The Statistical Mechanics
of the (2+1)-Dimensional Black Hole

S. Carlip*

Department of Physics
University of California
Davis, CA 95616
USA

Abstract

The presence of a horizon breaks the gauge invariance of (2+1)-dimensional general relativity, leading to the appearance of new physical states at the horizon. I show that the entropy of the (2+1)-dimensional black hole can be obtained as the logarithm of the number of these microscopic states.

---

*email: carlip@dirac.ucdavis.edu

1. Introduction

Black holes possess a temperature and an entropy, and obey the usual laws of thermodynamics. Despite twenty years of research, however, black hole thermodynamics remains something of an anomaly. The thermal properties of ordinary physical systems arise from the statistical mechanics of microscopic states. But a classical black hole is completely characterized by its mass and angular momentum, leaving little room for additional microscopic physics. Black holes have entropy, but we do not know why.

The dependence of black hole temperature and entropy on Planck's constant suggests that these quantities are fundamentally quantum mechanical in nature. The goal of this paper is to demonstrate a mechanism whereby black hole entropy—at least for the (2+1)-dimensional black hole of Bañados, Teitelboim, and Zanelli [1]—can be obtained by counting quantum gravitational states at the horizon. The restriction to 2+1 dimensions is, of course, a serious limitation, but if enough new states can be found to account for black hole entropy in this simple setting, it is reasonable to hope that the same will happen in the vastly richer arena of realistic (3+1)-dimensional gravity.

The basic argument is quite simple. Begin by considering general relativity on a manifold M with boundary. We ordinarily split the metric into true physical excitations and "pure gauge" degrees of freedom that can be removed by diffeomorphisms of M. But the presence of a boundary alters the gauge invariance of general relativity: the infinitesimal transformations $g \to g + \mathcal{L}_\xi g$ must now be restricted to those generated by vector fields ξ with no component normal to the boundary, that is, true diffeomorphisms that preserve ∂M. As a consequence, some degrees of freedom that would naively be viewed as "pure gauge" become dynamical, introducing new degrees of freedom associated with the boundary.

Now, the event horizon of a black hole is not a true boundary, although the black hole complementarity approach of Susskind et al. [2] suggests that it might be appropriately treated as such. Regardless of one's view of that program, however, it is clear that in order to ask quantum mechanical questions about the behavior of black holes, one must put in "boundary conditions" that ensure that a black hole is present. This means requiring the existence of a hypersurface with particular metric properties—say, those of an apparent horizon.

The simplest way to do quantum mechanics in the presence of such a surface is to quantize fields separately on each side, imposing the appropriate correlations as boundary conditions. In a path integral approach, for instance, one can integrate over fields on each side, equate the boundary values, and finally integrate over those boundary values compatible with the existence of a black hole. But this process again introduces boundary terms that restrict the gauge invariance of the theory, leading once more to the appearance of new degrees of freedom at the horizon that would otherwise be treated as unphysical.

My suggestion is that black hole entropy is determined by counting these

would-be gauge degrees of freedom. The resulting picture is similar to Maggiore's membrane model of the black hole horizon [3], but with a particular derivation and interpretation of the "membrane" degrees of freedom.

The analysis of this phenomenon is fairly simple in 2+1 dimensions. It is well known that (2+1)-dimensional gravity can be written as a Chern-Simons theory [4,5], and it is also a standard result that a Chern-Simons theory on a manifold with boundary induces a dynamical Wess-Zumino-Witten (WZW) theory on the boundary [6, 7]. In the presence of a cosmological constant $\Lambda = -1/\ell^2$ appropriate for the (2+1)-dimensional black hole, one obtains a slightly modified $SO(2,1) \times SO(2,1)$ WZW model, with coupling constant

$$k = \frac{\ell\sqrt{2}}{8G}. \tag{1.1}$$

This model is not completely understood, but in the large k—i.e., small Λ—limit, it may be approximated by a theory of six independent bosonic oscillators. I show below that the Virasoro operator L_0 for this theory takes the form

$$L_0 \sim N - \left(\frac{r_+}{4G}\right)^2, \tag{1.2}$$

where $N = \sum_{i=1}^{6} N_i$ is a number operator and r_+ is the horizon radius. It is a standard result of string theory [8] that the number of states of such a system behaves asymptotically as

$$n(N) \sim \exp\left\{\pi\sqrt{6 \cdot \frac{2N}{3}}\right\}. \tag{1.3}$$

If we demand that L_0 vanish—physically, requiring states to be independent of the choice of origin of the angular coordinate at the horizon—we thus obtain

$$\log n(r_+) \sim \frac{2\pi r_+}{4G}, \tag{1.4}$$

precisely the right expression for the entropy of the (2+1)-dimensional black hole [1, 9].

I now turn to the details of this analysis.

2. Chern-Simons Theory, WZW Models, and Gravity

The most thoroughly studied example of boundary dynamics arising from a simple "bulk" action is that of Chern-Simons theory. The Chern-Simons action for a three-manifold M is

$$I_{\text{CS}} = \frac{k}{4\pi}\int_M \text{Tr}\left(A \wedge dA + \frac{2}{3}A \wedge A \wedge A\right), \tag{2.1}$$

where $A = A^a_\mu T_a \, dx^\mu$ is a gauge field (connection one-form) for a group G whose Lie algebra is generated by $\{T_a\}$. If M is closed, this action is invariant under gauge transformations

$$\bar{A} \to A = g^{-1}dg + g^{-1}\bar{A}g. \tag{2.2}$$

If M has a boundary, however, this invariance is broken: the integrand of (2.1) is invariant only up to a total divergence, which can lead to a nontrivial term on ∂M. Moreover, in order for the theory to admit classical solutions, I_{CS} must be supplemented by a surface term, whose exact form depends on the choice of boundary conditions. In particular, if ∂M has the topology of a cylinder, parametrized by an angular coordinate ϕ and a linear coordinate v, and if A_ϕ is fixed at the boundary, then the required term is

$$I'_{\text{CS}} = -\frac{k}{4\pi} \int_{\partial M} \text{Tr} A_\phi A_v, \tag{2.3}$$

which further breaks the gauge invariance of the "bulk" action.

To isolate the gauge dependence of the action, it is useful to partition the space of connections into gauge orbits parametrized by gauge-fixed connections $\bar{A}$, with points on each orbit labeled by group elements g. In the path integral context, this may be recognized as the first step in Faddeev-Popov gauge-fixing. Writing a general connection in the form (2.2), it is easy to show that [10, 11]

$$(I_{\text{CS}} + I'_{\text{CS}})[A] = (I_{\text{CS}} + I'_{\text{CS}})[\bar{A}] - k I^+_{\text{WZW}}[g, \bar{A}], \tag{2.4}$$

where

$$I^+_{\text{WZW}}[g, \bar{A}] = \frac{1}{4\pi} \int_{\partial M} \text{Tr} \left(g^{-1}\partial_\phi g \right) \left(g^{-1}\partial_v g \right)$$
$$+ \frac{1}{2\pi} \int_{\partial M} \text{Tr} \left(g^{-1}\partial_v g \right) \left(g^{-1}\bar{A}_\phi g \right) + \frac{1}{12\pi} \int_M \text{Tr} \left(g^{-1}dg \right)^3 \tag{2.5}$$

is the standard chiral Wess-Zumino-Witten action on ∂M. This action can be obtained in a number of ways [7], but the derivation given here points to a particular physical interpretation: the boundary variables are would-be "pure gauge" degrees of freedom that become dynamical because the presence of a boundary restricts the allowed gauge transformations.

As first shown by Achúcarro and Townsend [4], (2+1)-dimensional gravity can itself be written as a Chern-Simons theory. In particular, for the case of a cosmological constant $\Lambda = -1/\ell^2$, one can define two $SO(2,1)$ gauge fields

$$A^a = \omega^a + \frac{1}{\ell} e^a, \quad \tilde{A} = \omega^a - \frac{1}{\ell} e^a, \tag{2.6}$$

where $e^a = e^a_\mu dx^\mu$ is a triad and $\omega^a = \frac{1}{2}\epsilon^{abc}\omega_{\mu bc}dx^\mu$ is a spin connection. The first-order form of the Einstein action is then

$$I_{\text{grav}} = I_{\text{CS}}[A] - I_{\text{CS}}[\tilde{A}]. \tag{2.7}$$

3

The value of the coupling constant k depends on normalization. For this paper, I choose

$$(T_a)_b{}^c = -\epsilon_{abd}\eta^{dc}, \quad \eta_{ab} = \text{diag}\,(-1,1,1), \quad \epsilon_{012} = 1, \tag{2.8}$$

and define Tr as the ordinary matrix trace, so

$$[T_a, T_b] = f_{ab}{}^c T_c = \epsilon_{abd}\eta^{dc}T_c, \quad \text{Tr}\,T_a T_b = \hat{g}_{ab} = 2\eta_{ab}, \quad g_{\mu\nu} = \hat{g}_{ab}e^a_\mu e^b_\nu = 2\eta_{ab}e^a_\mu e^b_\nu;$$

$$\hat{g}^{ad}\hat{g}^{be}f_{ab}{}^c f_{de}{}^f = Q\hat{g}^{cf} \quad \text{with } Q = -1. \tag{2.9}$$

Then k is given by (1.1), as is most easily checked by comparing the value of the action (2.7) at a classical solution to the corresponding Einstein action. The diffeomorphism invariance of general relativity translates into gauge invariance of the Chern-Simons action: appropriate combinations of gauge transformations of A and $\tilde{A}$ are equivalent to diffeomorphisms [5]. We should thus expect (2+1)-dimensional gravity to induce a pair of SO(2,1) WZW actions on ∂M, whose degrees of freedom correspond in some sense to deformations of the horizon.

3. The Boundary Action

To determine the exact form of these boundary WZW actions, we must now choose boundary conditions for A and $\tilde{A}$. Ideally, we would require the existence of a black hole by imposing the requirement that an event horizon be present. Unfortunately, a genuine event horizon is a complicated global object, and it is difficult to translate its existence into local boundary conditions. Let us therefore impose the simpler requirement that ∂M be an apparent horizon (strictly speaking, a trapping horizon [12]).

To do so, we write the metric in double null (light cone) coordinates [12],

$$ds^2 = -2e^{-f}dudv + r^2(u,v)\left(d\phi + N^u du + N^v dv\right)^2. \tag{3.1}$$

It may then be shown that the spin connection $\omega^+ = \omega^0 + \omega^1$ is

$$\omega^+ = 2re^{f/2}\,\theta^+ \left(d\phi + N^u du + N^v dv\right) + B^+ du, \tag{3.2}$$

where $\theta^+(u,v,\phi)$ is the expansion of the outgoing null geodesic congruence at (u,v,ϕ) and B^+ is a complicated but irrelevant function. For a circle (u_0, v_0, ϕ) to lie on an apparent horizon, we require that $\theta^+(u_0, v_0, \phi)$ vanish. If in addition the stress-energy tensor vanishes at (u_0, v_0, ϕ)—or, more narrowly, if T^{++} vanishes—then it is easy to show (classically) that the apparent horizon is null at that point. From (3.2), the boundary conditions are thus $\omega^+_\phi = \omega^+_v = 0$, or

$$A^+_\phi = A^+_v = \tilde{A}^+_\phi = \tilde{A}^+_v = 0. \tag{3.3}$$

These conditions are not quite sufficient; we must also prescribe appropriate boundary values for A^2 and $\tilde{A}^2$ at ∂M. As shown in [9], the right choice for black

hole thermodynamics is to fix the horizon radius r_+ and either a component p_+ of the extrinsic curvature or its canonical conjugate, the shift vector N^ϕ. In the coordinates (3.1), p_+ is determined by the spin connection component ω_ϕ^2, and suitable boundary conditions are

$$e_\phi^2 = \frac{r_+}{\sqrt{2}}, \quad \omega_\phi^2 = \bar{\omega}. \tag{3.4}$$

(The factor of $\sqrt{2}$ comes from the normalizations (2.9).)

The boundary conditions (3.3)–(3.4) are not, of course, diffeomorphism-invariant. This is as it should be, since our aim is to separate out the "diffeomorphism" degrees of freedom at ∂M. The boundary conditions are, however, invariant under rigid translations $\phi \to \phi + c(v)$, that is, time-dependent shifts of the origin of ϕ. Thus while most of the diffeomorphisms of the horizon will be absorbed into the new dynamical fields of the WZW model at ∂M, these rigid translations remain as symmetries.

Now, the physical meaning of r_+ is clear—$2\pi r_+$ is the circumference of the boundary ∂M, on or off shell. The interpretation of $\bar{\omega}$ is more problematic. On shell, it may be shown that in Kruskal coordinates, $\bar{\omega} = 2r_-/\ell$, where r_- is the value of r at the inner horizon. But this relationship is coordinate-dependent, and its off-shell generalization is not at all obvious; for an arbitrary metric, $\bar{\omega}$ depends not only on the horizon geometry, but also on normal derivatives. Moreover, even on shell, $\bar{\omega}$ is determined only modulo an integer in the Euclideanized theory [9].

It would therefore be preferable to fix the shift vector N^ϕ—or equivalently, e_v^2—at ∂M. Unfortunately, this would lead to boundary conditions that mix A and $\tilde{A}$, making the induced boundary action much more complicated. I shall instead argue as follows. A standard choice of boundary conditions for the black hole is to set $N^\phi = 0$ at ∂M, but this choice is somewhat conventional, since N^ϕ is determined only up to an integration constant [1], which may be shifted by the rigid rotations of the horizon discussed above. So let us instead require that $\partial_\phi N^\phi = 0$, and sum over the constant values of the shift vector to count macroscopically indistinguishable states. Since N^ϕ and $\bar{\omega}$ are canonically conjugate, this should be equivalent to summing over constant values of $\bar{\omega}$. I will therefore adopt the boundary conditions (3.3)–(3.4), and integrate over $\bar{\omega}$ to count states.

Given the boundary conditions (3.3)–(3.4), the induced action on ∂M is not hard to determine. One obtains

$$I[g, \bar{A}] = -kI_{\mathrm{WZW}}^+[g, \bar{A}] - \frac{k}{2\pi} \int_{\partial M} \left(g^{-1}\partial_v g + g^{-1}A_v g\right)^+ \left(g^{-1}\partial_\phi g + g^{-1}\bar{A}_\phi g\right)^- \tag{3.5}$$

with a similar expression for $\tilde{A}$. The last term in (3.5) is most easily understood by noticing that the Chern-Simons boundary action (2.3) is appropriate for fixing A_ϕ^+ and A_ϕ^-; if we wish instead to fix A_ϕ^+ and A_v^-, we need an additional boundary

term of the form $\int A_\phi^- A_v^+$. Note that (3.5) is not a "gauged WZW action" in the usual sense of the term, since the fields $\bar{A}$ are fixed by the boundary data, and not integrated out.

The action $I[g, \bar{A}]$ is no longer quite the standard SO(2, 1) WZW action, but it is classically equivalent. Indeed, if we define an element h of SO(2, 1) by

$$\partial_v h \cdot h^{-1} = (g^{-1}\partial_v g + g^{-1}\bar{A}_v g)^+ T^-, \tag{3.6}$$

it follows from the Polyakov-Wiegmann formula [13] that

$$I[g, \bar{A}] = -k I_{\text{WZW}}^+[gh, \bar{A}]. \tag{3.7}$$

Conversely, h is determined from gh by the condition

$$\left(h \cdot J_v[gh] \cdot h^{-1} \right)^+ = 0, \tag{3.8}$$

where $J[g] = g^{-1}\partial_v g + g^{-1}\bar{A}_v g$.

In the quantum theory, the change of variables from g to gh will lead to a Jacobian, which can be most easily determined by using the Gauss decomposition of g,

$$g = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^\lambda & 0 \\ 0 & e^{-\lambda} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ \hat{h} & 1 \end{pmatrix}. \tag{3.9}$$

In this representation, the change of variables from b to $b + \hat{h}$ is relatively easy to evaluate; one finds a Jacobian of the form

$$\mathcal{J} = \left| \det \left(\partial_v^{-1} A_v^2 \right) \right| \tag{3.10}$$

with A_v^2 defined by (2.2). This Jacobian will renormalize terms in the action (3.5), and its careful treatment is necessary for a full evaluation of the boundary WZW theory. (A similar deformation has been considered by Förste [14].) In the large k, or semiclassical, limit, however, it should be possible to neglect this correction. For the purposes of this paper, I will therefore work with the usual SO(2, 1)$\times$SO(2, 1) WZW action.

4. Counting States

It remains for us to count the states of this induced boundary theory. Observe first that a WZW model is completely characterized by a current algebra [15, 16, 17]

$$[J_m^a, J_n^b] = if^{ab}{}_c J_{m+n}^c - km\hat{g}^{ab}\delta_{m+n,0}, \quad [\tilde{J}_m^a, \tilde{J}_n^b] = if^{ab}{}_c \tilde{J}_{m+n}^c + km\hat{g}^{ab}\delta_{m+n,0}, \tag{4.1}$$

6

where the currents J_n^a and $\tilde{J}_n^a$ are defined by the expansions [18]

$$A_\phi^a = -\frac{1}{k} \sum_{n=-\infty}^{\infty} J_n^a e^{in\phi}, \quad \tilde{A}_\phi^a = \frac{1}{k} \sum_{n=-\infty}^{\infty} \tilde{J}_n^a e^{in\phi}. \tag{4.2}$$

Here, A^a and $\tilde{A}^a$ are the gauge fields of equation (2.2)—insertion of that equation into (4.2) gives the usual dependence of the currents on g and $\tilde{g}$, with signs determined by (2.7) and (3.7)—and indices are raised and lowered with the metric $\hat{g}_{ab}$ of (2.9). To find the Hilbert space, we must thus find an appropriate representation of the affine Lie algebra (4.1).

The standard choice is a highest weight representation. We start with a vacuum multiplet $|\Omega\rangle$ that is annihilated by the J_n^a and $\tilde{J}_n^a$ with $n > 0$, and that transforms under a representation of the $\mathrm{SO}(2,1) \times \mathrm{SO}(2,1)$ generated by the zero-modes J_0^a and $\tilde{J}_0^a$. This representation is determined by the boundary conditions (3.3)–(3.4), which imply that

$$(J_0)^2 |\Omega\rangle = 2k^2 \left(\bar{\omega} + \frac{r_+}{\sqrt{2}\ell} \right)^2 |\Omega\rangle, \quad (\tilde{J}_0)^2 |\Omega\rangle = 2k^2 \left(\bar{\omega} - \frac{r_+}{\sqrt{2}\ell} \right)^2 |\Omega\rangle. \tag{4.3}$$

Note that the positivity of $(J_0)^2$ and $(\tilde{J}_0)^2$ implies that the relevant representation must be in the continuous series. (See [19] for a nice discussion of $\mathrm{SO}(2,1)$ representations.) The zero-modes can alternatively be obtained heuristically from section 2.3 of reference [7], treating the term $\bar{A}_\phi$ in (2.5) as a "source" term, which determines the relevant representation of the affine algebra. Equivalently, in the coadjoint orbit approach to quantization [20], the $\bar{A}$ term determines the appropriate orbit in $\mathrm{SO}(2,1) \times \mathrm{SO}(2,1)$.

The remaining states are now obtained by acting on $|\Omega\rangle$ with raising operators J_{-n}^a and $\tilde{J}_{-n}^a$. Because of the commutators (4.1), this process is rather nontrivial, and the Hilbert space of the $\mathrm{SO}(2,1)$ WZW model is not fully understood (but see [19, 21, 22, 23, 24]). In the large k limit, however, the components of the currents J and $\tilde{J}$ decouple, and the states can be approximated as those of a six-dimensional bosonic string theory. This phenomenon is most easily seen by rescaling the currents in (4.1) by $k^{-1/2}$; in the large k limit, the terms involving the structure constants drop out, leaving a set of $\widehat{u(1)}$ commutators at level ± 1.

In string theory, one must impose the added restriction that physical states be annihilated by the Virasoro generators L_n ($n > 0$). This condition comes from the requirement of diffeomorphism invariance; classically, the L_n generate the diffeomorphisms of the circle. As we saw in the last section, this is not an appropriate requirement here: the boundary conditions (3.3)–(3.4) are not diffeomorphism-invariant, and indeed, the existence of our boundary degrees of freedom directly reflect this noninvariance. Our boundary conditions are, however, still invariant under time-dependent rigid translations of ϕ, and we must

therefore require the corresponding invariance of the states, i.e., the vanishing[*] of L_0. The condition $L_0|\psi\rangle = 0$ may be viewed as a last remnant of the Wheeler-DeWitt equation.

In the conventions of this paper, the Virasoro operator L_0 for the action (2.7) is

$$L_0 = -\frac{1}{2k-Q} \sum_{n=-\infty}^{\infty} :J_{-n}^a J_n^b: \hat{g}_{ab} + \frac{1}{2k+Q} \sum_{n=-\infty}^{\infty} :\tilde{J}_{-n}^a \tilde{J}_n^b: \hat{g}_{ab}, \qquad (4.4)$$

which satisfies

$$[L_0, J_n^a] = -nJ_n^a, \qquad [L_0, \tilde{J}_n^a] = -n\tilde{J}_n^a. \qquad (4.5)$$

These commutation relations imply that the non-zero mode contributions to L_0 take the form of number operators, which in the large k limit can be taken to be independent, while the zero-mode contributions are determined by (4.3). Combining terms and using the normalizations (2.9), we find

$$L_0 = \sum_{i=1}^{6} N_i + \frac{4k^2}{4k^2-1}\left(\bar{\omega} - \frac{\sqrt{2}kr_+}{\ell}\right)^2 - \frac{2k^2 r_+^2}{\ell^2} \qquad (4.6)$$

with $k = \ell\sqrt{2}/8G$. The condition $L_0|\psi\rangle = 0$ thus determines $\sum N_i$ in terms of r_+ and $\bar{\omega}$.

Now, given a set of independent number operators N_i, it is fairly easy to determine the number of states. A fixed component J of the current creates states of the form

$$|(n_1, a_1), (n_2, a_2), \ldots\rangle = (J_{-n_1})^{a_1}(J_{-n_2})^{a_2}\ldots|\Omega\rangle, \qquad (4.7)$$

for which $N = \sum a_i n_i$. The number of states is then given by the number of ways of writing N in this form. This is essentially the partition function of number theory [25], whose asymptotic behavior is given by (1.3); the factor of 6 in the square root is the number of independent N_i. A similar expression occurs for unitary representations of arbitrary affine Lie algebras based on compact groups [26], and has been generalized to at least some representations of affine $SO(2,1)$ [27]. The asymptotic behavior can also be derived (for the discrete series) from the character formulas of Henningson et al. [21] and Dixon and Lykken [28], which can be rewritten to give a generating function for $n(N)$ in terms of theta functions.

As argued in the last section, we should now integrate over $\bar{\omega}$ to obtain the total number of macroscopically indistinguishable states. With $n(N)$ given by (1.3) and N determined by (4.6), the dominant contribution will come from

$$\bar{\omega} \sim \frac{\sqrt{2}kr_+}{\ell},$$

---

[*]Normal ordering introduces an ambiguity in L_0, and in string theory the appropriate condition is that $L_0 = 1$. I do not know whether a similar adjustment is needed in (2+1)-dimensional gravity, but a small normal-ordering constant will not qualitatively affect the conclusions of this paper.

and the total number of states will have the asymptotic behavior (1.4), as claimed. We have thus found a good candidate for a set of microscopic states whose statistical mechanics could explain the entropy of the (2+1)-dimensional black hole.

Two cautionary remarks are necessary regarding the representation of the current algebra (4.1) used here. First, the analysis has thus far ignored the degeneracy of the vacuum $|\Omega\rangle$. All unitary irreducible representations of $SO(2,1)\times SO(2,1)$ are infinite-dimensional—one may act on a highest weight state $|\Omega\rangle$ with an arbitrary number of factors $(J_0^1 - iJ_0^2)$ and $(\tilde{J}_0^1 - i\tilde{J}_0^2)$ to obtain new states with the same value of L_0. The states (4.7) have the same degeneracy, and $n(N)$ actually counts global $SO(2,1)\times SO(2,1)$ representations rather than individual vectors in each representation. Second, the representations considered here—like all highest weight representations of affine $SO(2,1)$ [29]—are nonunitary: states such as $J_{-N}^0|\Omega\rangle$ are easily seen to have negative norm.

It is not clear whether either of these issues presents a serious obstacle for a thermodynamic interpretation of our new boundary states. Nonunitary conformal field theories occur elsewhere in physics [30, 31]; since our boundary states are not yet interacting with any external fields, the appearance of negative-norm states need not be fatal. Indeed, one expects the (2+1)-dimensional black hole coupled to matter to be unstable against Hawking radiation, and the appearance of negative-norm states in the uncoupled theory may be seen as a sign of this instability. It would be interesting to find a suitable "Euclidean" continuation of this model in which the gauge group $SO(2,1)\times SO(2,1)$ is replaced by $SO(3)\times SO(3)$; it is likely that such a substitution would simultaneously provide a unitary Hilbert space and remove the infinite vacuum degeneracy.

5. Next Steps

The results of this work strongly suggest that black hole entropy has a natural microscopic, "statistical mechanical" origin. A number of important questions remain, however, both in the (2+1)-dimensional model and in 3+1 dimensions.

In 2+1 dimensions, it is important to understand the boundary WZW model in more detail. In particular, the effect of the Jacobian (3.10) needs further investigation, as does the physical significance of the boundary variable $\bar{\omega}$. $SO(2,1)$ WZW models are not yet well understood—in particular, the proper choice of representation of the affine $SO(2,1)$ algebra, discussed at the end of the last section, is not clear—but a good deal of research on this subject is now in progress. Ultimately, of course, a thermodynamic interpretation will require coupling the horizon degrees of freedom to external fields. This is a difficult problem, since matter couplings remove much of the simplicity of (2+1)-dimensional gravity.

I do not know whether the picture presented here will help to resolve the "information loss" paradox of black hole physics. It is worth noting that $N \sim k^2$ for a Planck-mass black hole in 2+1 dimensions, so there is considerable room

for information to be stored in microscopic states; it is only when $GM \sim 1/k^2 \ll 1$ that the number of microscopic states of the black hole becomes small.

The most important question, of course, is whether the results of this paper can be extended to 3+1 dimensions. In the form presented here, they clearly cannot: (3+1)-dimensional general relativity has no Chern-Simons formulation, and there is no easy way to view the diffeomorphisms as an ordinary group of gauge transformations. Nevertheless, it is plausible that the underlying physical ideas introduced here can be translated into the formalism of standard general relativity. In particular, the boundary term (2.3) has a metric counterpart in the $\mathrm{Tr}\,K$ term that must be added to the Einstein action on a manifold with boundary. This term is invariant only under those transformations that take ∂M to itself, so the metric degrees of freedom that would normally be eliminated by "diffeomorphisms" normal to ∂M ought to become dynamical, providing new physics at the black hole horizon. Whether this description can be made quantitatively correct remains to be seen.

Acknowledgements

References

[1] M. Bañados, C. Teitelboim, and J. Zanelli, *Phys. Rev. Lett.* **69** (1992) 1849; M. Bañados, M. Henneaux, C. Teitelboim, and J. Zanelli, *Phys. Rev.* **D48** (1993) 1506.

[2] L. Susskind, L. Thorlacius, and J. Uglum, *Phys. Rev.* **D48** (1993) 3743.

[3] M. Maggiore, *Phys. Lett.* **B333** (1994) 39; "Horizon quantization and thermodynamics of the 2+1 black hole," Pisa preprint IFUP-TH 43/94 (1994).

[4] A. Achúcarro and P. K. Townsend, *Phys. Lett.* **B180** (1986) 89.

[5] E. Witten, *Nucl. Phys.* **B311** (1989) 46.

[6] G. Moore and N. Seiberg, *Phys. Lett.* **B220** (1989) 422.

[7] S. Elitzur et al., *Nucl. Phys.* **B326** (1989) 108.

[8] K. Huang and S. Weinberg, *Phys. Rev. Lett.* **25** (1970) 895.

[9] S. Carlip and C. Teitelboim, "Aspects of black hole quantum mechanics and thermodynamics in 2+1 dimensions," Institute for Advanced Study preprint IASSNS-HEP-94/34 and Davis preprint UCD-94-13 (1994).

[10] W. Ogura, *Phys. Lett.* **B229** (1989) 61.

[11] S. Carlip, *Nucl. Phys.* **B362** (1991) 111.

[12] S. A. Hayward, *Phys. Rev.* **D49** (1994) 6467.

[13] A. M. Polyakov and P. B. Wiegmann, *Phys. Lett.* **B141** (1984) 223.

[14] S. Förste, "A truly marginal deformation of SL$(2, \mathbb{R})$ in a null direction," Hebrew University preprint RI-8-94 (1994).

[15] D. Gepner and E. Witten, *Nucl. Phys.* **B278** (1986) 493.

[16] V. G. Knizhnik and A. B. Zamolodchikov, *Nucl. Phys.* **B247** (1984) 83.

[17] G. Papadopoulos and B. Spence, *Phys. Lett.* **B295** (1992) 44.

[18] For a nice derivation, see M. Bañados, "Sugawara construction and the relation between diffeomorphisms and gauge transformations in Chern-Simons theory," Imperial College preprint Imperial/TP/93-94/40 (1994).

[19] L. J. Dixon and M. E. Peskin, *Nucl. Phys.* **B325** (1989) 329.

[20] A. Alekseev and S. Shatashvili, *Nucl. Phys.* **B341** (1990) 719.

[21] S. Hwang, *Nucl. Phys.* **B354** (1991) 100.

[22] M. Henningson, S. Hwang, and P. Roberts, *Phys. Lett.* **B267** (1991) 350.

[23] S. Ramgoolam, "New modular Hopf algebras related to rational k $\widehat{sl(2)}$," Yale preprint YCTP-P2-93 (1993).

[24] C. Imbimbo, *Phys. Lett.* **B308** (1993) 51; *Nucl. Phys.* **B384** (1992) 484.

[25] S. Ramanujan and G. H. Hardy, *Proc. London Math. Soc. (ser. 2)* **17** (1918) 75, reprinted in *Collected Papers of Srinivase Ramanujan*, G. H. Hardy et al., editors (Chelsea Publishing Company, NY, 1962).

[26] V. G. Kač and D. H. Peterson, *Adv. in Math.* **53** (1984) 125, section 4.7.

[27] V. G. Kač and M. Wakimoto, *Proc. Natl. Acad. Sci. USA* **85** (1988) 4656.

[28] L. J. Dixon, personal communication.

[29] V. Chari and A. Pressley, *Math. Proc. Camb. Phil. Soc.* **102** (1987) 259.

[30] J. L. Cardy, *Phys. Rev. Lett.* **54** (1985) 1354.

[31] J. D. Cohn, *Phys. Lett.* **B226** (1989) 267.