Black Hole Thermodynamics in a Box

David Brown

Abstract
Simple calculations indicate that the partition function for a black hole is defined only if the temperature is fixed on a finite boundary. Consequences of this result are discussed.

1 The Black Hole Partition Function
From the work of Gibbons and Hawking in the late 1970’s came a very simple prescription for the computation of the temperature of a static black hole [1]:
• Write the black hole metric in static coordinates;
• Euclideanize ($t \rightarrow -it$) and periodically identify $t$;
• Adjust the period to remove conical singularities.
The resulting period is the inverse temperature $\beta$. The origin of this prescription is a formal calculation of the partition function $Z(\beta)$ as a functional integral over all Euclidean geometries $g$ with period $\beta$ and action $I[g]$. Some of the key features of this calculation can be captured in a ‘microsuperspace’ version based on the metric ansatz [2]

$$ds^2 = N^2(r) dt^2 + (1 - 2M/r)^{-1} dr^2 + r^2 d\Omega^2.$$  (1)

Let $t$ have the range $0 \leq t < 2\pi$ and $r$ have the range $2M \leq r < \infty$. Also restrict $N$ so that $2\pi N(\infty) = \beta$ and $2\pi N(r) \sim 8\pi M \sqrt{1 - 2M/r}$ near $r = 2M$. The first restriction fixes the proper period at infinity to the inverse temperature. The second restriction insures that the metric (1) describes a smooth geometry with no conical singularities, and with topology $R^2 \times S^2$. The action for (1) is a function of $M$ only, $I(M) = M\beta - 4\pi M^2$. A ‘toy’ partition function can be constructed as the integral over $M$ of $\exp(-I(M))$. The extremum of the action satisfies $0 = \partial I/\partial M = \beta - 8\pi M$. (This is the classical equation of motion obtained by integrating $(Nr/2)(1 - 2M/r)^{-3/2}G_{rr}$ over $t$ and $r$, where $G_{rr}$ is the $r$-$r$ component of the Einstein tensor.) The solution for the extremum is a Euclidean black hole with $M = \beta/(8\pi)$, and the partition function is classically approximated by $\ln Z(\beta) \approx -I(\beta/(8\pi)) = -\beta^2/(16\pi)$. The expectation value of energy is $\langle E \rangle \equiv -\langle \partial \ln Z/\partial \beta \rangle \approx \beta/(8\pi)$, which equals the extremal value of the mass parameter $M$. An interpretation of these results is that $Z(\beta)$ describes a system that contains a black hole of mass $M$ and inverse temperature $\beta = 8\pi M$.

What about pre–exponential factors in $Z(\beta)$? A simple calculation shows that the second derivative $\partial^2 I/\partial M^2$ is negative at the extremum $M = \beta/(8\pi)$. Therefore, the extremum lies along a path of steepest ascents, not along a path of steepest descents. The Euclidean black hole does not dominate the integral for $Z(\beta)$, and should not be used to approximate $Z(\beta)$. As a consequence, the conclusion $\langle E \rangle = \beta/(8\pi)$ is unfounded. Formally, the Eu-
clidean black hole makes an imaginary contribution to the partition function, and should be interpreted as an instanton that governs black hole nucleation [3].

So the prescription given above for the temperature of a black hole is not justified, perhaps not even correct. Yet, it is tempting to believe in that prescription because apparently it gives the correct result $\beta = 8\pi M$ for the Hawking temperature. In order to understand this puzzling situation, consider a modified microsuperspace calculation [2]. As before, the metric ansatz is given by Eq. (1) with $0 \leq t < 2\pi$ and $2\pi N(r) \sim \frac{8\pi M}{\sqrt{1 - \frac{2M}{r}}} \sim \frac{8\pi M}{\sqrt{1 - \frac{2M}{r}}}$ near $r = 2M$. In this case, however, the system is placed in a finite ‘box’ of size $R$ by restricting $r$ to the range $2M \leq r \leq R$ and fixing the proper period at $r = R$ to the inverse temperature: $2\pi N(R) = \beta$. The action is $I(M) = R\beta (1 - \sqrt{1 - \frac{2M}{R}}) - 4\pi M^2$, and a toy partition function is constructed by integrating $\exp(-I(M))$ over $M$. The action is extremized for $M$ satisfying

$$
\beta = 8\pi M \sqrt{1 - \frac{2M}{R}}.
$$

(Again, this is related to the $G\tau = 0$ Einstein equation.) There are two solutions to Eq. (2), $M_1$ and $M_2$ with $M_1 < M_2$. $\partial^2 I / \partial M^2$ is negative at the extremum $M_1$, and $M_1 \to \beta/(8\pi)$ as $R \to \infty$ with $\beta$ fixed. Thus, the Euclidean black hole with mass parameter $M_1$ is an instanton. On the other hand, $\partial^2 I / \partial M^2$ is positive at the extremum $M_2$, and $M_2 \sim R/2 \to \infty$ as $R \to \infty$ with $\beta$ fixed. It follows that the Euclidean black hole with mass parameter $M_2$ can be used for a steepest descents approximation to $Z(\beta)$. In the classical approximation, $\ln Z(\beta) \approx -I(M_2)$ and the expectation value of energy is $\langle E \rangle \approx R(1 - \sqrt{1 - \frac{2M_2}{R}})$. $\langle E \rangle$ can be expanded in powers of $GM_2/R$ (where Newton’s constant $G$ is set to unity) with the result $\langle E \rangle \approx M_2 + M_2^2/(2R) + \cdots$. This shows that the energy $\langle E \rangle$ inside the box $R$ equals the energy at infinity $M_2$ minus the binding energy $-M_2^2/(2R)$ of a shell of mass $M_2$ and radius $R$, which is the energy associated with the gravitational field outside $R$ [4].

The calculation of $Z(\beta)$ above supports the conclusion that Eq. (2) gives the inverse temperature at $R$ of a black hole with mass $M$. Note that the square root in (2) is the Tolman redshift factor for temperature in a stationary gravitational field. In what sense can one say that the inverse temperature of a black hole is $8\pi M$? That statement does not follow from taking the limit $R \to \infty$ of Eq. (2), since in that limit $M \to \infty$ as well. Rather, $8\pi M$ is the inverse temperature obtained by dropping the Tolman redshift factor from Eq. (2). Physically, this corresponds to drilling a small hole in the box and letting some radiation leak out to infinity. The inverse temperature of the black hole as measured by the radiation at infinity is $8\pi M$.

It is important to recognize that the partition function, by itself, does not give the result $8\pi M$ (or any result) for the inverse temperature at infinity. With $\beta$ fixed at infinity, as in the first microsuperspace calculation presented above, the (real part of the) partition function does not exist. This is a consequence of the physical fact that a gravitating system in an infinitely large cavity at nonzero temperature cannot be in equilibrium, because black holes will form and grow without bound. In order to conclude that the inverse temperature at infinity is $8\pi M$, it is necessary to supplement the partition function analysis with a further physical argument. In the argument given above, a small amount of radiation is allowed to leak from the box to infinity. Having argued in this way that the inverse temperature of an equilibrium black hole, as measured at infinity, is $8\pi M$, one can go a step farther and consider removing the box altogether. The black hole will no longer be in thermal equilibrium, but to the extent that it evolves relatively slowly it is justified to identify $8\pi M$ as the inverse temperature at infinity.
The calculation of the partition function for a system in a finite box leads to a corrected prescription for the temperature of a static black hole:

- Write the black hole metric in static coordinates;
- Euclideanize \( t \rightarrow -it \) and periodically identify \( t \);
- Fix the proper period to \( \beta \) at \( R \);
- Adjust the mass parameter \( M \) to remove conical singularities.

This yields two values for \( M \). The larger \( M \) is the mass of a black hole with inverse temperature \( \beta \) at \( R \). (This prescription holds as stated for 3+1 dimensional Einstein gravity with a negative or vanishing cosmological constant [5]. In other cases there might be more or fewer than two extrema \( M \). In order to distinguish among the instantons and the stable or quasi–stable black holes, the sign of \( \partial^2 I/\partial M^2 \) must be checked for each extremum.)

2 Temperature of Gravitating Systems

For the canonical partition function \( Z(\beta) \) of a gravitating system the temperature must be fixed at a finite boundary \( B \). This has an important consequence: Since temperature redshifts and blueshifts in stationary gravitational fields, one must allow the temperature to be fixed to different values at different points on \( B \). In other words, gravitating systems are not characterized by a single temperature but instead by a temperature field on the boundary of the system [6]. Correspondingly, the partition function is a functional \( Z[\beta] \) of the inverse temperature field on \( B \). For a typical problem (such as the microsuperspace calculation of the previous section), it is possible to choose the temperature to be a constant on \( B \), in which case \( B \) coincides with an isothermal surface for the system. However, experience with the Kerr black hole shows that this must be viewed as a particular choice of boundary conditions, not the most general choice. What happens in the Kerr case is that the angular velocity of the black hole with respect to observers who are at rest in the stationary time slices enters as a chemical potential conjugate to angular momentum. It turns out that the constant temperature surfaces and the constant angular velocity surfaces do not coincide. Therefore it is necessary to allow for some thermodynamical data, either the temperature or the chemical potential or both, to vary across the boundary. This conclusion might seem disturbing at first, since traditionally one of the purposes of thermodynamics has been to provide a characterization of systems in terms of only a few parameters. Nevertheless, the thermodynamical formalism that results from a generalization to non–constant thermodynamical data has a number of compelling features. In particular, thermodynamical data is brought into direct correspondence with canonical boundary data, and in the process an intimate connection between thermodynamics and dynamics is revealed [6].

References

[1] G. W. Gibbons and S. W. Hawking, \textit{Phys. Rev.} \textbf{15} (1977) 2752.
[2] This is based on B. F. Whiting and J. W. York, \textit{Phys. Rev. Lett.} \textbf{61} (1988) 1336; H. W. Braden, J. D. Brown, B. F. Whiting, and J. W. York, \textit{Phys. Rev.} \textbf{D42} (1990) 3376.
[3] D. J. Gross, M. J. Perry, and L. G. Yaffe, \textit{Phys. Rev.} \textbf{D25} (1982) 330; J. W. York, \textit{Phys. Rev.} \textbf{D33} (1986) 2092.
[4] J. D. Brown and J. W. York, \textit{Phys. Rev.} \textbf{D47} (1993) 1407.
[5] J. D. Brown, J. Creighton, and R. B. Mann, submitted to \textit{Phys. Rev.} \textbf{D}.
[6] J. D. Brown, E. A. Martinez, and J. W. York, \textit{Phys. Rev. Lett.} \textbf{66} (1991) 2281; J. D. Brown and J. W. York, \textit{Phys. Rev.} \textbf{D47} (1993) 1420; J. D. Brown and J. W. York, in \textit{Physical Origins of Time Asymmetry}, edited by J. J. Halliwell, J. Perez–Mercader, and W. Zurek (Cambridge University Press, Cambridge, 1994).