimensions), and it may be possible to prove them by extending the methods of Ref. [10]. One might expect that S coincides with one quarter the surface area of the horizon [11] as in Einstein gravity, but this identity fails in Lovelock gravity [7, 8], and other higher-curvature theories [12].

As described above, the Euclidean approach is applied to investigate known stationary solutions. While this yields S as a function of the solution's parameters, it provides no guidance as to any *geometric* significance of the black hole entropy. It also fails to provide a definition for S that generalizes to an arbitrary, time-dependent horizon, without which it is difficult to see how to even address the question of whether a classical second law is obeyed in "nonequilibrium" processes. However, it is

0031-9007/93/70(24)/3684(4)$06.00

also possible to work with the Euclidean approach without specific solutions in mind [10, 13]. (Indeed, this has recently been done for static metrics in Lovelock gravity with only c_1 and c_2 nonvanishing [14].) In this more general fashion, it may be possible to avoid the above mentioned drawbacks of the Euclidean approach.

Our alternate approach is to find the *variation* δS of the entropy by deriving a general form of the first law of black hole mechanics following the Hamiltonian method introduced by Sudarsky and Wald [15]. The method of Ref. [15] applies to *Killing horizons.* A Killing horizon is a null hypersurface whose null generators are tangent to a Killing field. In four-dimensional Einstein gravity, Hawking proved that a stationary black hole event horizon must be a Killing horizon [16]. This proof cannot obviously be generalized to Lovelock gravity, but the result is certainly true for all the known solutions (which, however, are all static).

If χ^a is the Killing field which is null on the horizon, the surface gravity κ can be defined by $\chi^b\nabla_b\chi^a = \kappa\chi^a$. Assuming that the null generators of the horizon can be extended to the past as complete geodesics, three important properties follow: (i) κ is *constant* over the entire horizon [17]. (ii) There exists a $(D-2)$-dimensional spacelike cross-section B of the horizon on which χ^a vanishes [17]. B is called the *bifurcation surface.* (iii) The extrinsic curvature of B vanishes [18]. A further property of Killing horizons that is important for our purposes is that any two spacelike slices of the horizon are isometric.

Let ξ^a be the Killing field that is asymptotically a pure time translation. Then by suitably rescaling χ^a, one is able to reduce the Killing field $\chi^a - \xi^a$ to a rotation. In higher dimensions, there are $[(D-1)/2]$ commuting Killing fields $\phi^a_{(\alpha)}$ generating rotations in totally orthogonal planes [19], and so one has

$$\chi^a = \xi^a + \Omega^{(\alpha)}\phi^a_{(\alpha)} \ .$$

Next, we briefly present the Hamiltonian description of Lovelock gravity [20], which will be necessary in the following derivation of the first law. As usual after splitting space and time, the dynamical variables on a spatial surface Σ are the spatial metric h_{ab}, and its conjugate momentum π^{ab} [5]. In the present case, π^{ab} is a complicated function of both the extrinsic and intrinsic curvatures of Σ [20]. The normal and spatial components of the time flow vector field t^a, the lapse N and shift N^a, are arbitrary Lagrange multipliers in the Hamiltonian. For asymptotically flat space, the Hamiltonian has two parts $H = H_V + H_S$. The volume term H_V is a combination of constraints $\mathcal{H}_\perp$ and $\mathcal{H}_a$,

$$H_V = \int_\Sigma d^{D-1}x \,(N\mathcal{H}_\perp + N^a\mathcal{H}_a) \ ,$$

and hence vanishes when evaluated for solutions of the field equations. $\mathcal{H}_a$ is the generator of spatial diffeomorphisms in Σ, and so, as in Einstein gravity, $\mathcal{H}_a = -2D_b(\pi_a{}^b)$, where D_a is the covariant derivative compatible with h_{ab}. $\mathcal{H}_\perp$ generates normal deformations of Σ, and one finds $\mathcal{H}_\perp = \Sigma c_m \mathcal{H}_\perp^{(m)}$ with

$$\mathcal{H}_\perp^{(m)} = -\frac{\sqrt{h}}{2^m}\, \bar{\delta}^{a_1 b_1 \cdots a_m b_m}_{c_1 d_1 \cdots c_m d_m}\; R_{a_1 b_1}{}^{c_1 d_1} \cdots R_{a_m b_m}{}^{c_m d_m}$$

where $R_{ab}{}^{cd}$ is the full D-dimensional curvature tensor. Since $R_{ab}{}^{cd}$ is projected into Σ, it can be replaced by $\bar{R}_{ab}{}^{cd} + 2K_{[a}{}^c K_{b]}{}^d$. Here $\bar{R}_{ab}{}^{cd}$ is the curvature of h_{ab}, and K_{ab} is the extrinsic curvature of Σ, which is regarded as a function of h_{ab} and π^{ab} [21]. The $\bar{\delta}$ symbol is the antisymmetric product of projected Kronecker deltas, $\bar{\delta}^a_c = \delta^a_c + n^a n_c$ where n^a is the unit normal to Σ.

The Hamiltonian also has an asymptotic boundary term, just as in Einstein gravity [22], $H_S = \oint_\infty d^{D-2}x\,(NS_\perp + N^a S_a)$. H_S is added to cancel surface terms which arise by integrating by parts in H_V when generating the Hamiltonian flow equations. The precise value of H_S depends on the choice of $t^a = Nn^a + N^a$. If asymptotically $t^a \to \xi^a$, the time translation Killing field, then $H_S = M$, the total mass of the solution. If $t^a \to \phi^a_{(\alpha)}$, a rotational Killing field, then $H_S = -J_{(\alpha)}$, the associated angular momentum.

The derivation of the first law comes from a judicious evaluation of the Hamiltonian [15]. Let h_{ab} and π^{ab} describe a stationary black hole solution, and choose the "time" flow field to be the Killing field that is null on the horizon, $t^a = \chi^a = \xi^a + \Omega^{(\alpha)}\phi^a_{(\alpha)}$. Evaluate H on a spacelike slice Σ that extends from asymptotic infinity to the horizon, intersecting the horizon at the bifurcation surface, B. With these choices one has $H = M - \Omega^{(\alpha)}J_{(\alpha)}$. Let $(\delta h_{ab}, \delta\pi^{ab})$ be a perturbation of the initial solution to *any* nearby solution—not necessarily stationary. With t^a and Σ fixed, one has

$$\delta H = \delta M - \Omega^{(\alpha)}\delta J_{(\alpha)} \ .$$

On the other hand, one has Hamilton's equations

$$\delta H = \int_\Sigma d^{D-1}x\,\left(\mathcal{L}_\chi h_{ab}\,\delta\pi^{ab} - \mathcal{L}_\chi \pi^{ab}\,\delta h_{ab}\right) + \delta H_B \ .$$

The volume term vanishes because the flow is along a Killing field, $\mathcal{L}_\chi h_{ab} = 0 = \mathcal{L}_\chi \pi^{ab}$. Integration by parts is needed to produce these volume terms. The surface terms at asymptotic infinity are canceled by the variation δH_S, but one is left with surface terms, denoted δH_B, at the inner boundary of Σ.

Since δH_B is evaluated on the bifurcation surface B where $t^a = \chi^a$ vanishes, any nonvanishing terms involve derivatives of t^a. Such terms only arise from the metric variation of the curvatures in $N\mathcal{H}_\perp$. Two integrations by parts arise from $\delta\bar{R}_{ab}{}^{cd} = -2D_{[a}D^{[c}\delta h_{b]}{}^{d]} + \bar{R}_{ab}{}^{e[c}\delta h_e{}^{d]}$, where $\delta h_a{}^d = h^{db}\delta h_{ab}$. The first integration yields new volume integrals involving $D_a N$. In these integrals, the second integration by parts produces a boundary term at B involving $D_a N$. The contribution from $\mathcal{H}_\perp^{(m)}$ is

$$\delta H_B^{(m)} = \frac{mc_m}{2^{m-1}} \oint_B d^{D-2}x \left(\sqrt{\tilde{h}}\, v^{c_1} D_{a_1} N\, \bar{\delta}^{a_1 b_1 a_2 \cdots b_m}_{c_1 d_1 c_2 \cdots d_m} \delta h_{b_1}{}^{d_1} R_{a_2 b_2}{}^{c_2 d_2} \cdots R_{a_m b_m}{}^{c_m d_m} \right) , \tag{3}$$

where v^c is the unit normal to the bifurcation surface B, pointing into Σ, and $\tilde{h}$ is the determinant of $\tilde{h}_{ab}$, the induced metric on B. Now it is not hard to show that, at B, $D_a N = \kappa v_a$, where κ is the surface gravity. Since $v_{a_1} v^{c_1} \bar{\delta}^{a_1 b_1 \cdots b_m}_{c_1 d_1 \cdots d_m} = \tilde{\delta}^{b_1 \cdots b_m}_{d_1 \cdots d_m}$, the antisymmetrized product of Kronecker deltas $\tilde{\delta}^a_c = \delta^a_c + n^a n_c - v^a v_c$ projected into B, all remaining tensors are fully projected into B. The projected $\delta h_b{}^d$ yields $\delta \tilde{h}_b{}^d \equiv \tilde{h}^{de} \delta \tilde{h}_{be}$. Since the extrinsic curvature of the bifurcation surface B vanishes, projecting $R_{ab}{}^{cd}$ yields the intrinsic curvature $\tilde{R}_{ab}{}^{cd}$ of $\tilde{h}_{ab}$.

Thus far variations of the K_{ab} terms in $R_{ab}{}^{cd}$ were neglected. Regarded as a function of π^{ab} and h_{ab}, K_{ab} involves $\bar{R}_{ab}{}^{cd}$ and so could produce extra boundary contributions. These extra terms always vanish, however, since they contain at least one other factor of K_{ab} which vanishes when projected into B. Since κ is a constant on B, Eq. (3) thus becomes

$$\delta H_B^{(m)} = \frac{mc_m \kappa}{2^{m-1}} \oint_B d^{D-2}x \left(\sqrt{\tilde{h}} \tilde{\delta}^{b_1 \cdots a_m b_m}_{d_1 \cdots c_m d_m}\, \delta \tilde{h}_{b_1}{}^{d_1} \cdots \tilde{R}_{a_m b_m}{}^{c_m d_m} \right) .$$

We now recognize $\delta S^{(m)} \equiv (2\pi/\kappa) \delta H_B^{(m)}$ as the variation of $4\pi m c_m \oint_B d^{D-2}x\, \mathcal{L}_{m-1}(\tilde{h})$, with $\mathcal{L}_m$ as defined in Eq. (1).

Collecting all of the contributions to the entropy, our final result is the first law of black hole mechanics in Lovelock gravity, $\delta M - \Omega^{(\alpha)} \delta J_{(\alpha)} = (\kappa/2\pi)\delta S$, where

$$S = \sum_{m=1}^{[D/2]} 4\pi m\, c_m \oint d^{D-2}x\, \mathcal{L}_{m-1}(\tilde{h}) . \tag{4}$$

Our derivation of the first law relied heavily on the fact that the variation of the entropy is evaluated on the bifurcation surface B. Nevertheless, the integrated expression (4) for the entropy of a stationary black hole can be evaluated on *any* spacelike slice of the Killing horizon, since all such slices are isometric. Note that with $c_1 = 1/16\pi G$, as appropriate for the Einstein action, the first term in the sum is simply $A/(4G)$, where A is the surface area of the horizon. Remarkably the entropy (4) is identical in form to the original action, evaluated in a (Riemannian) space of dimension $D-2$, with c_{m-1} replaced by $4\pi m c_m$.

For even dimensions, we have included a nontrivial integration constant in Eq. (4). This constant is the contribution of $\mathcal{L}_{D/2-1}$, which yields the Euler constant of the cross section of the horizon. In four dimensions, this constant is fixed for all stationary black holes since the horizon must have the topology of $S^2 \times R$ [16]. In $D>4$, the horizon topology is not unique, so this topological entropy may have more significance. With this choice for integration constant one easily confirms that for the known black hole solutions [6], Eq. (4) produces the same result as Eq. (2) in the Euclidean approach [8] (provided, of course, that $c_{D/2}$ is identical for both approaches). These results are easily extended to include Maxwell or Yang-Mills fields [15].

Our results extend the framework of black hole thermodynamics in a natural way to Lovelock gravity. The zeroth law (i.e., the constancy of κ) holds modulo the assumption that the horizon generators are geodesically complete (or, if the black hole forms from collapse, that the horizon generators of the stationary solution which it approaches have this property). The first law defines an entropy that is localized in the intrinsic geometry of the horizon and is valid for all stationary black holes, at least provided the event horizon is a Killing horizon. For example, Eq. (4) applies for rotating black hole solutions, which are as yet unknown. If all the coefficients c_m are of the same order in units of some common length scale, then the area term ($m = 1$) dominates the entropy for black holes that are much larger than that scale. The entropy of small black holes can be negative, but for fixed c_m it is bounded below for the known solutions [8]. Such negative values are thermodynamically benign, since only changes of the entropy occur in the first law.

The second law for quasistationary processes follows from the first law, provided the quantity $\delta M - \Omega^{(\alpha)} \delta J_{(\alpha)}$ is positive. This is the case for fluxes of positive energy matter [23]. To control the sign for fluxes including gravitational energy would require a positive energy theorem for Lovelock gravity, which has not been established [24]. (Negative energy states would probably make the theory unstable, however.) Alternatively, the second law in Einstein gravity can be derived from Hawking's area theorem [16] which shows that, even in nonquasistationary processes, the area (and therefore the entropy) will never decrease. In Lovelock gravity the expression (4) for the entropy is at least meaningful on a slice of a time-dependent horizon, so one might hope to prove a nonequilibrium second law by similarly following the evolution of the horizon. This remains an interesting problem for future work. In this connection curiously, note that with $c_{D/2} > 0$ the topological term in Eq. (4) could lead to violations of the second law when two black holes coalesce, even in four dimensions.

When quantum fields are included, negative energy can be transferred to the black hole, as in Hawking evaporation [25] or "mining" [26] processes. The interesting question is then whether the *generalized* second law (GSL) $[\delta(S_{\rm BH} + S_{\rm outside}) \geq 0]$ holds. (Validity of the *classical* second law for $S_{\rm BH}$ alone would seem to be a prerequisite for validity of the GSL.) There exist arguments

[27] in support of the GSL in Einstein gravity, and these carry over to Lovelock gravity as far as the contribution of matter fields is concerned. However, these arguments apply only to quasistationary processes and, moreover, the gravitational contributions in Lovelock theories are not yet fully understood. Thus the validity of the GSL remains an important open problem.

Finally, Lovelock gravity is only a very special case of possible generally covariant gravity theories. The fact that the entropy is not one quarter the surface area, is already known in many other examples of theories with higher-curvature interactions [12]. If one derives a first law via the method of Sudarsky and Wald as we have done here, the variation of the entropy will again be given as an integral over the bifurcation surface of the horizon. However, this expression will not in general depend only on the intrinsic geometry of the horizon, and it is not even clear that it will be the variation of some quantity defined locally at the horizon (although, from the Euclidean approach, Ref. [14] argues that the entropy is *always* a quantity defined locally at the horizon). In fact, the particular properties of Lovelock gravity played a crucial role in our calculations establishing the local and intrinsic nature of the entropy.

We would like to acknowledge useful discussions with D. Brill, D. Brown, P. Chrusciel, S. Deser, G. Horowitz, G. Kang, I. Moss, J.Z. Simon, A. Strominger, R. Wald, and J. York. R.C.M. thanks the Institute for Theoretical Physics at UCSB for their hospitality during this work. R.C.M. was supported by NSERC of Canada, and Fonds FCAR du Québec. T.J. was supported by NSF Grant No. PHY91-12240. Research at the ITP, UCSB was supported by NSF Grant No. PHY89-04035.

(a) Electronic address: jacobson@umdhep.umd.edu
(b) Electronic address: rcm@hep.physics.mcgill.ca

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

[2] See, for example, N.D. Birrell and P.C.W. Davies, *Quantum Fields in Curved Space* (Cambridge Univ. Press, Cambridge, 1982).

[3] See, for example, M.B. Green, J.H. Schwarz, and E. Witten, *Superstring Theory* (Cambridge Univ. Press, Cambridge, 1987).

[4] D. Lovelock, J. Math. Phys. **12**, 498 (1971); **13**, 874 (1972).

[5] We set c_0, the cosmological constant, to zero, and consider only asymptotically flat spaces in the following. The spacetime signature is taken as $(- + \cdots +)$. Throughout the paper lower case latin letters are employed for all tensor indices, including those on "spatial" tensors. That is, spatial tensors are just regarded as spacetime tensors which vanish when any index is contracted with the normal(s) to the spatial submanifold.

[6] D.G. Boulware and S. Deser, Phys. Rev. Lett. **55**, 2656 (1985); J.T. Wheeler, Nucl. Phys. **B268**, 737 (1986); **B273**, 732 (1986).

[7] D.L. Wiltshire, Phys. Lett. **169B**, 36 (1986); Phys. Rev. D **38**, 2445 (1988).

[8] R.C. Myers and J.Z. Simon, Phys. Rev. D **38**, 2434 (1988); General Relativ. Gravity, **21**, 761 (1989); B. Whitt, Phys. Rev. D **38**, 3000 (1988).

[9] G.W. Gibbons and S.W. Hawking, Phys. Rev. D **15**, 2752 (1977); S.W. Hawking, in *General Relativity, An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1979).

[10] J.D. Brown and J.W. York, Phys. Rev. D **47**, 1407 (1993); **47**, 1420 (1993).

[11] In D dimensions, a cross section of the horizon is a $(D-2)$-dimensional space, and so the surface area refers to the volume of such a slice.

[12] C. Callan, R.C. Myers, and M. Perry, Nucl. Phys. **B311**, 673 (1988); R.C. Myers, Nucl. Phys. **B289**, 701 (1987); E. Poisson, Classical Quantum Gravity **8**, 639 (1991); M. Lu and M. Wise, Report No. hep-th/9301021, 1993 (to be published); M. Visser, Report No. hep-th/9303029, 1993 (to be published).

[13] B.F. Whiting and J.W. York, Phys. Rev. Lett. **61**, 1336 (1988).

[14] J.Z. Simon and B.F. Whiting (to be published).

[15] D. Sudarsky and R. Wald, Phys. Rev. D **46**, 1453 (1992). For a closely related method see also J.D. Brown, E.A. Martinez, and J.W. York, Phys. Rev. Lett. **66**, 2281 (1991).

[16] S.W. Hawking, *Commun. Math. Phys.* **25**, 152 (1972); S.W. Hawking and G.R.F. Ellis, *The Large Scale Structure of Spacetime* (Cambridge Univ. Press, Cambridge, 1973).

[17] I. Rácz and R.M. Wald, Classical Quantum Gravity **9**, 2643 (1992).

[18] R.M. Wald (unpublished).

[19] R.C. Myers and M.J. Perry, Ann. Phys. (N.Y.) **172**, 304 (1986).

[20] C. Teitelboim and J. Zanelli, Classical Quantum Gravity **4**, L125 (1987).

[21] In fact, one cannot in general solve uniquely for K_{ab} in terms of π^{ab} and h_{ab} [20], but this ambiguity does not enter into our analysis.

[22] T. Regge and C. Teitelboim, Ann. Phys. (N.Y.) **88**, 286 (1974).

[23] See, for example, B. Carter, in *General Relativity, An Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1979).

[24] For a discussion of energy positivity in Lovelock theories see Ref. [6] and S. Deser and Z. Yang, Classical Quantum Gravity **6**, L83 (1989).

[25] S.W. Hawking, Commun. Math. Phys. **43**, 199 (1975).

[26] W.G. Unruh and R.M. Wald, Phys. Rev. D **25**, 942 (1982).

[27] W. Zurek, Phys. Rev. Lett. **49**, 1683 (1982); W.H. Zurek and K.S. Thorne, Phys. Rev. Lett. **54**, 2171 (1985); R.M. Wald, in *Highlights in Gravitation and Cosmology*, edited by B.R. Iyer *et al.* (Cambridge Univ. Press, Cambridge, 1988); V.P. Frolov and D.N. Page, Report No. gr-qc/9302017, 1993 (to be published).