

## Classical and quantum thermodynamics of horizons in spherically symmetric spacetimes

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2002 Class. Quantum Grav. 19 5387

(<http://iopscience.iop.org/0264-9381/19/21/306>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 180.149.51.198

The article was downloaded on 01/10/2011 at 04:30

Please note that [terms and conditions apply](#).

# Classical and quantum thermodynamics of horizons in spherically symmetric spacetimes

**T Padmanabhan**

Inter-University Center for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind,  
Pune - 411 007, India

E-mail: nabhan@iucaa.ernet.in

Received 15 July 2002

Published 14 October 2002

Online at [stacks.iop.org/CQG/19/5387](http://stacks.iop.org/CQG/19/5387)

## Abstract

A general formalism for understanding the thermodynamics of horizons in spherically symmetric spacetimes is developed. The formalism reproduces known results in the case of black-hole spacetimes and can handle more general situations such as: (i) spacetimes which are not asymptotically flat (such as the de Sitter spacetime) and (ii) spacetimes with multiple horizons having different temperatures (such as the Schwarzschild–de Sitter spacetime) and provide a consistent interpretation for temperature, entropy and energy. I show that it is possible to write Einstein's equations for a spherically symmetric spacetime in the form  $T dS - dE = P dV$  near *any* horizon of radius  $a$  with  $S = \frac{1}{4}(4\pi a^2)$ ,  $|E| = (a/2)$  and the temperature  $T$  determined from the surface gravity at the horizon. The pressure  $P$  is provided by the source of Einstein's equations and  $dV$  is the change in the volume when the horizon is displaced infinitesimally. The same results can be obtained by evaluating the quantum mechanical partition function *without using Einstein's equations or the WKB approximation for the action*. Both the classical and quantum analyses provide a simple and consistent interpretation of entropy and energy for de Sitter spacetime as well as for (1 + 2) dimensional gravity. For the Rindler spacetime the entropy per unit transverse area turns out to be  $\frac{1}{4}$  while the energy is zero. The approach also shows that the de Sitter horizon—like the Schwarzschild horizon—is effectively one dimensional as far as the flow of information is concerned, while the Schwarzschild–de Sitter, Reissner–Nordstrom horizons are not. The implications for spacetimes with multiple horizons are discussed.

PACS numbers: 0470, 0462, 0460, 9880

## 1. The need for local description of horizon thermodynamics

This paper provides a classical and quantum description of a mini superspace of spacetimes which are spherically symmetric. Since the subject of thermodynamics of horizons has nearly three decades of history, it is probably best if I begin with a detailed description of motivation. Readers interested in concrete results can skip ahead to section 2.

The best-studied spacetimes with horizons are black hole spacetimes [1, 2]. In the simplest context of a Schwarzschild black hole of mass  $M$ , one can attribute an energy  $E = M$ , temperature  $T = (8\pi M)^{-1}$  and entropy  $S = \frac{1}{4}(A_H/L_P^2)$ , where  $A_H$  is the area of the horizon and  $L_P = (G\hbar/c^3)^{1/2}$  is the Planck length. (Hereafter, I will use units with  $G = \hbar = c = 1$ .) These are clearly related by the thermodynamic identity  $T dS = dE$ , usually called the first law of black-hole dynamics. This result has been obtained in much more general contexts and has been investigated from many different points of view in the literature. The simplicity of the result depends on the following features. (a) The Schwarzschild metric is a vacuum solution with no pressure so that there is no  $P dV$  term in the first law of thermodynamics. (b) The metric has only one parameter  $M$  so that changes in all physical parameters can be related to  $dM$ . (c) Most importantly, there is a well-defined notion of energy  $E$  to the spacetime and the changes in the energy  $dE$  can be interpreted in terms of the physical process of black hole evaporation. (One can also interpret the relations by, say, dropping ‘test particles’ into the black hole, but it is really not necessary.) The idea can be generalized to other black hole spacetimes in a rather simple manner only because of well-defined notions of energy, angular momentum, etc.

Can one generalize the thermodynamics of horizons to cases other than black holes in a straightforward way? In spite of years of research in this field, this generalization remains non-trivial and challenging when the conditions listed above are not satisfied. To see the importance of the above conditions, we only need to contrast the situation in Schwarzschild spacetime with that of de Sitter spacetime and observe that:

- (1) The notion of temperature is well defined in the case of de Sitter spacetime since the de Sitter horizon also radiates like a black body; we have  $T = H/2\pi$  where  $H^{-1}$  is the radius of the de Sitter horizon. (The same result can be obtained more formally in terms of the periodicity in the Euclidean time.) But the correspondence probably ends there. A study of literature shows that there exist very few concrete calculations of energy, entropy and laws of horizon dynamics in the case of de Sitter spacetimes, in sharp contrast to BH spacetimes.
- (2) There have been several attempts in the literature to define the concept of energy using local or quasi-local concepts (for a small subsample of ten references, see [3]). The problem is that not all definitions of energy (or even definitions of horizon; see [8]) agree with each other and not all of them can be applied to de Sitter type universes.
- (3) Even when a notion of energy can be defined, it is not clear how to write and interpret an equation analogous to  $dS = (dE/T)$  in this spacetime, especially since the physical basis for  $dE$  would require a notion of evaporation of the de Sitter universe.
- (4) Further, we know that de Sitter spacetime is a solution to Einstein’s equations with a source having non-zero pressure. Hence one would very much doubt whether  $T dS$  is indeed equal to  $dE$ . It would be necessary to add a  $P dV$  term for consistency.

An argument is sometimes advanced that the de Sitter horizon is conceptually different from the black hole horizon because it is observer dependent. I believe this argument is incorrect and that all horizons (including even the Rindler horizon) should be treated on a par for at least three reasons (see [4, 5] for more details). (i) To begin with, if the notion

of entropy in black hole spacetimes is not accidental, then one would expect *any* one-way membrane which blocks out information to lead to a notion of entropy [5]. (ii) As regards observer dependence, even in the case of Schwarzschild spacetimes, it is possible to have observers moving in time-like trajectories inside the event horizon who will access part of the information which is not available to the outside observer. It seems unlikely that these suicidal observers will attribute the same amount of entropy to the Schwarzschild black hole as an observer playing it safe by staying far away from the event horizon. (iii) If the notion of entropy associated with a one-way membrane arises from *local* degrees of freedom and Planck scale physics, our inability to define a *global* notion of energy should not have a bearing on the issue of entropy.

All these suggest that there must be a local approach by which one can define the notion of entropy *and* energy for spacetimes with horizons. This conclusion is strengthened further by the following argument: consider a class of spherically symmetric spacetimes of the form

$$ds^2 = f(r) dt^2 - f(r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2). \tag{1}$$

If  $f(r)$  has a simple zero at  $r = a$  with  $f'(a) \equiv B$  remaining finite, then this spacetime has a horizon at  $r = a$ . Spacetimes like Schwarzschild or de Sitter have only one free parameter in the metric (such as  $M$  or  $H^{-1}$ ) and hence the scaling of all other thermodynamical parameters is uniquely fixed by purely dimensional considerations. But, for a general metric of the form in (1), with an arbitrary  $f(r)$ , the area of the horizon (and hence the entropy) is determined by the location of the zero of the function  $f(r)$  while the temperature—obtained from the periodicity considerations—is determined by the value of  $f'(r)$  at the zero. For a general function, of course, there will be no relation between the location of the zero and the slope of the function at that point. It will, therefore, be incredible if there exists any *a priori* relationship between the temperature (determined by  $f'$ ) and the entropy (determined by the zero of  $f$ ) even in the context of horizons in spherically symmetric spacetimes. If we take the entropy to be  $S = \pi a^2$  (where  $f(a) = 0$  determines the radius of the horizon) and the temperature to be  $T = |f'(a)|/4\pi$  (determined by the periodicity of Euclidean time), the quantity  $T dS = \frac{1}{2}|f'(a)|a da$  will depend both on the slope  $f'(a)$  as well as the radius of the horizon. This implies that any local interpretation of thermodynamics will be quite non-trivial.

Finally, the need for local description of thermodynamics of horizons becomes crucial in the case of spacetimes with multiple horizons. Let me briefly describe this situation, relegating the details to appendix A. The strongest and the most robust result we have, regarding spacetimes with a horizon, is the notion of temperature associated with them. This, in turn, depends either on a complicated analysis of the mode functions of a wave equation, Bogoliubov coefficients, etc or on the study of the periodicity of the Euclidean time coordinate. Neither approach works very well if the spacetime has more than one horizon such as, for example, in the Schwarzschild–de Sitter metric which has the form in (1) with

$$f(r) = \left(1 - \frac{2M}{r} - H^2 r^2\right). \tag{2}$$

This spacetime has two horizons at  $r_{\pm}$  with

$$r_+ = \sqrt{\frac{4}{3}} H^{-1} \cos \frac{x + 4\pi}{3}; \quad r_- = \sqrt{\frac{4}{3}} H^{-1} \cos \frac{x}{3} \tag{3}$$

where  $\cos x = -3\sqrt{3}MH^{-1}$ . (The parameter  $x$  is in the range  $(\pi, \frac{3}{2}\pi]$  and  $0 \leq 27M^2H^{-2} < 1$ .) Close to either horizon the spacetime can be approximated as Rindler. Since the surface gravities on the two horizons are different, we get two different Rindler temperatures  $T_{\pm} = |f'(r_{\pm})|/4\pi$ . It is also possible to introduce two *different* Kruskal-like

coordinates using the two surface gravities in which the metric is well behaved at the horizons (see appendix A for details; for related work, see [6]). In the (1 + 1) case, it is possible to introduce the analogues of Boulware and Hartle–Hawking vacuum states with *either* of these transformations. In the overlapping region of the coordinate patches, however, there is no simple notion of temperature.

This point is brought in quite dramatically when we study the periodicity in the imaginary time in the overlapping region of the two Kruskal coordinates. To maintain invariance under  $it \rightarrow it + \beta$  (with some finite  $\beta$ ) it is necessary that  $\beta$  is an integer multiple of both  $4\pi/|f'(r_+)|$  and  $4\pi/|f'(r_-)|$  so that  $\beta = (4\pi n_{\pm}/|f'(r_{\pm})|)$  where  $n_{\pm}$  are integers. Hence the ratio of surface gravities  $|f'(r_+)/f'(r_-)| = (n_+/n_-)$  must be a rational number. Though irrationals can be approximated by rationals, such a condition definitely excludes a class of values for  $M$  if  $H$  is specified and vice versa. It is not clear why the existence of a cosmological constant should imply something for the masses of black holes (or vice versa). Since there is no physical basis for such a condition, it seems reasonable to conclude that these difficulties arise because of our demanding the existence of a finite periodicity  $\beta$  in the Euclidean time coordinate. This demand is related to an expectation of thermal equilibrium which is violated in spacetimes with multiple horizons having different temperatures.

If even the simple notion of temperature falls apart in the presence of multiple horizons, it is not likely that the notion of energy or entropy can be defined by *global* considerations. On the other hand, it will be equally strange if we cannot attribute a temperature to a black hole formed in some region of the universe just because the universe at the largest scales is, for instance, described by a de Sitter spacetime. One is again led to searching for a *local* description of the thermodynamics of *all types of* horizons.

The results in this paper have bearing on all these important issues. One of the key results of this paper will be to provide a consistent interpretation of the relation  $T dS - dE = P dV$  in the case of a de Sitter universe. To the extent I know, this has not been done before in any of the published literature explicitly, giving an expression for the energy in the de Sitter universe. I obtain the result  $E = -\frac{1}{2}H^{-1}$  by three separate arguments and these arguments work consistently for (1 + 2) dimensional spacetime as well. I will also show (in section 2) that one can use Einstein's constraint equations to relate  $f'(r)$  with  $r$  and thus provide a simple interpretation for the expression  $T dS$ . (The constraint equation arises essentially from the symmetries of the Einstein–Hilbert action under coordinate transformations and can be imposed as a constraint on physical states even in quantum theory.)

## 2. Einstein's equations as a thermodynamic identity

The key idea developed in this section is to use the notion of periodicity in Euclidean time (obtained by a local Rindler approximation near the horizon) to define the temperature, without any ambiguity in the proportionality constant, and then to rewrite Einstein's equations in a form analogous to the  $T dS - dE = P dV$  equation. I will show that it is fairly straightforward to achieve this in the case of spacetimes of the form

$$\begin{aligned} ds^2 &= f(r) dt^2 - f(r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \\ &\equiv f(r) dt^2 - f(r)^{-1} dr^2 - dL_{\perp}^2 \end{aligned} \quad (4)$$

with a general  $f(r)$  determined via Einstein's equations.

This metric will satisfy Einstein's equations provided the source stress tensor has the form

$$T_i^t = T_r^r = \frac{\epsilon(r)}{8\pi}; \quad T_{\theta}^{\theta} = T_{\phi}^{\phi} = \frac{\mu(r)}{8\pi}. \quad (5)$$

The equality  $T_t^t = T_r^r$  arises from our assumption  $g_{00} = (-1/g_{11})$  and can be relaxed if  $g_{00} \neq (-1/g_{11})$ . (It will turn out that our analysis goes through even in this more general case and I will comment on it in the next section.) The equality  $T_\theta^\theta = T_\phi^\phi$  arises from spherical symmetry. Equation (5) also defines the functions  $\epsilon(r)$  and  $\mu(r)$ . Einstein's equations now reduce to

$$\frac{1}{r^2}(1 - f) - \frac{f'}{r} = \epsilon; \quad \nabla^2 f = -2\mu. \tag{6}$$

The remarkable feature about the metric in (4) is that Einstein's equations become linear in  $f(r)$  so that solutions for different  $\epsilon(r)$  can be superposed. (I could not find this result in standard textbooks.) Given any  $\epsilon(r)$  the solution becomes

$$f(r) = 1 - \frac{a}{r} - \frac{1}{r} \int_a^r \epsilon(r)r^2 dr \tag{7}$$

with  $a$  being an integration constant and  $\mu(r)$  is fixed by  $\epsilon(r)$  through

$$\mu(r) = \epsilon + \frac{1}{2}r\epsilon'(r). \tag{8}$$

I have chosen the integration constant  $a$  in (7) such that  $f(r) = 0$  at  $r = a$  so that this surface is a horizon. It is, of course, quite possible for  $f(r)$  to vanish at other values of  $r$  if there are multiple horizons in the spacetime. Given any  $\epsilon(r)$ , the two equations (7) and (8) give a classical solution to Einstein's equations with (at least) one horizon at  $r = a$ .

One can easily verify that: (i)  $\epsilon = 0$  implies  $\mu = 0$  and leads to a Schwarzschild spacetime; (ii)  $\epsilon = \epsilon_0 = \text{constant}$  requires  $\mu = \epsilon_0$  and  $T_b^a \propto \text{dia}(\epsilon_0, \mu_0, \mu_0, \mu_0)$  and leads to

$$f(r) = 1 - \frac{A}{r} - Br^2; \quad A = a - \frac{\epsilon_0 a^3}{3}; \quad B = \frac{\epsilon_0}{3}. \tag{9}$$

This represents the Schwarzschild–de Sitter spacetime and when  $a = 0$  it reduces to the de Sitter spacetime; (iii)  $\epsilon = (Q^2/r^4) + \epsilon_0$  gives  $\mu = -(Q^2/r^4) + \epsilon_0$  which corresponds to a Reissner–Nordstrom–de Sitter metric. In fact, the linearity of Einstein's equations (6) allows superposition of different  $\epsilon$  and one can build all these solutions by superposition.

Let us now assume that the solution (7) is such that  $f(r) = 0$  at  $r = a_i, i = 1, 2, \dots$  with  $f'(a_i) = B_i$  finite. Then near  $r = a_i$  the metric can be expanded in a Taylor series with  $f(r) \approx B_i(r - a_i)$ . If there are multiple horizons, such expansion is possible near each zero of  $f$  and I will suppress the subscript  $i$  and just denote by the symbols  $a$  and  $B$  the corresponding values for each horizon. If only one horizon is present, then the Euclidean time coordinate is periodic with a period  $\beta = (4\pi/|B|)$  leading to a *global* notion of temperature. In the context of multi-horizon spacetimes, it is natural to endow each horizon with a locally defined temperature  $T_i = |B_i|/4\pi$ . This can be defined precisely in the context of multi-horizon spacetimes by considering the metric near each horizon and identifying the local Rindler temperature or by developing the quantum field theory. (See appendix A; in the case of a multi-horizon spacetime, there is, of course, no notion of a *global* equilibrium temperature.) From the first of the equations (6) evaluated at  $r = a$ , we get

$$f'(a) = \frac{1}{a} - \epsilon(a)a = B \tag{10}$$

or,

$$\frac{1}{2}Ba - \frac{1}{2} = -\frac{1}{2}\epsilon(a)a^2. \tag{11}$$

It is possible to provide an interesting interpretation of this equation which throws light on the notion of entropy and energy. Multiplying the above equation by  $da$  and using  $\epsilon = 8\pi T_r^r$ ,

it is trivial to rewrite equation (11) in the form

$$\begin{aligned} \frac{B}{4\pi} d\left(\frac{1}{4}4\pi a^2\right) - \frac{1}{2} da &= -T_r^r(a) d\left(\frac{4\pi}{3}a^3\right) \\ &= -T_r^r(a)[4\pi a^2] da. \end{aligned} \quad (12)$$

(It should be remembered that this equation holds separately for each of the horizons with  $a$ ,  $B$ , etc actually standing for  $a_i$ ,  $B_i$ ,  $i = 1, 2, \dots$ .) Let us first consider the case in which a particular horizon has  $f'(a) = B > 0$  so that the temperature is  $T = B/4\pi$ . Since  $f(a) = 0$ ,  $f'(a) > 0$ , it follows that  $f > 0$  for  $r > a$  and  $f < 0$  for  $r < a$ ; that is, the ‘normal region’ in which  $t$  is time-like is outside the horizon as in the case of, for example, the Schwarzschild metric. The first term on the left-hand side of (12) clearly has the form of  $T dS$  since we have an independent identification of temperature from the periodicity argument in the local Rindler coordinates. Since the pressure is  $P = -T_r^r$ , the right-hand side has the structure of  $P dV$  or—more relevantly—is the product of the radial pressure times the transverse area times the radial displacement. This is important because, for the metrics in the form (4), the proper transverse area is just that of a 2-sphere though the proper volumes and coordinate volumes differ. The product of pressure  $P$  times the transverse *proper* area  $4\pi a^2$  gives the correct force; multiplying this force by the (virtual) displacement  $da$  gives the (virtual) work done. Hence the relevant quantity is  $dV = (4\pi a^2) da$  and its integral  $V = (4\pi/3)a^3$ , sometimes called ‘areal volume’, is the relevant volume for our analysis. The second equality in (12) makes this point clear. (This interpretation is discussed further in subsection (2.3) below.) In the case of horizons with  $B = f'(a) > 0$  which we are considering (with  $da > 0$ ), the volume of the region where  $f < 0$  will increase and the volume of the region where  $f > 0$  will decrease. Since the entropy is due to the existence of an inaccessible region,  $dV$  must refer to the change in the volume of the inaccessible region where  $f < 0$ . (We will see below that consistent interpretation is possible for the  $B < 0$  case as well.) We can now identify  $T$  in  $T dS$  and  $P$  in  $P dV$  without any difficulty and interpret the remaining term (second term on the left-hand side) as  $dE = da/2$ . We thus get the expressions for the entropy  $S$  and energy  $E$  (when  $B > 0$ ) to be

$$S = \frac{1}{4}(4\pi a^2) = \frac{1}{4}A_H; \quad E = \frac{1}{2}a = \left(\frac{A_H}{16\pi}\right)^{1/2}. \quad (13)$$

In the case of the Schwarzschild black hole with  $a = 2M$ , the energy turns out to be  $E = a/2 = M$  which is as expected. More generally,  $E = (A_{\text{horizon}}/16\pi)^{1/2}$  corresponds to the so-called irreducible mass in black-hole spacetimes [7]. Of course, the identifications  $S = 4\pi M^2$ ,  $E = M$ ,  $T = 1/8\pi M$  are consistent with the result  $dE = T dS$  in this particular case since the pressure vanishes. As I said before, reproducing this result for black holes is not so significant, because much more formal and general results exist in the case of black hole thermodynamics. What is significant is the fact that our analysis is completely local and did not use any feature regarding asymptotic flatness, etc even to define the energy.

### 2.1. Energy of the de Sitter horizon

This aspect becomes clearer when we study a horizon which is *not* associated with a black hole, namely the de Sitter horizon. There has been considerable interest in this spacetime recently but no clear formulation of ‘laws’ analogous to the laws of black-hole dynamics exist in this context. (Except possibly for one approach, based on the concept of isolated horizons, that attempts to provide such an analysis [8] using a very specific definition for spacetime horizons.) Equation (12), however, can easily provide an interpretation of entropy and energy

in the case of the de Sitter universe. In this case,  $f(r) = (1 - H^2 r^2)$ ,  $a = H^{-1}$ ,  $B = -2H < 0$  so that the temperature—which should be positive—is  $T = |f'(a)|/(4\pi) = (-B)/4\pi$ . For horizons with  $B = f'(a) < 0$  (like the de Sitter horizon) which we are now considering,  $f(a) = 0$ ,  $f'(a) < 0$ , and it follows that  $f > 0$  for  $r < a$  and  $f < 0$  for  $r > a$ ; that is, the ‘normal region’ in which  $t$  is time-like is inside the horizon as in the case of, for example, the de Sitter metric. Multiplying equation (12) by  $(-1)$ , we get

$$\frac{-B}{4\pi} d\left(\frac{1}{4}4\pi a^2\right) + \frac{1}{2} da = T_r^r(a) d\left(\frac{4\pi}{3}a^3\right) = P(-dV). \tag{14}$$

The first term on the left-hand side is again of the form  $T dS$  (with positive temperature and entropy). The term on the right-hand side has the correct sign since the inaccessible region (where  $f < 0$ ) is now outside the horizon and the volume of this region changes by  $(-dV)$ . Once again, we can use (14) to identify the entropy and the energy:

$$S = \frac{1}{4}(4\pi a^2) = \frac{1}{4}A_{\text{horizon}}; \quad E = -\frac{1}{2}H^{-1}. \tag{15}$$

As a byproduct, our approach provides an interpretation of energy for the de Sitter spacetime—an issue which is currently attracting attention—and a consistent thermodynamic interpretation of the de Sitter horizon. Our identification,  $E = -\frac{1}{2}H^{-1}$  is also supported by the following argument: If we use the ‘reasonable’ assumptions  $S = \frac{1}{4}(4\pi H^{-2})$ ,  $V \propto H^{-3}$  and  $E = -PV$  in the equation  $T dS - P dV = dE$  and treat  $E$  as an unknown function of  $H$ , we get the equation

$$H^2 \frac{dE}{dH} = -(3EH + 1) \tag{16}$$

which integrates to give precisely  $E = -\frac{1}{2}H^{-1}$ . Note that I only needed the proportionality,  $V \propto H^{-3}$ , in this argument since  $P dV \propto (dV/V)$ . The ambiguity between the coordinate and proper volume is again irrelevant. This energy is also numerically the same as the total energy within the coordinate Hubble volume of the classical solution, with a cosmological constant:

$$E_{\text{Hub}} = \frac{4\pi}{3}H^{-3}\rho_{\Lambda} = \frac{4\pi}{3}H^{-3}\frac{3H^2}{8\pi} = \frac{1}{2}H^{-1}. \tag{17}$$

(The extra negative sign of  $E = -E_{\text{Hub}}$  is related to a feature noted in the literature in a different context; see, for example, the discussion following equation (71) in the review [9].)

There is an interesting feature in this argument which is worth discussing<sup>1</sup>. If the formal results of black hole dynamics valid for arbitrary matter filled stationary spaces containing a horizon [10] are carried over to the de Sitter spacetime, then one either obtains  $E = +\frac{1}{2}H^{-1}$  or  $E = -H^{-1}$ . The first one arises essentially due to the result in (17) while the second arises if the effective mass density is taken to be  $(\rho + 3P)$ . The result obtained above  $E = -\frac{1}{2}H^{-1}$  differs from both these expressions. It is easy to see that the approach based on Einstein’s equations as well as the argument given above uses  $\rho$  as the relevant energy density and not  $\rho + 3P$  and hence differs from the result  $E = -H^{-1}$ . Settling whether  $E = +\frac{1}{2}H^{-1}$  or  $E = -\frac{1}{2}H^{-1}$  is more subtle for the following reason. In the case of a black hole, increase in the area of horizon  $A_H$  leads to increase in the volume of the inaccessible region. It therefore makes sense to assume that  $(dS/dA_H) > 0$ . In the case of a de Sitter universe, increase in the area of a horizon increases the region accessible to the canonical observer in the region  $r < H^{-1}$  and thus decreases the inaccessible region beyond the horizon. (This notion is not very precise since the volume in the region  $r > H^{-1}$  could be divergent.) From this point of view, it may be acceptable to take  $(dS/dA_H) < 0$ . This will change the sign in  $T dS$

<sup>1</sup> I thank W Israel for discussion on this point.

term in the relevant equations and repeating our analysis, straightforward algebra will lead to  $E = \frac{1}{2}H^{-1}$ . This is most easily seen from the argument given in the last paragraph. If we take  $dS = -\frac{1}{4}dA_H$ , equation (16) changes to

$$H^2 \frac{dE}{dH} = -(3EH - 1) \quad (18)$$

which has the solution  $E = +\frac{1}{2}H^{-1}$ . So, the choice of  $E = \pm\frac{1}{2}H^{-1}$  is related to the choice of  $dS = \mp\frac{1}{4}dA_H$ .

I believe the correct result is indeed  $E = -\frac{1}{2}H^{-1}$  because we expect total entropy of a system to be a well-defined positive quantity. Integrating the equation  $dS = -\frac{1}{4}dA_H$  leads to the result  $S = \frac{1}{4}(A_0 - A_H)$  with an undetermined constant  $A_0$ . This result is hard to interpret since there are no physical quantities available to characterize  $A_0$  in these spacetimes and the choice of  $A_0 = 0$  will now lead to negative entropy. There is nothing wrong in  $dS$  being negative but in a complete thermodynamic description of the spacetime horizon,  $S$  has to be positive and should not depend on any arbitrary constant.

While (12) gives a consistent interpretation, one may wonder about the uniqueness. For example, one could have multiplied the entire equation by an arbitrary function  $F(a)$  which could even differ from horizon to horizon in the case of multi-horizon spacetimes. In that case (taking  $B > 0$ ), the expressions for entropy, energy and volume will become:

$$S = \int 2\pi F(a)a da; \quad E = \frac{1}{2} \int F(a) da; \quad V = \int 4\pi a^2 F(a) da. \quad (19)$$

It seems reasonable to assume that the cross-section area on which the radial pressure should act must be  $4\pi a^2$  since it is the *proper* area in the class of metrics which we are considering. Given this criterion, it follows that  $F(a) = 1$ . There is no freedom—not even that of rescaling—left after this.

## 2.2. Results for (1 + 2) dimensional gravity

The ideas also work in the case of (1+2) dimensional gravity which has attracted a fair amount of attention (see [11]). For the metrics in (4) with  $dL_{\perp}^2 = r^2 d\theta^2$ , Einstein's equations demand that the stress tensor has the form  $8\pi T_b^a = \text{dia}(\epsilon(r), \epsilon(r), \mu(r))$ . Einstein's equations are

$$-\frac{1}{2} \frac{f'}{r} = \epsilon(r); \quad -\frac{1}{2} f''(r) = \mu(r). \quad (20)$$

These are also linear in the source term and can be integrated for a given function  $\epsilon(r)$ . The solution is

$$f(r) = -2 \int_a^r dx x \epsilon(x); \quad \mu(r) = (r\epsilon(r))' \quad (21)$$

where we have again chosen the boundary condition such that  $f(a) = 0$ . The relation  $f' = -2r\epsilon$  evaluated at  $r = a$  gives  $B = -2a\epsilon(a)$ . Multiplying by  $da$  and rearranging terms, this relation can be written in the form

$$\left(\frac{B}{4\pi}\right) d\left(\frac{1}{4}(2\pi a)\right) = (-T_r^r)(2\pi a) da = (-T_r^r) d(\pi a^2). \quad (22)$$

We see that our interpretation carries through in this case as well. Since the temperature is  $T = (B/4\pi)$  when  $B > 0$  and the pressure is  $P = -T_r^r$ , we can immediately identify the  $T dS$  and  $P dV$  terms. (The interpretation of the signs proceed as in the (1 + 3) case and can be provided for the  $B < 0$  situation as well.) The entropy is still one quarter of the 'area',

$S = \frac{1}{4}2\pi a$ , and the energy vanishes identically:  $E = 0$ . The vanishing of energy signifies the fact that at the level of the metric, Einstein's equations are vacuous in (1 + 2) and we have not incorporated any topological effects (such as deficit angles corresponding to point masses in (1 + 2) dimensions) in this approach. We will see in section 3 that the same result can be obtained from the quantum mechanical evaluation of the partition function.

It would have been nice if the analysis could also be extended to metrics of the form in (4) with  $dL_{\perp}^2 = dy^2 + dz^2$  and  $r$  interpreted in the range  $-\infty < r < +\infty$  which will include the Rindler spacetime when  $f(r) = (1 + 2gr)$ . This, however, cannot be done because both the  $T dS$  and  $P dV$  terms—which are proportional to the transverse area—will diverge in this case, making Einstein's equations vacuous. As we shall see in section 3, one can get around this difficulty in the quantum formulation which does not rely on Einstein's equations.

### 2.3. Interpretation and comments

How does one interpret the differential form of the equation (12) which arises from Einstein's equations written in the form (11)? To begin with, we note that this equation is purely local and arises from the classical Einstein's equation connecting  $f'$  and  $\epsilon$  when  $f$  vanishes. (Though I used the trick of multiplying by  $da$  to get (12) one could have got the same result by taking the differential of, say, (11) and using Einstein's equations again.) Therefore any interpretation should be purely local. Secondly, the differential form of the equation shows how the solution changes if the location of the horizon is displaced radially. Endowing the horizon with physical characteristics such as entropy, temperature and energy, one can interpret the left-hand side of (12) as the change in these parameters. The right-hand side represents the work done by or against the pressure depending on the sign. This relation is independent of how the horizon radius is changed. The standard discussions of the first law of black hole dynamics, in which a test particle of some mass is made to fall into a black hole, say, is just one way of effecting the change in the horizon size; black hole evaporation could provide another process leading to such a change. The formalism does not care what causes the change of the horizon radius. (For a different approach in interpreting Einstein's equations as a thermodynamic identity see [5, 12].)

The same result can be stated more formally along these lines: In standard thermodynamics, we can consider two equilibrium states of a system differing infinitesimally in the extensive variables volume, energy and entropy by  $dV$ ,  $dE$  and  $dS$  while having the *same* values for the intensive variables temperature ( $T$ ) and pressure ( $P$ ). Then, the first law of thermodynamics asserts that  $T dS = P dV + dE$  for these states. In a similar vein, we can consider two spherically symmetric solutions to Einstein's equations with the radius of the horizon differing by  $da$  while having the same source  $T_{ik}$  and the same value for  $B$ . Then the entropy and energy will be infinitesimally different for these two spacetimes; but the fact that both spacetimes satisfy Einstein's equations shows that  $T dS$  and  $dE$  will be related to the external source  $T_{ik}$  and  $da$  by equation (12). Just as in standard thermodynamics, this relation could be interpreted as connecting a sequence of quasi-static equilibrium states. The mathematical description does not distinguish between the 'active' interpretation in which the pressure is considered to have done some work against expansion or the 'passive' interpretation in which one is comparing the changes in  $S$ ,  $V$ ,  $E$  of two infinitesimally different states. (It should be stressed that, in any thermodynamic description of horizons,  $P$  does not refer to the pressure of the radiation emitted by the horizon. This is clear from the fact that, in the case of Schwarzschild spacetime with  $T_{ik} = 0$ , we have  $T dS = dE$  though the Hawking radiation will indeed have some pressure.)

In fact, the virtue of the formalism is *not* in handling black hole spacetimes with a single horizon—for which much more sophisticated, formal and generally covariant approaches and results are available [1]—but in its ability to handle spacetimes like the de Sitter universe, Schwarzschild–de Sitter universe, etc which have horizons *not* associated with black holes. Compared to black hole horizons, we have much less understanding of the dynamics of other horizons. (For example, there is no simple analogue of the four laws of black-hole dynamics in the de Sitter spacetime or the Schwarzschild–de Sitter spacetime.) In the local interpretation advocated here, I will consider  $E$  to be the energy of the horizon rather than the energy of the spacetime geometry defined asymptotically, etc. Thus  $E_i$ ,  $S_i$  and  $T_i$  are defined and attributed to each horizon and  $P_i dV_i = P_i(4\pi a_i^2) da_i$  is the work done due to the pressure acting on the transverse proper area of the horizon. The entire description is local and allows the formalism to be extended to spacetimes with any number of horizons. Virtual displacements of each of the horizons in spherically symmetric spacetimes have to satisfy the relationship (12). Each horizon comes with its own temperature, entropy and energy.

The analysis is classical except for the crucial periodicity argument which is used to identify the temperature uniquely. This is again done locally by approximating the metric by a Rindler metric close to the horizon and identifying the Rindler temperature. (This idea bypasses the difficulties in defining and normalizing Killing vectors in spacetimes which are not asymptotically flat). Without this quantum mechanical ingredient one will not be able to fix the constant of proportionality between surface gravity and temperature. In normal units, equation (12) should read

$$\frac{c^4}{G} \frac{B}{4\pi} d\left(\frac{1}{4}4\pi a^2\right) - \frac{1}{2} \frac{c^4}{G} da = -T_r(a)[4\pi a^2] da. \quad (23)$$

The second term on the left-hand side and the right-hand side have the correct dimensions for each factor. In the first term on the left-hand side, multiplying and dividing by  $\hbar$  allows us to write

$$\frac{c^4}{G} \frac{B}{4\pi} d\left(\frac{1}{4}4\pi a^2\right) = \frac{c^3}{G\hbar} \frac{B\hbar c}{4\pi} d\left(\frac{1}{4}4\pi a^2\right) = T d\left(\frac{A}{4L_p^2}\right) \quad (24)$$

with  $T = (B\hbar c/4\pi)$  being the temperature in energy units. It is obvious that quantum mechanics enters only through the factor  $\hbar$  which appears in expressing the temperature in terms of the surface gravity. Of course, one needs  $\hbar$  along with  $G$  and  $c$  to obtain a quantity with dimensions of area.

One can extend much of this analysis to a slightly more general case in which  $g_{00} = f_1(r)$  and  $g_{11} = -1/f_2(r)$  with  $f_1 \neq f_2$  in a fairly straightforward manner. Using the relevant Einstein's equations for this case, simple algebra shows that: (i) at the horizon, both  $f_1$  and  $f_2$  will vanish, (ii)  $T_t^t = T_r^r$  at the horizon, though in general they are not equal and (iii) equation (11) is satisfied. To prove (i) let us assume that  $f_2(r) = 0$  at some  $r = a$  for the first time as we approach from the right and  $f_1(r) > 0$  at  $r > a$ . If  $f_1(r) > 0$  at  $r = a$ , then the  $t$  direction will remain time-like and  $(\theta, \phi)$  directions will be space-like at  $r = a$ . But since the horizon at  $r = a$  must be a null surface, it must have a null tangent direction and symmetry requires this to be the  $t$  direction. Hence  $f_1$  must vanish at  $r = a$ . Using the behaviour of Einstein's equations it is also possible to prove the converse: that is, if  $f_1(r) = 0$  at  $r = a$ , then  $f_2(r) = 0$  at  $r = a$ . The result (ii) can be proved by expanding  $f_1(r)$ ,  $f_2(r)$  around  $r = a$  with derivatives  $B_1, B_2$  at  $r = a$  and substituting into the expression for the scalar curvature. The scalar curvature will be finite at  $r = a$  only if  $T_t^t = T_r^r$  at the horizon. Using (ii) in the Einstein's equations will lead to (iii). Thus the only non-trivial assumption used in the above analysis is that of spherical symmetry.

2.4. Dimensionality of horizons

Incidentally, this analysis throws light on another issue, namely the effective dimension of the horizon. Bekenstein and Mayo have argued [13] that the horizon of the Schwarzschild black hole is effectively one dimensional. The argument relies on the fact that the rate of flow,  $\dot{S}$ , of entropy or information in a channel with power  $\mathcal{P}$  scales as  $\dot{S} \propto \mathcal{P}^{1/2}$  for a one-dimensional channel, independent of other details. For higher dimensional channels it is not possible to obtain the unique index of  $\frac{1}{2}$  in the above relation. In the case of blackbody radiation emerging from a surface of area  $A$  and temperature  $T$ , we have  $\dot{S} \propto AT^3$  while  $\mathcal{P} \propto AT^4$ . In general, there is no relation between  $A$  and  $T$  and one cannot relate  $\dot{S}$  to  $\mathcal{P}$ . In the case of spherically symmetric spacetimes with horizons considered above,  $A \propto a^2$  and  $T^3 \propto |f'(a)|^3 \propto B^3$ . Taking  $\dot{S} \propto B^3 a^2$ ,  $\mathcal{P} \propto B^4 a^2$  and using (11) it follows that

$$\frac{\dot{S}}{\sqrt{\mathcal{P}}} \propto |B|a = |1 - a^2\epsilon(a)|. \tag{25}$$

Therefore Bekenstein’s result can be generalized to any spacetime in which  $\epsilon(a)$  vanishes or is proportional to  $a^{-2}$ . The de Sitter universe is an example of the latter case for which  $\epsilon(a)a^2 = 3$  allowing us to conclude that the pure de Sitter horizon is also of effectively one dimension. Equation (25) shows that the result does *not* extend to other cases such as, for example, the Schwarzschild–de Sitter spacetime or even the Reissner–Nordstrom black hole. The reason for this failure is related to a feature which was mentioned earlier. Spacetimes such as Schwarzschild or de Sitter have only one free parameter in the metric (such as  $M$  or  $H^{-1}$ ) and hence the scaling of all other thermodynamic parameters is uniquely fixed by purely dimensional considerations. This is the reason why combinations such as  $Ba$  are pure numbers in these cases. But, in general, the area of the horizon (and hence the entropy) is determined by the location of the zero of the function  $f(r)$  while the temperature—determined by periodicity considerations—is determined by the value of  $f'(r)$  at the zero. For a general function, of course, there will be no relation between the location of the zero and the slope of the function at that point. Thus, we cannot expect  $\dot{S}/\sqrt{\mathcal{P}}$  to have a simple form in the general case.

3. Partition function in quantum theory

The analysis given above provided expressions for  $S$  and  $E$  using Einstein’s equations. I will now show that the same result can be obtained in quantum theory *without using Einstein’s equations*, for a slightly wider class of metrics [14].

A wider class of spacetimes, analysed in the literature, has the form in (4) where  $f(r)$  vanishes at some surface  $r = a$ , say, with  $f'(a) \equiv B$  remaining finite and  $dL_{\perp}^2$  interpreted more generally. When  $dL_{\perp}^2$  is taken as the metric on a 2-sphere and  $r$  is interpreted as the radial coordinate [ $0 \leq r \leq \infty$ ], equation (4) covers a variety of spherically symmetric spacetimes (including Schwarzschild, Reissner–Nordstrom, de Sitter, etc) with a compact horizon at  $r = a$ . (This is the case discussed earlier. I do not assume that there is only one horizon; symbols such as  $a, B \dots$ , etc refer to different horizons if more than one horizon is present and should be thought of as  $B_i, a_i, \dots$ , with  $i = 1, 2, \dots$ . For simplicity of notation, I suppress the subscript  $i$  in what follows.) If  $r$  is interpreted as one of the Cartesian coordinates  $x$  with ( $-\infty \leq x \leq \infty$ ) and  $dL_{\perp}^2 = dy^2 + dz^2$ ,  $f(x) = 1 + 2gx$ , equation (4) can describe the Rindler frame in flat spacetimes. Finally  $0 \leq r \leq \infty$  with  $dL_{\perp}^2 = r^2 d\theta^2$  will describe circularly symmetric (2 + 1) dimensional spacetimes. We shall first concentrate on compact horizons in (3 + 1) with  $r$  interpreted as a radial coordinate, and comment on the other cases at the end.

Since the metric is static, Euclidean continuation is trivially effected by  $t \rightarrow \tau = it$  and an examination of the conical singularity near  $r = a$  (where  $f(r) \approx B(r - a)$ ) shows that  $\tau$  should be interpreted as periodic with period  $\beta = 4\pi/|B|$  corresponding to the temperature  $T = |B|/4\pi$ . Let us consider any one of the horizons with the temperature  $T = |B|/4\pi$ . The class of metrics in (4) with the behaviour [ $f(a) = 0, f'(a) = B$ ] constitutes a canonical ensemble at constant temperature since they all have the same temperature  $T = |B|/4\pi$ . The partition function for this ensemble is given by the path integral sum

$$\begin{aligned} Z(\beta) &= \sum_{g \in \mathcal{S}} \exp(-A_E(g)) \\ &= \sum_{g \in \mathcal{S}} \exp\left(-\frac{1}{16\pi} \int_0^\beta d\tau \int d^3x \sqrt{g_E} R_E[f(r)]\right) \end{aligned} \quad (26)$$

where I have made the Euclidean continuation of the Einstein action and imposed the periodicity in  $\tau$  with period  $\beta = 4\pi/|B|$ . (For static spacetimes, there is some ambiguity regarding the overall sign when the Euclidean continuation is performed. The convention adopted here is as follows: first note that a positive definite Euclidean metric requires one to work with a signature  $(-+++)$  while I am using  $(+---)$ . Changing the signature in a diagonal metric is equivalent to changing the signs of components of a metric tensor:  $g_{ik} \rightarrow -g_{ik}$ . This leads to  $\Gamma_{i,kl} \rightarrow -\Gamma_{i,kl}; \Gamma_{kl}^i \rightarrow \Gamma_{kl}^i; R^i_{jkl} \rightarrow R^i_{jkl}; R_{jl} \rightarrow R_{jl}; R \rightarrow -R$ . The Euclidean continuation is effected through  $it = \tau$  in a spacetime with signature  $(-+++)$ .) The sum in (27) is restricted to the set  $\mathcal{S}$  of all metrics of the form in (4) with the behaviour [ $f(a) = 0, f'(a) = B$ ] and the Euclidean Lagrangian is a functional of  $f(r)$ . No source term or cosmological constant (which cannot be distinguished from a certain form of source) is included since the idea is to obtain a result which depends purely on the geometry. The spatial integration will be restricted to a region bounded by the 2-spheres  $r = a$  and  $r = b$ , where the choice of  $b$  is arbitrary except for the requirement that within the region of integration the Lorentzian metric must have the proper signature with  $t$  being a time coordinate. Using the result

$$R = \frac{1}{r^2} \frac{d}{dr}(r^2 f') - \frac{2}{r^2} \frac{d}{dr}[r(1-f)] \quad (27)$$

valid for metrics of the form in (4) [with signature  $(+---)$ ], a straightforward calculation shows that

$$\begin{aligned} -A_E &= \frac{\beta}{4} \int_a^b dr [-[r^2 f']' + 2[r(1-f)]'] \\ &= \frac{\beta}{4} [a^2 B - 2a] + Q[f(b), f'(b)] \end{aligned} \quad (28)$$

where  $Q$  depends on the behaviour of the metric near  $r = b$  and we have used the conditions [ $f(a) = 0, f'(a) = B$ ]. The sum in (27) now reduces to summing over the values of [ $f(b), f'(b)$ ] with a suitable (but unknown) measure. This sum, however, will only lead to a factor which we can ignore in deciding about the dependence of  $Z(\beta)$  on the form of the metric near  $r = a$ . Using  $\beta = 4\pi/B$  (and taking  $B > 0$ , for the moment) the final result can be written in a very suggestive form:

$$Z(\beta) = Z_0 \exp\left[\frac{1}{4}(4\pi a^2) - \beta \left(\frac{a}{2}\right)\right] \propto \exp[S(a) - \beta E(a)] \quad (29)$$

with the identifications for the entropy and energy being given by:

$$S = \frac{1}{4}(4\pi a^2) = \frac{1}{4} A_{\text{horizon}}; \quad E = \frac{1}{2} a = \left(\frac{A_{\text{horizon}}}{16\pi}\right)^{1/2}. \quad (30)$$

This is exactly the result obtained earlier except that we have *not* used Einstein’s equations. Instead, the partition function was evaluated with two very natural conditions:  $f(a) = 0$  making the surface  $r = a$  a compact horizon and  $f'(a) = \text{constant}$  which is the proper characterization of the canonical ensemble of spacetime metrics. Since temperature is well defined for the class of metrics which I have considered, this canonical ensemble is defined without any ambiguity. This allows me to sum over a class of spherically symmetric spacetimes at one go rather than deal with, say, black-hole spacetimes and de Sitter spacetime separately.

The idea of using a canonical ensemble is not new and one of the earliest attempts is by [15]; similar ideas have been explored by different people [16] mostly in the context of BH spacetime. What is new are the following features: (a) I have not added any boundary terms to the action and in fact  $R\sqrt{-g}$  in (27) is a total derivative. A detailed description of action functionals is given in appendix B showing that the same result is obtained from the surface terms, if done correctly. (b) I have not invoked the WKB limit to evaluate the action using a solution to Einstein’s equations. Normally, the path integral in (27) is formidable to calculate and is ill-defined without the boundary terms. I can manage with (a) and (b) above only because I confine the sum to metrics in the set  $\mathcal{S}$ . Conceptually, a canonical ensemble for a minisuperspace of metrics of the form in (4) should be constructed by keeping the temperature constant *without* assuming the metrics to be the solutions of Einstein’s equation; this is what I do and exploit the form of  $R$  given by (27). Since this action involves second derivatives, it is not only allowed but even required to fix both  $f$  and  $f'$  at the boundaries. Moreover, since I am summing over an arbitrary class of  $f(r)$ , the behaviour of the function  $f(r)$  at  $r = b$  is not in any way related to the parameter  $a$  or parameter  $B$  which enters the analysis at the horizon. In the conventional analysis  $f$  is chosen to be a classical solution and hence its value at any radius is predetermined in terms of the parameters of the solution (such as  $M$  in the Schwarzschild metric).

The temperature is determined by using the Euclidean periodicity condition near the horizon and is related to the derivative of  $f$  by  $T = |f'(a)|/4\pi$ . In the literature one sometimes uses the locally defined Tolman temperature  $T_{\text{local}}(x) = T_R(g_{00}(x_R)/g_{00}(x))^{1/2}$  where  $x_R$  is some reference point. This local temperature will diverge at the horizon if  $T_R$  is defined as a finite quantity elsewhere. In the approach taken here, all these issues are bypassed by directly defining the temperature by expanding  $f(r)$  in a Taylor series near the horizon and using the locally defined Rindler temperature. It is important that one uses this quantity because there is no well-defined notion of temperature in spacetimes with multiple horizons (see appendix A).

We get the same result as obtained earlier in the case of horizons with  $B < 0$  such as the de Sitter universe as well. In this case,  $f(r) = (1 - H^2r^2)$ ,  $a = H^{-1}$ ,  $B = -2H$ . Since the region where  $t$  is time-like is ‘inside’ the horizon, the integral for  $A_E$  in (28) should be taken from some arbitrary value  $r = b$  (which could even be the origin) to the horizon at  $r = a$  with  $a > b$ . So the horizon contributes in the upper limit of the integral introducing a change of sign in (28). Further, since  $B < 0$ , there is another negative sign in the area term from  $\beta B \propto B/|B|$ . Taking all these into account we get, in this case,

$$Z(\beta) = Z_0 \exp \left[ \frac{1}{4}(4\pi a^2) + \beta \left( \frac{a}{2} \right) \right] \propto \exp [S(a) - \beta E(a)] \tag{31}$$

giving  $S = \frac{1}{4}(4\pi a^2) = \frac{1}{4}A_{\text{horizon}}$  and  $E = -\frac{1}{2}H^{-1}$  as before.

Let us next consider the  $(1 + 2)$  dimensional spacetime. In  $D = (1 + 2)$ , metrics of the type in (4) with  $dL_{\perp}^2 = r^2 d\theta^2$  will have the scalar curvature

$$R = \frac{1}{r} \frac{d}{dr}(rf') + \frac{f'}{r}. \tag{32}$$

The action will now become

$$-A_E = \frac{B}{16\pi} \int_a^b 2\pi dr [-(rf')' - f'] = \frac{1}{4}(2\pi a) + Q[f'(b)] \quad (33)$$

leading to  $Z = Z_0 \exp\left[\frac{1}{4}A_{2D}\right]$  where the ‘area’ of the 2D horizon is  $2\pi a$ . This will give  $S = \frac{1}{4}(2\pi a) = \frac{1}{4}A_{\text{horizon}}$  with  $E = 0$ . The vanishing of the energy signifies the fact that, at the level of the metric, Einstein’s equations are vacuum in (1 + 2) and we have not incorporated any topological effects (such as deficit angles corresponding to point masses in (1 + 2) dimensions) in our approach.

Interestingly enough, this formalism can also handle spacetimes such as Rindler unlike the approach discussed in the last section which could only work for compact transverse areas. For the spacetimes with planar symmetry for which (4) is still applicable with  $r = x$  being a Cartesian coordinate and  $dL_{\perp}^2 = dy^2 + dz^2$ . In this case  $R = f''(x)$  and the action becomes

$$\begin{aligned} -A_E &= \frac{1}{16\pi} \int_0^{\beta} d\tau \int dy dz \int_b^a dx f''(x) \\ &= \frac{\beta}{16\pi} A_{\perp} f'(a) + Q[f'(b)] \end{aligned} \quad (34)$$

where we have confined the transverse integrations to a surface of area  $A_{\perp}$ . If we now sum over all the metrics with  $f(a) = 0$ ,  $f'(a) = B$  and  $f'(b)$  arbitrary, the partition function will become

$$Z(\beta) = Z_0 \exp\left(\frac{1}{4}A_{\perp}\right) \quad (35)$$

which shows that planar horizons have an entropy of  $\frac{1}{4}$  per unit transverse area but zero energy. This includes the Rindler frame as a special case. Note that if we freeze  $f$  to its Rindler form  $f = 1 + 2gx$  (for instance, by demanding the validity of Einstein’s equations in the WKB approach), then  $R = f'' = 0$  as it should. In the action in (34),  $f'(a) - f'(b)$  will give zero. It is only because I am *not* doing a WKB analysis—but varying  $f'(b)$  with fixed  $f'(a)$ —that I obtain an entropy for these spacetimes.

It is interesting that this approach, which does not use the classical Einstein’s equations, gives the same result as that obtained earlier using Einstein’s equations. One difference is that the partition function allows us to identify  $S$  and  $E$  without recourse to  $P$  or  $V$  while in the earlier approach we essentially rewrote Einstein’s equations as the first law of thermodynamics (which included the  $P dV$  term) by multiplying by  $da$ . The fact that the approach based on the action functional leads to the thermodynamic interpretation directly (without the use of Einstein’s equations) suggests that there could be a deep connection between the form of the Einstein action and the thermodynamic interpretation of spacetimes with horizons. This is indeed true and some aspects of this connection are explored in detail elsewhere [5, 14]. In particular, it was shown in [5] that the Einstein–Hilbert action can be expressed as the ‘free energy of spacetime’ and Einstein’s equations may be interpreted as describing the *thermodynamics of spacetime* while quantum gravity is needed to describe the statistical mechanics of the microscopic degrees of freedom making up the spacetime.

Both the classical and quantum descriptions developed here used spherical symmetry; in the first approach, we used Einstein’s equations for the spherically symmetric spacetime while in the second approach we used the form of  $R$  valid for spherically symmetric spacetimes. One may wonder whether this assumption is essential or not. Given the broader interpretation of Einstein–Hilbert action as a free energy of spacetime, it seems likely that one will be able to generalize the results of this paper to any horizon without having to assume spherical or planar symmetry. This issue is under investigation.

## Acknowledgments

I thank Apoorva Patel, S Shankaranarayanan and Suneeta Varadarajan for useful comments.

## Appendix A. Spacetimes with multiple horizons

Metrics of the form in (4) with  $f(r)$  having a simple zero at one or more points,  $r = a_i, i = 1, 2, 3, \dots$ , exhibit coordinate singularities at  $r = a_i$ . The coordinate  $t$  alternates between being time-like and space-like when each of these horizons are crossed. Since all curvature invariants are well behaved at the horizons, it will be possible to introduce coordinate patches such that the metric is also well behaved at the horizon. There is a general procedure for doing this in the case of spacetimes with a single horizon, which can be used to handle multi-horizon spacetimes as well.

Concentrating on the  $(t, r)$  plane, we first introduce the tortoise coordinate  $\xi$  so that the metric becomes

$$ds^2 = f(r) dt^2 - \frac{dr^2}{f(r)} = f(\xi)(dt^2 - d\xi^2); \quad \xi = \int \frac{dr}{f(r)}. \quad (\text{A.1})$$

Switching over to the light cone coordinates  $(u, v)$  by

$$u = t - \xi; \quad v = t + \xi \quad (\text{A.2})$$

the metric becomes

$$ds^2 = f(\xi)(dt^2 - d\xi^2) = f\left[\frac{1}{2}(v - u)\right] du dv. \quad (\text{A.3})$$

The coordinate systems  $(t, \xi), (u, v)$  are still singular near the horizons  $r \approx a_i$ . Assuming  $f(r) \approx B_i(r - a_i)$  near  $r = a_i$  (with  $B_i = f'(a_i)$ ), it is easy to see that near the horizon the line element has the form

$$ds^2 \approx B_i a_i e^{(B_i/2)(v-u)} dv du \quad (\text{A.4})$$

which is badly behaved because  $(v - u) \propto \xi$  has a logarithmic singularity near  $r = a_i$ . However, the form of the metric in (A.4) suggests a natural coordinate transformation from  $(u, v)$  to a non-singular coordinate system  $(U, V)$  with

$$V_i = (2/B_i) \exp[(B_i/2)v]; \quad U_i = -(2/B_i) \exp[(-B_i/2)u] \quad (\text{A.5})$$

in terms of which the metric becomes  $ds^2 = -(4f/B_i^2 U_i V_i) dU_i dV_i$ . By construction, the  $(U_i, V_i)$  coordinate system is regular on the horizon at  $r = a_i$ , since the combination  $(f/U_i V_i)$  is finite on the horizon. The transformation (A.5) will lead to the familiar Kruskal-type coordinate system in all cases with a single horizon such as the Schwarzschild, Reissner–Nordstrom or de Sitter manifold. In these cases  $i = 1$  and we need only one coordinate patch with  $(U, V)$  to cover the manifold.

When there is more than one horizon, we need to introduce one Kruskal-like coordinate patch for each of the horizons; the  $(u, v)$  coordinate system is unique in the manifold but the  $(U_i, V_i)$  coordinate systems are different for each of the horizons since the transformation in (A.5) depends explicitly on  $B_i$ 's which are (in general) different for each of the horizons. In such a case, there will be regions of the manifold in which more than one Kruskal-like patch can be introduced. The compatibility between these coordinates can provide interesting constraints.

It is easy to develop the quantum field theory in the  $t-r$  plane if we treat it as a  $(1 + 1)$  dimensional spacetime. In this case, the solutions to the wave equations are just arbitrary functions of the null coordinates in a conformally flat coordinate system. Since  $(u, v)$  as well

as all the  $(U_i, V_i)$  coordinates retain the conformally flat nature of the  $(1 + 1)$  dimension, we can define suitable mode functions and vacuum state in a straightforward manner. The outgoing and ingoing modes of the kind  $(4\pi\omega)^{-1/2}[\exp(-i\omega u), \exp(-i\omega v)]$  define a static vacuum state. This is called the Boulware vacuum in the case of the Schwarzschild black hole, but it can be defined in a general spacetime of the kind we are studying (see [4]). The modes of the kind  $(4\pi\omega)^{-1/2}[\exp(-i\omega U_i), \exp(-i\omega V_i)]$  define another vacuum state called the Hartle–Hawking vacuum in the case of a Schwarzschild black hole. (The modes of the kind  $(4\pi\omega)^{-1/2}[\exp(-i\omega U), \exp(-i\omega V)]$  define the analogue of the Unruh vacuum which, however, we will not need.) The transformations between the Boulware and Hartle–Hawking modes are identical to those in the case of the Schwarzschild black hole with the surface gravity at the corresponding horizon occurring in the transformations. The study of either Bogoliubov coefficients or the stress-tensor expectation values will now show that the Hartle–Hawking vacuum has an equilibrium temperature  $T_i = |B_i|/4\pi$ . The Bogoliubov coefficients between two different Hartle–Hawking modes, of course, have no simple interpretation, showing that there is a conceptual problem in defining a unique temperature in the overlapping region.

This difficulty is apparent even from an examination of the coordinate transformations. Consider, for example, the region between two consecutive horizons  $r_n < r < r_{n+1}$  in which  $t$  is time-like. The coordinates  $(U_i, V_i)$  with  $i = n, n + 1$  overlaps in this region. Euclideanization of the metric can be easily effected in the region  $r_n < r < r_{n+1}$  by taking  $\tau = it$ . This will lead to the transformations

$$\begin{aligned} U_{n+1} &= -U_n \exp[(B_{n+1} + B_n)((-i\tau - \xi)/2)]; \\ V_{n+1} &= -V_n \exp[-(B_{n+1} + B_n)((-i\tau + \xi)/2)]. \end{aligned} \quad (\text{A.6})$$

Obviously, single valuedness can be maintained only if the period of  $\tau$  is an integer multiple of  $4\pi/(B_{n+1} + B_n)$ . More importantly, we get from (A.5) the relation

$$U_i + V_i = \frac{4}{B_i} \exp\left(\frac{B_i \xi}{2}\right) \sinh\left(-i \frac{B_i \tau}{2}\right) \quad (\text{A.7})$$

which shows that  $(U_i, V_i)$  can be used to define values of  $\tau$  only up to integer multiples of  $4\pi/B_i$  in each patch. But since  $(U_n, V_n)$  and  $(U_{n+1}, V_{n+1})$  are to be well-defined coordinates in the overlap, the periodicity  $\tau \rightarrow \tau + \beta$  which leaves both the sets  $(U_n, V_n)$  and  $(U_{n+1}, V_{n+1})$  invariant must be such that  $\beta$  is an integer multiple of both  $4\pi/B_n$  and  $4\pi/B_{n+1}$ . This will require  $\beta = 4\pi n_i/B_i$  for all  $i$  with  $n_i$  being a set of integers. This, in turn, implies that  $B_i/B_j = n_i/n_j$  making the ratio between any two surface gravities a rational number. Since there is no physical basis for such a condition, it seems reasonable to conclude that these difficulties arise because of our demanding the existence of a finite periodicity  $\beta$  in the Euclidean time coordinate. This demand is related to an expectation of thermal equilibrium which is violated in spacetimes with multiple horizons having different temperatures. Hence, such spacetimes will not have a global notion of temperature, entropy, etc. It seems, however, unlikely that we cannot attribute a temperature to a black hole formed in some region of the universe just because the universe at the largest scales is described by a de Sitter spacetime. One is led to searching for a local description of the dynamics of all horizons.

## Appendix B. The choice of action functional

Action functionals in classical theory are essentially tools to obtain equations of motion and provide a book keeping of the symmetries of the theory. The numerical value of the action does not play a significant role in classical theory. The situation is different in quantum theory

in which the actual value of the action may be of relevance in evaluating a path integral or in interpreting a partition function.

In the case of general relativity, the generally covariant action is the Einstein–Hilbert action  $A_{EH}$  given by

$$A_{EH} \equiv \frac{1}{16\pi} \int R \sqrt{-g} \, d^4x. \tag{B.1}$$

Since  $R\sqrt{-g}$  is linear in second derivatives of the metric tensor, it is possible to write this action in the form

$$\begin{aligned} A_{EH} &= \frac{1}{16\pi} \int L_{SD} \sqrt{-g} \, d^4x - \frac{1}{16\pi} \int \partial_c P^c \, d^4x \\ &\equiv A_{SD} - \frac{1}{16\pi} \int \partial_c P^c \, d^4x \end{aligned} \tag{B.2}$$

where

$$L_{SD} = g^{ab} \left( \Gamma_{ja}^i \Gamma_{ib}^j - \Gamma_{ab}^i \Gamma_{ij}^j \right) \tag{B.3}$$

$$P^c = \sqrt{-g} \left( g^{ck} \Gamma_{km}^m - g^{ik} \Gamma_{ik}^c \right) = 2g^{cb} \partial_b \sqrt{-g} + \sqrt{-g} \partial_b g^{bc}. \tag{B.4}$$

Equation (B.2) defines the Schrödinger–Dirac action  $A_{SD}$  which is quadratic in the Christoffel symbols and does not contain second derivatives of the metric tensor.

It is possible to write the second term in (B.2) in a different manner which is often used in the literature. If we foliate the spacetime by a series of spacelike hyper-surfaces  $\mathcal{S}$  with  $n^i$  as normal, then  $g^{ik} = h^{ik} + n^i n^k$  where  $h^{ik}$  is the induced metric on  $\mathcal{S}$ . It is conventional to define a quantity called extrinsic curvature  $K_{ab}$  such that

$$K_{ab} = \nabla_a n_b + n_a a_b \tag{B.5}$$

where  $a^b \equiv n^c \nabla_c n^b$  is the ‘acceleration’ corresponding to trajectories with ‘4-velocity’  $n^a$ . (This makes sense on the  $x^0 = \text{constant}$  space-like surfaces since the unit normal  $n^a$  to such surfaces can be thought of as the four velocities tangential to  $\mathbf{x} = \text{constant}$  curves. On the surfaces such as  $x^1 = \text{constant}$ , say, the normal will not be time-like and  $a^b$  should be just treated as a vector normal to  $n^b$  and tangential to the surface element. It does not correspond to a physical acceleration.) In all the surfaces,  $n^b a_b = 0$ , allowing us to write  $K = \nabla_a n^a$ . Given the covariant derivative  $\nabla_i n_j$  of the normals to  $\mathcal{S}$ , one can construct only three vectors ( $n^j \nabla_j n^i, n^j \nabla^i n_j, n^i \nabla^j n_j$ ) which are linear in covariant derivative operator. The first one is the acceleration  $a^i = n^j \nabla_j n^i$ ; the second identically vanishes since  $n^j$  has unit norm; the third,  $u^i K$ , is proportional to the trace of the extrinsic curvature  $K = \nabla^j n_j$  of  $\mathcal{S}$ . Hence, the second term in (B.2)—which is linear in the derivatives of the Christoffel symbols—must be the 4-divergence of a linear combination of  $u^i K$  and  $a^i$ . So the corresponding term in the action must have the form

$$A_{\text{surface}} = \int d^4x \sqrt{-g} \nabla_i [\lambda_1 K u^i + \lambda_2 a^i] \tag{B.6}$$

where  $\lambda_1$  and  $\lambda_2$  are numerical constants. This final result—though probably not the reasoning given above—is known in the conventional (3 + 1) formalism (see, e.g., equation (21.88) of [7]). Since  $u^i a_i = 0$ , the spacelike boundaries at  $x^0 = \text{constant}$  get a contribution only from  $K$  while the time-like surfaces such as  $x^1 = \text{constant}$  get a contribution from the normal component of the acceleration  $a^i \hat{n}_i$  where  $\hat{n}_i$  is the normal to the time-like surface. The numerical constants can be easily evaluated by choosing some simple metric, thereby fixing the surface term.

Such a detailed analysis shows that the surface integral can actually be written as

$$\frac{1}{16\pi} \int_{\mathcal{V}} \partial_c P^c d^4x = \sum \frac{1}{8\pi} \int_{\partial\mathcal{V}} d^3x \sqrt{h} K. \quad (\text{B.7})$$

It is therefore possible to define the  $A_{\text{SD}}$  by either of the relations

$$\begin{aligned} A_{\text{SD}} &= A_{\text{EH}} + \frac{1}{16\pi} \int \partial_c P^c d^4x \\ &\equiv A_{\text{EH}} + \sum \frac{1}{8\pi} \int_{\partial\mathcal{V}} d^3x \sqrt{h} K. \end{aligned} \quad (\text{B.8})$$

The following points need to be noted regarding this result. (a) From the origin of this equation we know that neither the left-hand side nor the second term on the right-hand side is generally covariant. (b) The action  $A_{\text{EH}}$  vanishes in flat spacetime. The action  $A_{\text{SD}}$  does not, in general, vanish in flat spacetime if non-Cartesian coordinates are used. (c) The numerical values of all these actions can be formally divergent if the integration domain is unbounded either in space or time. In the case of  $A_{\text{SD}}$  this can happen even in flat spacetime if curvilinear coordinates are used.

The action  $A_{\text{EH}}$  vanishes for any vacuum solution of Einstein's equation and in particular for the Schwarzschild metric. If one insists on using the WKB value of action to interpret spacetime thermodynamics, then  $A_{\text{EH}}$  will not help even in the simplest case of a Schwarzschild black hole. Motivated by this, most of the previous workers in this field [16] have used  $A_{\text{SD}}$  in order to obtain a non-zero value for the action for the Schwarzschild black hole. In fact, it is often done by using the second expression (involving the surface integral). While this is done fairly routinely in the literature, it is easy to go wrong unless some caution is exercised, especially in deciding which are the surfaces over which one must do the summation.

To illustrate this, I shall consider evaluation of the action for the class of metrics studied in this paper—which have been repeatedly analysed in the literature, though wrongly at times. The line element is

$$ds^2 = f(r) dt^2 - f(r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2). \quad (\text{B.9})$$

I choose the four-dimensional region  $\mathcal{V}$  to be bounded by two spacelike surfaces ( $S_0, S_\beta$ ) at  $t = 0, t = \beta$  and two time-like surfaces ( $R_a, R_b$ ) at  $r = a, r = b$ . Since the metric is static, the time derivative in  $\partial_c P^c$  does not contribute (or, rather the contributions from  $S_0$  and  $S_\beta$  cancel each other). We thus only need to do the integral over  $R_a, R_b$  and—in these integrals—the integration over time reduces to multiplication by  $\beta$ . Thus

$$\frac{1}{16\pi} \int_{\mathcal{V}} \partial_c P^c d^4x = \frac{\beta}{16\pi} \int \partial_\alpha P^\alpha d^3x; \quad (\alpha = 1, 2, 3). \quad (\text{B.10})$$

If we convert this into a surface integral of the radial component  $P^r$  over the bounding surfaces  $R_a, R_b$ , then we will get

$$\begin{aligned} \frac{1}{16\pi} \int_{\mathcal{V}} d^4x \partial_c P^c &= \frac{\beta}{16\pi} \int d^3x \partial_\alpha P^\alpha \\ &= \frac{\beta}{16\pi} \int_{\theta=0}^{\theta=\pi} d\theta \int_0^{2\pi} d\phi P^r \Big|_a^b \end{aligned} \quad (\text{B.11})$$

where the vertical line denotes that the integral should be evaluated at  $r = b$  and  $r = a$  and the latter should be subtracted from the former. The relevant component is

$$P^r = 2g^{rr} \partial_r \sqrt{-g} + \sqrt{-g} \partial_r g^{rr} = -r^2 \sin\theta \left( \frac{4f}{r} + f' \right) \quad (\text{B.12})$$

(note that the  $\sqrt{-g}$  factors are built into  $P^c$ ; the  $d^n x$  just stands for  $d^0 x \dots d^{n-1} x$ ) giving the result

$$\frac{1}{16\pi} \int d^4 x \partial_c P^c = \frac{\beta}{4} [-r^2 f' - 4rf]_a^b. \tag{B.13}$$

This is exactly what we would have got by integrating  $2K = 2\nabla_a n^a$  over the  $r = \text{constant}$  surface. Using

$$K = \frac{1}{\sqrt{-g}} \partial_a (\sqrt{-g} g^{ab} n_b); \quad n_a = (0, 1, 0, 0) \frac{1}{\sqrt{f}} \tag{B.14}$$

we get the extrinsic curvature to be

$$K = -\sqrt{f} \left( \frac{1}{2} \frac{f'}{f} + \frac{2}{r} \right). \tag{B.15}$$

The integral on the right-hand side of (B.7) for the  $r = \text{constant}$  surface is

$$\begin{aligned} \frac{1}{8\pi} \int_0^\beta dt \int_0^\pi d\theta \int_0^{2\pi} d\phi \sqrt{|h|} K &= \frac{\beta}{8\pi} \int 2\pi d\theta (\sqrt{f} r^2 \sin \theta) K \\ &= \frac{\beta}{2} r^2 \sqrt{f} \left[ -\sqrt{f} \left( \frac{1}{2} \frac{f'}{f} + \frac{2}{r} \right) \right] = \frac{\beta}{4} [-r^2 f' - 4rf]. \end{aligned} \tag{B.16}$$

This result has been obtained in the literature several times, especially in the context of Schwarzschild spacetime in which  $R\sqrt{-g}$  vanishes and the only contribution to  $A_{SD}$  is from this term. The term on the right-hand side will contribute  $-4\pi M^2 = -\frac{1}{4}$ (horizon area) when evaluated on the horizon  $r = 2M$  and is often related to the entropy.

In spite of the simplicity of the calculation, *this result is algebraically wrong.*

Proving it wrong is quite trivial and can be done by computing (B.7) by evaluating the left-hand side directly. In  $\partial_c P^c$ , derivatives with respect to  $t$  and  $\phi$  vanish, giving

$$\partial_c P^c = \frac{\partial P^r}{\partial r} + \frac{\partial P^\theta}{\partial \theta}. \tag{B.17}$$

Integrating over the 4-volume the first term will give

$$\frac{1}{16\pi} \int_0^\beta dt \int_0^\pi d\theta \int_0^{2\pi} d\phi \int_a^b dr \left( \frac{\partial P^r}{\partial r} \right) = \frac{1}{16\pi} \int_0^\beta dt \int_0^\pi d\theta \int_0^{2\pi} d\phi P^r \Big|_a^b \tag{B.18}$$

which is precisely the contribution in (B.11). This will lead to the result quoted above. But there is a contribution from the second term! Since

$$P^\theta = 2g^{\theta\theta} \partial_\theta \sqrt{-g} + \sqrt{-g} \partial_\theta g^{\theta\theta} = -\frac{2}{r^2} r^2 \cos \theta = -2 \cos \theta \tag{B.19}$$

the second term gives

$$\begin{aligned} \frac{1}{16\pi} \int_0^\beta dt \int_0^{2\pi} d\phi \int_a^b dr \int_0^\pi d\theta \frac{\partial P^\theta}{\partial \theta} \\ = \frac{\beta}{8} \int_a^b dr [P^\theta(\theta = \pi) - P^\theta(\theta = 0)] = \frac{\beta}{2} (b - a) = \frac{\beta r}{2} \Big|_a^b. \end{aligned} \tag{B.20}$$

Adding the two contributions together, we get the correct result to be

$$\frac{1}{16\pi} \int \partial_c P^c d^4 x = \frac{1}{4} \beta [-r^2 f' + 2r(1 - 2f)]_a^b. \tag{B.21}$$

When evaluated on the Schwarzschild horizon (with  $\beta = 8\pi M$  being the inverse temperature related to the periodicity in Euclidean time) the extra term  $(\beta r/2)$  gives a contribution  $\beta M = 8\pi M^2$  which, when added to the original contribution  $(-4\pi M^2)$  from the  $P^r$ , leads

to a net contribution of  $(+4\pi M^2)$ . Thus the difference between the correct and wrong results in the case of Schwarzschild metric is just a flip of sign. It is easy to miss this or ‘reinterpret’ it, given the fact that one usually works in the Euclidean sector.

Unfortunately, this misses the correct interpretation. For Schwarzschild spacetime, the entropy and temperature are given by  $S = 4\pi M^2$ ,  $T = \beta^{-1} = 8\pi M$ ; so the combination  $\beta F \equiv S - \beta E$ , where  $F$  is the free energy, is given by

$$S - \beta E = 4\pi M^2 - (8\pi M)M = -4\pi M^2 = -S. \quad (\text{B.22})$$

The sign flip, arising due to the extra term, makes all the difference between the entropy  $S$  and the free energy  $S - \beta E$ .

To see this more clearly, let us consider a situation, for instance, in which  $f(r)$  has a simple zero at  $r = r_H$  with a finite derivative  $|f'(r_H)| \equiv B$ . This will lead to a compact horizon at  $r = r_H$ . Periodicity in the Euclidean time will now require  $\beta B = 4\pi$ . The *wrong* expressions ((B.13), (B.16)) will give, on the horizon (where  $f = 0$ )

$$\begin{aligned} I_{\text{wrong}} &= -\frac{1}{4}r_H^2\beta f'(r_H) = -\frac{1}{4}r_H^2\beta B = -\pi r_H^2 = -\frac{1}{4}(\text{horizon area}) \\ &= -S. \end{aligned} \quad (\text{B.23})$$

But the correct result (B.21) gives

$$\begin{aligned} I_{\text{correct}} &= -\frac{1}{4}r_H^2\beta f'(r_H) + \beta(r_H/2) = -\frac{1}{4}r_H^2\beta B + \beta(r_H/2) \\ &= -\pi r_H^2 + \beta(r_H/2) = -(S - \beta E) \end{aligned} \quad (\text{B.24})$$

with the energy associated with any horizon being given by  $E = (r_H/2)$ . It is this interpretation and the possibility of defining the energy for any horizon which is missed when the wrong result is used.

But how can a rigorously proved equation (B.7) go wrong? Actually, it did not. In the summation on the right-hand side of (B.7), one also needs to sum over a strange surface  $\theta = \text{constant}$  to get the correct result. To see this, note that the correct unit normal to the  $\theta = \text{constant}$  surface is

$$\begin{aligned} n_\theta &= \frac{1}{\sqrt{|g^{\theta\theta}|}}(0, 0, 1, 0) = r(0, 0, 1, 0); \\ n^\theta &= g^{\theta\theta}n_\theta = -\frac{1}{r}(0, 0, 1, 0). \end{aligned} \quad (\text{B.25})$$

The trace of the extrinsic curvature of this surface is

$$\begin{aligned} K &= \nabla_a n^a = \frac{1}{\sqrt{-g}}\partial_a(\sqrt{-g}n^a) = -\frac{1}{r^2\sin\theta}\partial_\theta\left(r^2\sin\theta\frac{1}{r}\right) \\ &= -\frac{1}{r}\cot\theta. \end{aligned} \quad (\text{B.26})$$

On a  $\theta = \text{constant}$  surface, the integral on the right-hand side of (B.8) will give

$$\begin{aligned} \frac{1}{8\pi}\int_{\theta=\text{const}}d^3x\sqrt{h}K &= -\frac{\beta}{8\pi}\int_0^{2\pi}d\phi\int_a^bdr\left[\sqrt{f}\frac{1}{\sqrt{f}}r\sin\theta\right]\left[\frac{\cot\theta}{r}\right] \\ &= -\frac{\beta}{4}(b-a)\cos\theta. \end{aligned} \quad (\text{B.27})$$

When we evaluate this contribution at the surfaces  $\theta = \pi$  and  $\theta = 0$  and subtract one from the other, we get

$$\frac{1}{8\pi}\int d^3x\sqrt{h}K\Big|_{\theta=0}^{\theta=\pi} = \frac{\beta}{2}(b-a) \quad (\text{B.28})$$

which is precisely the piece which was originally missing in action.

The reason for this ‘trouble’ is that there are situations with integrable singularities for which Gauss theorem is not applicable (or will give the wrong result) even in the ordinary three-dimensional flat space, in standard spherical polar coordinates. Consider a vector field with components

$$v^r = A(r) \sin \theta, \quad v^\theta = -\cos \theta, \quad v^\phi = 0 \quad (\text{B.29})$$

with  $A(r=0) = 0$ . We want to integrate the quantity  $\partial_\alpha v^\alpha$  over the 3D ball ( $B$ ) of radius  $R$ . Since

$$\partial_\alpha v^\alpha = (A'(r) + 1) \sin \theta \quad (\text{B.30})$$

the integral is

$$\begin{aligned} \int_0^R dr \int_0^{2\pi} d\phi \int_0^\pi d\theta \partial_\alpha v^\alpha &= 4\pi \int_0^R dr (A' + 1) \\ &= 4\pi [A(R) + R]. \end{aligned} \quad (\text{B.31})$$

This is the correct result, obtained by inelegant index dynamics. We will now do it in a different way. Let us define another vector field  $u^\alpha$  through the relation

$$d^3x (\partial_\alpha v^\alpha) = (\sqrt{g} d^3x) \left[ \frac{1}{\sqrt{g}} \partial_\alpha (\sqrt{g} u^\alpha) \right] = (\sqrt{g} d^3x) \nabla \cdot \mathbf{u}. \quad (\text{B.32})$$

The components of the new vector are given by

$$u^\alpha = \frac{v^\alpha}{\sqrt{g}} = \left( \frac{A(r)}{r^2}, -\frac{1}{r^2} \cot \theta, 0 \right). \quad (\text{B.33})$$

The integral of  $\partial_\alpha v^\alpha$  over  $d^3x \equiv dr d\theta d\phi$  is the same as the integral of the covariant divergence  $\nabla \cdot \mathbf{u}$  over the volume element  $\sqrt{g} d^3x$ . If we now use Gauss theorem, we get

$$\int d^3x \partial_\alpha v^\alpha = \int_B d^3x \sqrt{g} \nabla \cdot \mathbf{u} = \int_{\partial B} d^2x \sqrt{h} \mathbf{n} \cdot \mathbf{u} = 4\pi R^2 u^r(R) = 4\pi A(R). \quad (\text{B.34})$$

Only the radial component contributes and we get the *wrong result*. The second term  $4\pi R$  of (B.31) is missing; the original analysis clearly shows that it originated from  $(\partial v^\theta / \partial \theta)$  which we miss unless we deal with  $\theta = \text{constant}$  ‘surfaces’. How can an old faithful such as the Gauss theorem lead us astray? Before applying the Gauss theorem to the vector field  $\mathbf{u}$  in (B.34) we need to ensure that the vector field is not singular. Equation (B.33) clearly shows that the  $u^\theta$  component diverges on the  $z$ -axis corresponding to  $\theta = (0, \pi)$ . There is a singularity on this axis which gives the extra contribution.

In summary, we get the correct interpretation of the action as the free energy of horizon, when the surface terms are taken into account correctly.

## References

- [1] For a review of early results, see the articles in DeWitt C and DeWitt B S (eds) 1973 *Black Holes* (New York: Gordon and Breach)  
Carter B 1987 *Gravitation in Astrophysics, Cargese 1986* ed B Carter and J B Hartle (New York: Plenum)  
For a more recent discussion, see Bekenstein J 1994 Do we understand black hole entropy *Preprint gr-qc/9409015*
- [2] Bekenstein J D 1973 *Phys. Rev. D* **7** 2333  
Hawking S W 1975 *Commun. Math. Phys.* **43** 199  
Gerlach U H 1976 *Phys. Rev. D* **15** 1479  
t’Hooft G 1985 *Nucl. Phys. B* **256** 727  
York J 1985 *Phys. Rev. D* **15** 2929  
Zurek W H and Thorne K S 1985 *Phys. Rev. Lett.* **54** 2171

- Bombelli L *et al* 1986 *Phys. Rev. D* **34** 73  
 For a recent review, see Sorkin R D 1997 *Preprint* gr-qc/9705006
- [3] Hayward G and Wong K 1992 *Phys. Rev. D* **46** 620  
 Hayward G and Wong K 1993 *Phys. Rev. D* **47** 4778  
 Hayward G 1993 *Phys. Rev. D* **47** 3275  
 Brill D and Hayward G 1994 *Phys. Rev. D* **50** 4914  
 Martinez E A 1994 *Phys. Rev. D* **50** 4920  
 Hayward G 1995 *Phys. Rev. D* **52** 2001  
 Frolov V and Martinez E A 1996 *Class. Quantum Grav.* **13** 481  
 Booth I S and Mann R B 1999 *Phys. Rev. D* **60** 124004  
 Booth I S and Creighton J 2000 *Phys. Rev. D* **62** 067503  
 Epp R J 2000 *Phys. Rev. D* **62** 124018
- [4] Padmanabhan T 2002 *Mod. Phys. Lett. A* **17** 923 (*Preprint* gr-qc/0202078)  
 This view has also been expressed in the literature before, see for example, Jacobsen T 1999 *Preprint* gr-qc/9908031
- [5] Padmanabhan T 2002 *Mod. Phys. Lett. A* **17** 1147 (*Preprint* hep-th/0205278)
- [6] Lin Feng-Li and Soo Chopin 1999 *Class. Quantum Grav.* **16** 551
- [7] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco, CA: Freeman) p 889
- [8] Ashtekar A, Beetle C and Fairhurst S 2000 *Class. Quantum Grav.* **17** 253
- [9] Spradlin M *et al* 2001 *Preprint* hep-th/0110007
- [10] Bardeen J H, Carter B and Hawking S W 1973 *Commun. Math. Phys.* **31** 181
- [11] Carlip S 1998 *Quantum Gravity in 2 + 1 Dimension* (Cambridge: Cambridge University Press)
- [12] Jacobsen T 1995 *Phys. Rev. Lett.* **75** 1260
- [13] Bekenstein J D and Mayo A E 2001 *Preprint* gr-qc/0105055
- [14] Padmanabhan T 2002 The holography of gravity encoded in a relation between entropy, horizon area and action for gravity *Preprint* gr-qc/0205090  
 Padmanabhan T 2002 Second prize essay in the Gravity Research Foundation Essay Contest *Gen. Rel. Grav.* at press
- [15] Gibbons G W and Hawking S W 1977 *Phys. Rev. D* **15** 2738
- [16] Brown J D, Lau S R and York J W *Preprint* gr-qc/0010024  
 Banados M, Teitelboim C and Zanelli J 1994 *Phys. Rev. Lett.* **72** 957  
 Ashworth M C and Hayward S A 2000 *Phys. Rev. D* **62** 064024  
 Iyer V and Wald R M 1995 *Phys. Rev. D* **52** 4430  
 Brown J D and York J W 1993 *Phys. Rev. D* **47** 1420  
 Brown J D and York J W 1993 *Phys. Rev. D* **47** 1407  
 Lau S R 1993 *Class. Quantum Grav.* **10** 2379  
 Brown J D, Creighton J and Mann R 1994 *Phys. Rev. D* **50** 6394  
 Hawking S W and Horowitz G T 1996 *Class. Quantum Grav.* **13** 1487  
 Hawking S W and Hunter C J 1996 *Class. Quantum Grav.* **13** 2735  
 Brown J D, Lau S R and York J W 1997 *Phys. Rev. D* **55** 1977  
 Brown J D, Lau S R and York J W 1999 *Phys. Rev. D* **59** 064028
- [17] Gibbons G W and Hawking S W 1977 *Phys. Rev. D* **15** 2738