Thermodynamics of black holes in two (and higher) dimensions

Daniel Grumiller

Center for Theoretical Physics, Massachusetts Institute of Technology,
77 Massachusetts Ave., Cambridge, MA 02139, U.S.A.
E-mail: `grumil@lns.mit.edu`

Robert McNees

High Energy Theory Group, Department of Physics, Brown University
Providence, RI 02912, U.S.A.
E-mail: `mcnees@het.brown.edu`

ABSTRACT: A comprehensive treatment of black hole thermodynamics in two-dimensional dilaton gravity is presented. We derive an improved action for these theories and construct the Euclidean path integral. An essentially unique boundary counterterm renders the improved action finite on-shell, and its variational properties guarantee that the path integral has a well-defined semi-classical limit. We give a detailed discussion of the canonical ensemble described by the Euclidean partition function, and examine various issues related to stability. Numerous examples are provided, including black hole backgrounds that appear in two dimensional solutions of string theory. We show that the Exact String Black Hole is one of the rare cases that admits a consistent thermodynamics without the need for an external thermal reservoir. Our approach can also be applied to certain higher-dimensional black holes, such as Schwarzschild-AdS, Reissner-Nordström, and BTZ.

KEYWORDS: Black Holes, Black Holes in String Theory, 2D Gravity.

Contents

1. Introduction

Two-dimensional gravity provides a useful expedient for testing ideas about quantum gravity in higher dimensions. Technical simplifications in two dimensions often lead to exact results, and it is hoped that this might help to address some of the conceptual problems posed by quantum gravity. At the same time, these simplifications remove many of the features that make gravity interesting. Striking a balance between models that are tractable and models that seem relevant is an art in its own right.

In this paper we study black hole (BH) thermodynamics in two-dimensional models of dilaton gravity. In the absence of matter these models can be solved exactly, but they contain no dynamical degrees of freedom. This seems too restrictive, so we sacrifice integrability and consider models that are coupled to matter. A detailed description of the matter is not necessary for our purposes. We only require the existence of some propagating degrees of freedom, analogous to the radiation emitted by a BH in higher dimensions. We then perform a semi-classical analysis and obtain a comprehensive description of BH thermodynamics for arbitrary models. Despite the fact that we work in two dimensions, these systems exhibit a wide range of interesting thermodynamic phenomena. To some extent this is due to our consideration of 'quasilocal' thermodynamic potentials describing BHs in a finite cavity. Such a treatment reveals aspects of BH thermodynamics that are not visible in an asymptotic analysis. Several universal properties of dilaton gravity BHs are identified, as well as structures that are common to families of models. Our approach allows us to study the thermodynamics of backgrounds, both approximate [1] and exact [2], associated with two-dimensional solutions of string theory. We are also able to extend our analysis to spherically symmetric BHs in higher dimensions. In particular, we obtain an expression that describes the free energy of both asymptotically flat and asymptotically Anti-de Sitter (AdS) BHs.

Dilaton gravity in two dimensions is conventionally described by the (Euclidean) action

$$I = -\frac{1}{16\pi G_2} \int_{\mathcal{M}} d^2x \sqrt{g} \left[X R - U(X) (\nabla X)^2 - 2 V(X) \right] - \frac{1}{8\pi G_2} \int_{\partial\mathcal{M}} dx \sqrt{\gamma} X K . \quad (1.1)$$

We define the dilaton X in terms of its coupling to the Ricci scalar by the expression $X R$. Different models are distinguished by the kinetic and potential functions $U(X)$ and $V(X)$, cf. e.g. [3] for a review. The remaining part of the action is the analog of the Gibbons-Hawking-York boundary term [4, 5], where γ_{ab} is the induced metric on the boundary and

K is the trace of the extrinsic curvature. Our conventions for curvature tensors and other geometric quantities can be found in appendix A. To study the thermodynamics of these theories we employ the Euclidean path integral,

$$\mathcal{Z} = \int \mathscr{D}g \mathscr{D}X \, \exp\left(-\frac{1}{\hbar} I[g, X]\right) \ . \tag{1.2}$$

The path integral is evaluated by imposing boundary conditions on the fields and then performing the weighted sum over all relevant space-times $(\mathcal{M}, g)$ and dilaton configurations X. In the semi-classical limit it is dominated by contributions from stationary points of the action. This can be verified by expanding it around a classical solution

$$I[g_{cl} + \delta g, X_{cl} + \delta X] = I[g_{cl}, X_{cl}] + \delta I[g_{cl}, X_{cl}; \delta g, \delta X] + \frac{1}{2}\delta^2 I[g_{cl}, X_{cl}; \delta g, \delta X] + \dots \tag{1.3}$$

where δI and $\delta^2 I$ are the linear and quadratic terms in the Taylor expansion. The linear term is assumed to vanish on-shell. Then, if the leading term is finite and the quadratic term is positive definite, the saddle point approximation of the path integral is given by

$$\mathcal{Z} \sim \exp\left(-\frac{1}{\hbar} I[g_{cl}, X_{cl}]\right) \int \mathscr{D}\delta g \mathscr{D}\delta X \, \exp\left(-\frac{1}{2\hbar} \delta^2 I[g_{cl}, X_{cl}; \delta g, \delta X]\right) \ . \tag{1.4}$$

This is interpreted as the partition function in the canonical ensemble, with the temperature set by the path integral boundary conditions.

The transition from (1.2) to (1.4) is complicated by the fact that the action (1.1) does not possess any of the properties listed above. First, the on-shell action diverges. This is familiar from calculations in higher dimensions, and is commonly addressed using a technique known as 'background subtraction' [5, 6]. Second, the linear term in (1.3) does not vanish for all field configurations that contribute to the path integral [7, 8]. This is due to boundary terms whose contribution to δI are usually not considered in detail. Schematically, these boundary terms take the form

$$\delta I \Big|_{\text{on-shell}} \sim \int_{\partial \mathcal{M}} dx \sqrt{\gamma} \left[\pi^{ab} \, \delta\gamma_{ab} + \pi_X \, \delta X \right] \neq 0 \ . \tag{1.5}$$

Even when we consider models where the boundary conditions imply $\delta\gamma_{ab} \to 0$ and $\delta X \to 0$ at $\partial \mathcal{M}$, the coefficients of these terms tend to diverge so rapidly that δI does not vanish. Finally, the Gaussian integral in (1.4) is often divergent. In that case the canonical partition function is not well-defined. Instead of describing the thermodynamics of a stable system, (1.4) yields information about decay rates between various field configurations with the specified boundary conditions [9].

We solve the first two problems by replacing (1.1) with an 'improved' action Γ. The improved action is related to (1.1) by a boundary counterterm. The counterterm, which is essentially unique, is obtained by extending the results of [10] to arbitrary models of dilaton gravity. The improved action is finite on-shell, and $\delta\Gamma$ vanishes for any variation of the fields that preserve the path integral boundary conditions. However, it does not address the remaining problem. The quadratic term $\delta^2 \Gamma$ may result in a Gaussian integral,

as in (1.4), that is divergent. This is directly related to the thermodynamic stability of the system. Essentially, the density of states grows so rapidly that the canonical ensemble is not defined. We deal with this in the same manner as York [11]. The BH is placed inside a cavity that couples the system to a thermal reservoir, with boundary conditions fixed at the wall of the cavity. Following Gibbons and Perry [12], we identify the location of the wall with the value of a conserved dilaton charge. The canonical ensemble that emerges from this procedure is well defined if and only if the specific heat of the system is positive. Typically this requires a finite cavity, but there are some notable exceptions. Having resolved the three problems described above, the approximation (1.4), with Γ replacing I, gives the canonical partition function for the system.

We proceed as follows: in section 2 we review BH solutions of two-dimensional dilaton gravity [13, 14], and examine the action (1.1) and its variation evaluated on these backgrounds. We then derive the improved action Γ and verify its properties. The Euclidean path integral is used to define the canonical partition function in section 3, from which all of our results on thermodynamics follow. We obtain expressions for thermodynamic quantities in an arbitrary model of two-dimensional dilaton gravity, establish universal results like the first law of BH thermodynamics, and examine several issues related to stability. In section 4 these results are applied to two significant families of models. Section 5 addresses BHs that appear in models obtained from string theory. This includes a treatment of the Exact String BH [2]. The results obtained in section 3 can also be applied, in certain cases, to BH solutions of higher dimensional theories. This is discussed in section 6, where we study spherically symmetric BHs with or without a cosmological constant, and the toroidal Kaluza-Klein reduction of the BTZ BH [15]. Finally, in section 7 we summarize our results, discuss further applications, and suggest possible generalizations. A collection of appendices contain our conventions, a review of various models, and additional technical results.

Before proceeding to section 2, we point out a few important conventions. Euclidean signature is used throughout this paper. Nevertheless, we frequently use terms such as 'Minkowski' and 'horizon' as if we were working in Lorentzian signature. The dimensionless Newton's constant that appears in the action is set to $8\pi G_2 = 1$. It is restored in certain expressions, when necessary.

2. Semi-classical approximation of the dilaton gravity path integral

In this section we consider the Euclidean path integral formulation of dilaton gravity in two dimensions. Our analysis shows that the standard form of the action is not stationary under the full class of field variations that preserve the boundary conditions. We construct an improved action that solves this problem, and obtain the semi-classical approximation of the corresponding path integral.

2.1 Equations of motion

The equations of motion are obtained by extremizing the action (1.1) with respect to small

variations of the fields that vanish at the boundary. This yields

$$U(X) \nabla_\mu X \nabla_\nu X - \frac{1}{2} g_{\mu\nu} U(X)(\nabla X)^2 - g_{\mu\nu} V(X) + \nabla_\mu \nabla_\nu X - g_{\mu\nu} \nabla^2 X = 0 \quad (2.1)$$

$$R + \partial_X U(X)(\nabla X)^2 + 2\, U(X) \nabla^2 X - 2\, \partial_X V(X) = 0 \;. \quad (2.2)$$

Solutions of these equations always possess at least one Killing vector whose orbits are curves of constant X [16, 17]. We fix the gauge so that the metric is diagonal. The solutions

$$X = X(r), \qquad ds^2 = \xi(r)\, d\tau^2 + \frac{1}{\xi(r)}\, dr^2\,, \quad (2.3)$$

with

$$\partial_r X = e^{-Q(X)} \quad (2.4)$$

$$\xi(X) = w(X)\, e^{Q(X)} \left(1 - \frac{2\, M}{w(X)} \right) \quad (2.5)$$

are expressed in terms of two model-dependent functions,

$$Q(X) := Q_0 + \int^X d\tilde{X}\, U(\tilde{X})\,, \quad (2.6)$$

$$w(X) := w_0 - 2 \int^X d\tilde{X}\, V(\tilde{X})\, e^{Q(\tilde{X})}\,. \quad (2.7)$$

Here Q_0 and w_0 are constants, and the integrals are evaluated at X. Notice that w_0 and the integration constant M contribute to $\xi(X)$ in the same manner. Together, they represent a single parameter that has been partially incorporated into the definition of $w(X)$. By definition they transform as $w_0 \to e^{\Delta Q_0} w_0$ and $M \to e^{\Delta Q_0} M$ under the shift $Q_0 \to Q_0 + \Delta Q_0$. This ensures that the functions (2.4) and (2.5) transform homogeneously, allowing Q_0 to be absorbed into a rescaling of the coordinates. Therefore, the solution is parameterized by a single constant of integration. With an appropriate choice of w_0 we can restrict M to take values in the range $M \geq 0$ for physical solutions. As evident from (2.3) the Killing vector described above is ∂_τ. For brevity we refer to $\sqrt{\xi(X)}$ from (2.5) as the "Killing norm". If it vanishes we encounter a Killing horizon. The square of the Killing norm for the solution with $M = 0$ is

$$\xi_0(X) := w(X)\, e^{Q(X)} \;. \quad (2.8)$$

This solution, which we often refer to as the 'ground state', plays an important role throughout this paper.[1]

2.2 Black holes

The constant of integration that labels solutions of the equations of motion is analogous to the mass parameter that appears in the Schwarzschild metric. This constant has been

---

[1] If the function $V(X)$ has one or more zeroes at finite values of X there is a second, inequivalent family of solutions with constant dilaton. In that case, the identification of the actual ground state of the model becomes more involved. This issue will be addressed in subsection 3.11.

partially absorbed into the definition of $w(X)$ in such a way that $M = 0$ represents the ground state of a particular model. We refer to solutions with $M > 0$, which may exhibit horizons, as BHs.

In the models we consider the metric function $\xi(X)$ is non-negative over a semi-infinite interval

$$X_h \leq X < \infty \ . \tag{2.9}$$

The upper end of this interval corresponds to the asymptotic region of the space-time. The function $w(X)$ generally diverges in this limit

$$\lim_{X \to \infty} w(X) \to \infty \tag{2.10}$$

so that the asymptotic behavior of the metric is typified by the ground state solution (2.8). The lower end of the interval (2.9) is either the origin, or a Killing horizon. Assuming that the function $e^{Q(X)}$ is non-zero for finite values of X, the location of the horizon is related to the parameter M by

$$w(X_h) = 2\,M \ . \tag{2.11}$$

If this condition admits multiple solutions then X_h is taken to be the outermost horizon, so that $w(X) > 2M$ for $X > X_h$.

Regularity of the metric at the horizon fixes the periodicity $\tau \sim \tau + \beta$ of the Euclidean time, which is given by

$$\beta = \left.\frac{4\pi}{\partial_r \xi}\right|_{r_h} = \left.\frac{4\pi}{w'(X)}\right|_{X_h} \ . \tag{2.12}$$

The inverse periodicity is related to the surface gravity of the BH by $2\pi\beta^{-1} = \kappa$. In an asymptotically flat space-time, β^{-1} is also the temperature measured by an observer at infinity. In a slight abuse of notation, we denote this quantity by T

$$T = \beta^{-1} = \left.\frac{w'(X)}{4\,\pi}\right|_{X_h} \ . \tag{2.13}$$

This should not be confused with the proper local temperature, which is related to β^{-1} by a position dependent red-shift (the 'Tolman factor' [18])

$$T(X) = \frac{1}{\sqrt{\xi(X)}}\,\beta^{-1} \ . \tag{2.14}$$

The proper temperature at infinity coincides with (2.13) only if $\xi(X) \to 1$ as $X \to \infty$.

At this point it is instructive to evaluate the action for the BH solution, which we expect is related to the semi-classical approximation of the BH partition function. The bulk terms in (1.1) are straightforward, but we need to specify how the boundary is treated in order to calculate the Gibbons-Hawking-York term. It is apparent from (2.3) that the only component of $\partial\mathcal{M}$ that contributes to this calculation is a surface orthogonal to the coordinate r. Depending on the choice of coordinate system this may occur either at some finite value of r, or in a limit such as $r \to \infty$. In either case, the boundary is assumed to correspond to $X \to \infty$, which means that the boundary terms in the action should be

calculated via a suitable limiting procedure. We implement this by placing a regulator $X \leq X_{\text{reg}}$ on the dilaton, and treating the isosurface $X = X_{\text{reg}}$ as the boundary. Then the on-shell action is

$$I_{\text{reg}} = \beta \left(2\, M - w(X_{\text{reg}}) - 2\,\pi X_h\, T \right) \ . \tag{2.15}$$

The regulator should be removed by taking the limit $X_{\text{reg}} \to \infty$, which recovers the full BH space-time. At this point we encounter a problem: the asymptotic behavior (2.10) of $w(X_{\text{reg}})$ causes the action to diverge. How should this term be addressed? One possibility is that it should be discarded, and the finite remainder then interpreted as the action for the space-time. This strategy is unsatisfactory for two reasons. First, it fails to reproduce the correct thermodynamics of the BH. Second, as we shall demonstrate in the next subsection, the variational properties of the action (1.1) are not consistent with our assumptions about the path integral. Neither issue is resolved by the sort of ad-hoc subtraction described above.

2.3 Variational properties of the action

The equations of motion were described in section 2.1 as the conditions for extremizing the action. In this section we show that this is true for 'localized' variations of the fields, but not for generic variations that preserve the boundary conditions. This has important consequences for the path integral formulation of the theory. We begin by considering small, independent variations of the metric and dilaton. The corresponding change in the action is

$$\delta I = -\frac{1}{2} \int_{\mathcal{M}} d^2x \sqrt{g} \left[\mathcal{E}^{\mu\nu}\, \delta g_{\mu\nu} + \mathcal{E}_X\, \delta X \right] + \int_{\partial\mathcal{M}} dx \sqrt{\gamma} \left[\pi^{ab}\, \delta\gamma_{ab} + \pi_X\, \delta X \right] \ . \tag{2.16}$$

The terms $\mathcal{E}_{\mu\nu}$ and $\mathcal{E}_X$ in the bulk integral are the left-hand-sides of (2.1) and (2.2), respectively. In addition to the bulk term, there is also a boundary term that depends on momenta conjugate to γ_{ab} and X

$$\pi^{ab} = -\frac{1}{2}\, \gamma^{ab}\, n^\mu \nabla_\mu X \ , \qquad \pi_X = U(X)\, n^\mu \nabla_\mu X - K \ . \tag{2.17}$$

These momenta are defined with respect to evolution of the fields along the direction normal to the boundary, indicated here by the vector n^μ. When evaluated on a solution of the classical equations of motion, the change in the action (2.16) should vanish for any variation of the fields that preserves the boundary conditions. A direct calculation, which we shall carry out below, shows that this is not the case. Instead, we find that $\delta I \neq 0$ unless we restrict our attention to localized variations of the fields. In this context a field variation is 'localized' if its asymptotic behavior satisfies some condition that is more restrictive than simply preserving the boundary conditions. Essentially, the change in the action only vanishes for variations that fall off more rapidly at infinity than required by the boundary conditions.

The claim that δI does not vanish for physically reasonable variations of the fields may seem odd. The Gibbons-Hawking-York term in (1.1) is usually motivated by stating that it leads to an action with a "well defined" variational principle. But the Gibbons-Hawking-York term only ensures that fields obey Dirichlet conditions at $\partial\mathcal{M}$. It does

not guarantee that the boundary term in δI vanishes for arbitrary $\delta\gamma_{ab}$ and δX that preserve these boundary conditions. This issue has been studied in higher dimensions for gravitational theories with both asymptotically flat [7, 8] and asymptotically AdS [19, 20] boundary conditions. It is especially important for the path integral formulation of the theory. Given a solution (ξ_{cl}, X_{cl}) of the equations of motion, the path integral also receives contributions from field configurations whose asymptotic behavior is $(\xi_{cl} + \delta\xi, X_{cl} + \delta X)$, where $\delta\xi$ and δX are generic variations consistent with the boundary conditions. The $\hbar \to 0$ limit of the path integral is dominated by solutions of the classical equations of motion if, and only if, δI vanishes for all such variations.

We show now that the action (1.1) does not satisfy this condition. Evaluating (2.16) for a solution of the form (2.3) gives

$$\delta I_{\text{reg}} = \int d\tau \left[-\frac{1}{2} \partial_r X \, \delta\xi + \left(U(X)\, \xi(X)\, \partial_r X - \frac{1}{2}\, \partial_r \xi \right) \delta X \right] \ . \tag{2.18}$$

The subscript on δI indicates that the boundary term must be evaluated using a regulator, as in the previous subsection. The boundary term should vanish as the regulator is removed to infinity. To show that this is not the case, it is enough to set $\delta X = 0$ and consider the variation $\delta\xi$. Using (2.4), the coefficient of $\delta\xi$ is proportional to $e^{-Q(X)}$. Depending on the model, this term may either diverge, approach a constant, or vanish at $\partial\mathcal{M}$. Now we need to specify how a generic variation $\delta\xi$ is allowed to behave at $\partial\mathcal{M}$. The models we consider allow for a wide range of asymptotic behavior. In particular, ξ might diverge in the $X \to \infty$ limit. Thus, we cannot assume that $\delta\xi$ vanishes at $\partial\mathcal{M}$, as we would if ξ approached a constant at $\partial\mathcal{M}$. Instead we must appeal to the general solution (2.5), from which we infer the asymptotic behavior of a $\delta\xi$ that preserves the boundary conditions. Recall that the leading asymptotic behavior of ξ is determined by $\xi_0 = we^Q$, since $w \to \infty$ at $X \to \infty$. We assume that the boundary conditions are preserved by variations whose asymptotic behavior is similar to the sub-leading term in (2.5). This means that we must consider the change in the action due to

$$\delta\xi = \delta M \, e^{Q(X)} \tag{2.19}$$

where δM is an arbitrary infinitesimal constant. The corresponding change in the action is non-zero and independent of the regulator

$$\delta I = \int d\tau \, \delta M \ \neq \ 0 \ . \tag{2.20}$$

This shows that solutions of (2.1) and (2.2) do not extremize the action (1.1) against generic variations $\delta\xi$ that preserve the boundary conditions on ξ. In appendix B we repeat this analysis in terms of a new metric variable $\hat{\xi}$ that obeys the standard Dirichlet condition $\hat{\xi} = 1$ at $\partial\mathcal{M}$. Then, in appendix C, we show that the result (2.20) does not depend on the diagonal gauge used in (2.3).

This explains why one cannot drop the regulator-dependent term in (2.15) and assign a meaning to the part that remains. The semi-classical approximation of the path integral is only well-defined if $\delta I = 0$ for the full class of field variations encountered in the path

integral. With this in mind, we look for an improved action that is both finite and stationary for solutions of the classical equations of motion.

2.4 The Hamilton-Jacobi boundary term

In the previous section we found that the variational properties of the action are not compatible with the semi-classical approximation of the path integral. The same problem was encountered in [10], where the authors studied BH solutions of the tree-level equations for type 0A string theory in two dimensions. In this section we generalize the techniques used in [10] to arbitrary theories of dilaton gravity. The result is a new action Γ that is related to I by

$$I = \Gamma + I_{CT} \ . \tag{2.21}$$

The term I_{CT} is a boundary integral of the form

$$I_{CT} = \int_{\partial \mathcal{M}} dx \, L_{CT}(\gamma, X) \ . \tag{2.22}$$

The integrand may depend on the fields at the boundary, and their derivatives along the boundary, but not their normal derivatives. This ensures that Γ and I yield the same equations of motion. The term I_{CT} is then defined so that solutions of these equations are stationary points of Γ, with respect to any variation of the fields that preserves the boundary conditions. This construction is similar to the boundary counterterm technique used in the AdS/CFT correspondence. Although this approach was originally devised to cancel divergences in the action [21 – 24], it was later shown that the boundary counterterms follow from the requirement that the action admits a well-defined variational problem with asymptotically AdS boundary conditions [19]. A similar modification of the Einstein-Hilbert action was proposed in [8, 25], where the authors considered space-times with asymptotically flat boundary conditions. Other applications of boundary counterterms can be found in [26 – 30].

Following the derivation in [10], the boundary counterterm is identified with a solution of the Hamilton-Jacobi equation for the on-shell action. This equation can be derived by starting with the Hamiltonian constraint that follows from the action (1.1)

$$\mathcal{H} = 2 \, \pi_X \, \gamma_{ab} \, \pi^{ab} + 2 \, U(X) \left(\gamma_{ab} \, \pi^{ab} \right)^2 + V(X) = 0 \ . \tag{2.23}$$

Recall that the momenta, defined in (2.17), appear as boundary terms in the variation of the action. When δI is evaluated for a solution of the equations of motion the bulk term vanishes, leaving only the boundary term. Thus, the momenta can be expressed as functional derivatives of the *on-shell* action

$$\pi^{ab} = \frac{1}{\sqrt{\gamma}} \frac{\delta}{\delta \gamma_{ab}} \left(I \, \Big|_{\mathcal{E}=0} \right) \qquad \pi_X = \frac{1}{\sqrt{\gamma}} \frac{\delta}{\delta X} \left(I \, \Big|_{\mathcal{E}=0} \right) \ . \tag{2.24}$$

Rewriting the Hamiltonian constraint (2.23) in terms of these expressions gives a non-linear functional differential equation for the on-shell action which is difficult to solve. But in our case symmetry arguments reduce the problem to a linear differential equation for a function

of X, as we now demonstrate. The solution I_{CT} should be invariant under diffeomorphisms of $\partial\mathcal{M}$, so we express its integrand as $L_{CT}(\gamma, X) = \sqrt{\gamma}\,\mathcal{L}_{CT}(\gamma, X)$. Because there are no other intrinsic, diffeomorphism covariant scalars constructed from γ in one dimension, it follows that $\mathcal{L}_{CT}$ can depend only on X and its derivatives along the boundary. But terms involving derivatives can be ignored, because the boundary is an isosurface of X. Thus, the scalar density $\mathcal{L}_{CT}$ depends only on X and the boundary counterterm takes the form

$$I_{CT} = \int_{\partial\mathcal{M}} d\tau\,\sqrt{\gamma}\mathcal{L}_{CT}(X)\;. \tag{2.25}$$

The functional derivatives (2.24) applied to this integral give 'momenta' that we denote

$$p^{ab} = \frac{1}{2}\,\gamma^{ab}\,\mathcal{L}_{CT}(X) \qquad p_X = \partial_X\mathcal{L}_{CT}(X) \tag{2.26}$$

to distinguish them from π^{ab} and π_X. Applying these expressions in (2.23), the Hamiltonian constraint becomes a linear differential equation for $\mathcal{L}_{CT}^2$,

$$\frac{1}{2}\,\partial_X\left(\mathcal{L}_{CT}(X)^2\right) + \frac{1}{2}\,U(X)\,\mathcal{L}_{CT}(X)^2 + V(X) = 0\;. \tag{2.27}$$

The general solution of this equation is

$$\mathcal{L}_{CT}(X) = -\sqrt{e^{-Q(X)}\,(w(X) + c)} \tag{2.28}$$

where $Q(X)$ and $w(X)$ were defined in (2.6) and (2.7), and c is an arbitrary constant. Note that the Hamilton-Jacobi equation determines $(\mathcal{L}_{CT})^2$, and the negative root has been chosen in (2.28). With the definition (2.21), this root leads to an action Γ with the correct properties.

In many models there are natural arguments for setting the constant c to zero. For instance, in models associated with two-dimensional string theories the action I is invariant under Buscher duality [31]. The action Γ is expected to possess the same invariances as I, and this is only the case if $c = 0$, as pointed out in [10]. More generally, setting $c = 0$ guarantees that $\mathcal{L}_{CT}$ is independent of the scaling ambiguity associated with the constant Q_0 that appears in (2.6). Therefore, we set $c = 0$ and obtain the boundary counterterm

$$I_{CT} = -\int_{\partial\mathcal{M}} dx\,\sqrt{\gamma}\,\sqrt{w(X)e^{-Q(X)}}\;. \tag{2.29}$$

This leads to the 'improved' action

$$\Gamma = I + \int_{\partial\mathcal{M}} dx\,\sqrt{\gamma}\,\sqrt{w(X)e^{-Q(X)}}\;. \tag{2.30}$$

2.5 Path integral based on the improved action

Based on the analysis of the previous subsections, we conclude that the Euclidean path integral for two-dimensional dilaton gravity should weight field configurations using the exponential of the improved action,

$$\mathcal{Z} = \int \mathcal{D}g\mathcal{D}X\,\exp\left(-\frac{1}{\hbar}\,\Gamma[g, X]\right)\;, \tag{2.31}$$

rather than (1.2). In this section we justify this claim by establishing two important properties of Γ. First we show that solutions of (2.1) and (2.2) are stationary points of (2.30) with respect to variations of the fields that preserve the boundary conditions. This means that the semi-classical approximation of (2.31) is dominated by solutions of the classical equations of motion (2.1) and (2.2). Second, we evaluate the action for the BH space-times discussed in section 2.2 and show that the result is finite.

The variation of the action Γ is

$$\delta\,\Gamma = -\frac{1}{2} \int_{\mathcal{M}} d^2x \,\sqrt{g}\, \left[\, \mathcal{E}^{\mu\nu}\,\delta g_{\mu\nu} + \mathcal{E}_X\,\delta X \, \right] \\ + \int_{\partial\mathcal{M}} dx \,\sqrt{\gamma}\, \left[\, (\pi^{ab} - p^{ab})\,\delta\gamma_{ab} + (\pi_X - p_X)\,\delta X \, \right] \; . \tag{2.32}$$

This differs from (2.16) by boundary terms coming from the variation of (2.29). We can repeat the analysis of section 2.3, evaluating $\delta\Gamma$ on-shell and checking that the boundary term vanishes for appropriate variations of the metric and dilaton. The terms involving π^{ab} and π_X were evaluated in (2.18). Using (2.26), the contributions due to I_{CT} are

$$p^{ab} = -\frac{1}{2}\,\gamma^{ab}\,e^{-Q(X)}\,\sqrt{\xi_0(X)}, \qquad p_X = \frac{1}{2}\,e^{-Q(X)}\,U(X)\,\sqrt{\xi_0(X)} - \frac{1}{2}\,\frac{w'(X)}{\sqrt{\xi_0(X)}} \; . \tag{2.33}$$

As before, the boundary term is evaluated at the cut-off $X = X_{\mathrm{reg}}$

$$\delta\,\Gamma_{\mathrm{reg}} = \frac{1}{2} \int d\tau \left[e^{-Q(X)}\left(\sqrt{\frac{\xi_0(X)}{\xi(X)}} - 1\right)\delta\xi \tag{2.34} \right. \\ \left. + \left(\xi(X)U(X)e^{-Q(X)}\left(1 - \sqrt{\frac{\xi_0(X)}{\xi(X)}}\right) - w'(X)\left(1 - \sqrt{\frac{\xi(X)}{\xi_0(X)}}\right)\right)\delta X \right] \; .$$

This result simplifies a great deal when the regulator is moved far into the asymptotic region of the space-time. The function $w(X)$ becomes very large in this limit, and the coefficients of $\delta\xi$ and δX can be expanded in powers of $1/w$. The leading order terms in this expansion are

$$\delta\,\Gamma_{\mathrm{reg}} = \frac{M}{2} \int d\tau \left[\frac{1}{\xi_0(X)}\,\delta\xi - \left(U(X) + \frac{w'(X)}{w(X)}\right)\delta X \right] + \dots \tag{2.35}$$

The variation of the metric $\delta\xi$ behaves as (2.19), so the first term clearly vanishes in the $X_{\mathrm{reg}} \to \infty$ limit. However, the analysis of (2.35) is much simpler if we use the change of variables $g_{\mu\nu} = e^{2\sigma(X)}\,\hat{g}_{\mu\nu}$ defined in (B.1), with $\sigma = w(X)e^{Q(X)}$ as in (B.10). Then the variation of Γ is given by (2.34) with all un-hatted quantities replaced by their hatted equivalents. Using (B.12), the leading terms in the large-w expansion of $\delta\Gamma$ are

$$\lim_{X_{\mathrm{reg}}\to\infty}\delta\Gamma_{\mathrm{reg}} = \frac{M}{2}\lim_{X_{\mathrm{reg}}\to\infty}\int d\tau \left[\delta\hat{\xi} + M\,\delta\Big(\frac{1}{w}\Big)\right] \; . \tag{2.36}$$

The new variable $\hat{\xi}$ obeys the standard Dirichlet condition $\hat{\xi} = 1$ at the boundary, which means that $\delta\hat{\xi}$ vanishes as $X_{\mathrm{reg}} \to \infty$. The second term must also vanish in this limit since

we assume that $w \to \infty$ asymptotically. Thus we have established

$$\lim_{X_{\text{reg}} \to \infty} \delta\, \Gamma_{\text{reg}} = 0 \ . \tag{2.37}$$

We conclude that solutions of the classical equations of motion are stationary points of Γ with respect to all variations of the fields that preserve the boundary conditions.

Now we evaluate the action for the BH solution. In section 2.2 we found that the regulated on-shell action I_{reg} contains a term that diverges in the $X_{\text{reg}} \to \infty$ limit. Including the contribution from the boundary counterterm, the improved action is

$$\Gamma_{\text{reg}} = \beta \left(w(X_{\text{reg}}) \sqrt{1 - \frac{2M}{w(X_{\text{reg}})}} - w(X_{\text{reg}}) + 2\, M - 2\, \pi X_H\, T \right) \ . \tag{2.38}$$

The regulator can be safely removed to obtain a finite result

$$\lim_{X_{\text{reg}} \to \infty} \Gamma_{\text{reg}} = \beta \left(M - 2\pi X_H\, T \right) \ . \tag{2.39}$$

Notably, this differs by an amount $\beta\, M$ from the part of (2.15) that was independent of the regulator.

3. Thermodynamics in the canonical ensemble

Motivated by York's path integral formulation of gravitational thermodynamics [11], we now introduce an upper bound X_c on the dilaton. This constrains the dilaton to a 'cavity' $X \leq X_c$ whose wall is the dilaton isosurface $X = X_c$. The cavity is in contact with a thermal reservoir that maintains boundary conditions at X_c.[2] The semi-classical approximation of the corresponding path integral yields the partition function in the canonical ensemble. The consistency of this interpretation rests upon the condition of thermodynamic stability, which we discuss in subsections 3.5 and 3.10.

3.1 Dilaton charge and free energy

In standard thermodynamics the canonical ensemble fixes the temperature and volume of a system. The local temperature T_c at the cavity wall is fixed by requiring the proper periodicity of the Euclidean time to be β_c. But the volume of the cavity at constant Euclidean time is not a well-defined quantity, due to the possibility of a horizon $X_h < X_c$. In higher dimensions this issue is addressed by holding the area of the cavity wall fixed, instead [11]. We essentially do the same, but in an indirect way. In two-dimensional dilaton gravity any sufficiently regular function of the dilaton can be used to construct a conserved current. As a result, there are an infinite number of conserved charges associated with the

---

[2]Note that X_c has a clear physical interpretation, unlike the regulator X_{reg} introduced in the previous section. The regulator appeared as an ingredient of the limiting procedure used to evaluate boundary quantities. As such, the limit $X_{\text{reg}} \to \infty$ was always assumed to be the final step of any calculation. This is not the case with X_c, which may remain finite.

dilaton. Following Gibbons and Perry [12], we choose a particular conserved charge that we denote by $D(X)$ and hold this quantity fixed at X_c. One convenient choice is

$$D_c = X_c \,, \tag{3.1}$$

which plays a role similar to the area of the cavity wall in higher dimensions. This choice can be regarded as "natural" as it is precisely the quantity that couples to the Ricci scalar and the extrinsic curvature in the action. Another "natural" possibility for the dilaton charge,

$$\bar{D}_c = e^{-Q_c}\sqrt{\xi_0(X_c)} \,, \tag{3.2}$$

coincides with the solution of the Hamilton-Jacobi equation (2.29). We shall discuss in detail how various thermodynamical properties depend on the choice of dilaton charge.

Fixing D_c and β_c specifies the boundary conditions for the path integral. The classical solutions consistent with these boundary conditions correspond to real, non-negative values of M that satisfy the condition

$$\beta_c = \sqrt{\xi(D_c, M)}\, \beta(M) \,, \tag{3.3}$$

where $\beta(M)$ is the periodicity of the Euclidean time given in (2.12). In general, there may be zero, one, or multiple classical solutions that satisfy (3.3). Issues related to the existence of multiple solutions will be discussed in subsection 3.10. For the next several subsections, we assume that we are working with a single solution. In that case the Helmholtz free energy of the solution is related to the on-shell action Γ_c by

$$F_c(T_c, D_c) = T_c\, \Gamma_c(T_c, D_c) \ . \tag{3.4}$$

From (2.14), (2.38) and (3.4) we obtain

$$F_c = -2\pi\, X_h\, T_c + e^{-Q_c}\left(\sqrt{\xi_0} - \sqrt{\xi_c}\right) \ . \tag{3.5}$$

It is worth mentioning that the free energy for the ground state solution vanishes if either $T = 0$ or $X_h = 0$. This statement does not depend on the location of the cavity wall.

The interpretation of (3.4) as the Helmholtz free energy assumes thermodynamic stability. The consistency of this assumption will be addressed in subsections 3.5 and 3.10. In most models thermodynamic stability requires a *finite* cavity. However, we are often interested in the properties of an isolated BH without an enclosing cavity. This is accomplished by taking the limit $X_c \to \infty$, while simultaneously changing β_c according to (3.3). This limit, if it exists, can be thought of as removing the cavity wall to infinity while adiabatically relaxing the temperature to preserve M. Motivated by curiosity, we shall sometimes apply this limit to the result of a calculation without worrying about whether or not it is consistent with thermodynamic stability (usually it is not). We emphasize that, because the on-shell action is finite in this limit, any singular behavior in F_c as $X_c \to \infty$ is due to the Tolman factor in T_c, as evident from (3.4).

Another issue related to stability concerns the possibility of decay into a field configuration with a smaller free energy. The BH solution is unstable against decay into any field

configuration (that obeys the boundary conditions) with a free energy F_g smaller than F_c from (3.5). We shall discuss this issue in more detail later on. For now, we simply point out that there are usually relevant field configurations with a non-positive free energy, so a minimal condition for stability of the BH solution is $F_c \leq 0$. It implies

$$2M - 2\pi X_h T \leq w_c \left(1 - \sqrt{1 - \frac{2M}{w_c}} \right) \; . \tag{3.6}$$

This condition is always fulfilled for boundary conditions that admit solutions with horizons close to the cavity wall: $w_c = 2M + \epsilon$. As we shall see in later parts of this section, the inequality (3.6) is usually not fulfilled for solutions (3.3) with $2M \ll w_c$.

Keeping in mind the caveats described above, the collection of formulas in appendix D can be applied to the Helmholtz free energy (3.5) to calculate all thermodynamical quantities. We start with the most prominent one.

3.2 Entropy and area

The entropy is given by

$$S_c = - \left. \frac{\partial F_c}{\partial T_c} \right|_{D_c} \; . \tag{3.7}$$

The variation of F_c at fixed D_c can be expressed as

$$dF_c \Big|_{D_c} = -2\pi \, X_h \, dT_c + \left(\frac{1}{\sqrt{\xi_c}} - \frac{T_c}{T} \right) dM \; . \tag{3.8}$$

The last term, which was obtained using (D.1) and (D.3), vanishes by the defining relation (2.14). Thus, we find the well-known result [12, 32, 33, 10, 34]

$$S = 2\pi \, X_h \; . \tag{3.9}$$

The entropy is independent of X_c, and depends only on the value of the dilaton at the horizon. As expected entropy is exclusively a property of the horizon. Furthermore, this result is universal and applies to any model (1.1) regardless of the potentials U and V. Restoring the two-dimensional Newton's constant G_2 the result (3.9) becomes $S = X_h/(4G_2)$. It is useful to introduce the notion of area at this point in order to compare with the standard Bekenstein-Hawking formula. The area of a sphere of radius r embedded in d spatial dimensions is given by $A_d = 2\pi^{d/2} r^{d-1}/\Gamma(d/2)$. In the limit $d \to 1$ the area becomes independent of the radius and the "sphere" consists of two disjoint points. It is suggestive to restrict to a connected component and therefore to associate only one of these two points with the horizon. Thus we define $A_h = A_1/2 = 1$. The effective Newton coupling in a scalar-tensor theory is given by $G_{\text{eff}} = G_2/X$, where X is either approximately constant or evaluated at some characteristic scale. The only scale available here is provided by the horizon, so we define $G_{\text{eff}} = G_2/X_h$. Then (3.9) is equivalent to the Bekenstein-Hawking relation

$$S = \frac{A_h}{4 \, G_{\text{eff}}} \, , \tag{3.10}$$

so that the entropy is one quarter of the horizon area in 'effective Planck units'. In section 6, when we consider the reduction of a spherically symmetric BH from $d + 1 > 2$ dimensions down to two dimensions, we shall find that the conserved charge D_c always corresponds to the higher-dimensional area, in $d + 1$ dimensional Planck units, of the cavity surrounding the BH. This supports our earlier statement about "naturalness" of the choice (3.1).

3.3 Dilaton chemical potential and surface pressure

The chemical potential associated with the dilaton charge (3.1) is given by

$$\psi_c = -\left.\frac{\partial F_c}{\partial D_c}\right|_{T_c} . \tag{3.11}$$

The variation of F_c at fixed T_c can be expressed as

$$dF_c\Big|_{T_c} = -2\pi\, T_c\, dX_h + \frac{1}{2}\, e^{-Q_c} \left(\frac{d\xi_0}{\sqrt{\xi_0}} - \frac{d\xi_c}{\sqrt{\xi_c}}\right) - U_c e^{-Q_c} \left(\sqrt{\xi_0} - \sqrt{\xi_c}\right) dD_c . \tag{3.12}$$

Using (D.1), (D.2) and (D.5) we find

$$\psi_c = -\frac{1}{2}\, U_c\, e^{-Q_c} \left(\sqrt{\xi_c} - \sqrt{\xi_0}\right) + \frac{1}{2}\, w'_c \left(\frac{1}{\sqrt{\xi_c}} - \frac{1}{\sqrt{\xi_0}}\right) . \tag{3.13}$$

We have defined the dilaton chemical potential with a minus sign in (3.11) so that it follows the same conventions as a pressure in standard thermodynamics. This analogy becomes precise in several examples where D_c plays the role of the area of a cavity surrounding the horizon.

The dilaton chemical potential (3.13) exhibits some interesting properties: It diverges with the Tolman factor as the horizon is approached. For the ground state solution $M = 0$ it vanishes. Finally, it transforms covariantly with the dilaton charge $\bar{D}_c(D_c)$, viz., $\bar{\psi}_c = \psi_c\, dD_c/d\bar{D}_c$. With the choice (3.2) the dilaton chemical potential simplifies to

$$\bar{\psi}_c = \eta \sqrt{\frac{\xi_0}{\xi_c}} + (1 - \eta)\sqrt{\frac{\xi_c}{\xi_0}} - 1 \tag{3.14}$$

where

$$\eta := \frac{1}{1 - U_c w_c / w'_c} . \tag{3.15}$$

This definition is particularly useful if η is constant, such as for models with $U = 0$, Minkowskian ground state models and the so-called *ab*-family. We shall discuss them in detail in section 4.

It is sometimes of interest to determine the values of the dilaton charge for which the dilaton chemical potential (3.13) vanishes. This can be thought of as extremizing the free energy with respect to the dilaton charge. The dilaton chemical potential vanishes if either of the following conditions is satisfied

$$\sqrt{\xi_c} = \sqrt{\xi_0} , \qquad \sqrt{\xi_c} = -\frac{w'_c}{w_c U_c}\, \sqrt{\xi_0} . \tag{3.16}$$

The first possibility can only be realized asymptotically. Thus, one natural value of the cut-off is the asymptotic region, which might have been anticipated. The second possibility arises only if $-1 < w_c'/(w_c U_c) < 0$. In the latter case, the system is only stable against changes in D_c if the free energy is a minimum. This is equivalent to positivity of the isothermal compressibility (3.38) discussed below.

3.4 Internal energy and first law

All calculations so far have referred to the canonical ensemble, with the temperature T_c and dilaton charge D_c at the cavity wall held fixed. We can now perform a Legendre transformation to obtain the internal energy E_c as a function of the entropy and the dilaton charge

$$E_c(S, D_c) = F_c(T_c, D_c) + T_c S \ . \tag{3.17}$$

From (3.5) and (3.9) we obtain

$$E_c = e^{-Q(D_c)} \left(\sqrt{\xi_0} - \sqrt{\xi_c} \right) \geq 0 \ . \tag{3.18}$$

The last inequality holds because we assume $M, w_c \geq 0$. The internal energy E_c agrees with the proper energy obtained from the quasi-local stress tensor of Brown and York [35]. This stress tensor is defined as

$$T^{ab} := \frac{2}{\sqrt{\gamma}} \frac{\delta \Gamma}{\delta \gamma_{ab}} \ . \tag{3.19}$$

Using (2.32) obtains

$$T^{ab} = 2 \left(\pi^{ab} - p^{ab} \right) \ . \tag{3.20}$$

Contracting T^{ab} with two copies of the unit-norm Killing vector $u_a = \delta_a^\tau$ gives the proper (Euclidean) energy density. This can be read off from (2.17) and (2.33)

$$T^{ab} u_a u_b = e^{-Q(D_c)} \left(\sqrt{\xi_0} - \sqrt{\xi_c} \right) \ . \tag{3.21}$$

This is precisely the internal energy given in (3.18). Notice that it is not the same as the conserved charge associated with the Killing vector ∂_τ,

$$Q_{\partial_\tau} = \lim_{X_c \to \infty} \sqrt{\xi_c} \, T^{ab} u_a u_b = M \ , \tag{3.22}$$

where $\sqrt{\xi}$ is the lapse function[3] in the 1+1 decomposition of the metric (2.3). The definition of the conserved charge Q_{∂_τ} involves the limit $X_c \to \infty$, which should be thought of as the two-dimensional analog of the limiting procedure used to integrate over a sphere at spatial infinity in higher dimensions. The asymptotic limit of the internal energy is proportional to the conserved charge M, rescaled by the Tolman factor

$$\lim_{w_c \to \infty} E_c = \lim_{w_c \to \infty} \frac{M}{\sqrt{\xi_c}} \ . \tag{3.23}$$

---

[3]Here 'lapse function' refers to the standard temporal evolution, to be distinguished from the lapse function of the radial evolution introduced in appendix C.

At finite values of X_c the result (3.18) can be used to express the conserved quantity M as a function of the internal energy

$$M = \sqrt{\xi_0}\, E_c - \frac{1}{2\,w_c}\,(\sqrt{\xi_0}\, E_c)^2 \;. \tag{3.24}$$

This has a simple interpretation in space-times where the Killing norm $\sqrt{\xi}$ approaches a constant, which we may set to unity without loss of generality. In that case M coincides with the ADM mass, which is related to the internal energy in the region $X \leq X_c$ by

$$M = E_c - \frac{E_c^2}{2w_c}\;. \tag{3.25}$$

The ADM mass contains two contributions: the first one is just the internal energy E_c and the second one is the gravitational binding energy associated with collecting the energy E_c in a region determined by X_c. Consistently, the second term is absent asymptotically.

Using (D.2) and the results for entropy (3.9) and dilaton chemical potential (3.13) one can show that the internal energy obeys the first law of thermodynamics

$$dE_c = T_c dS - \psi_c dD_c \;. \tag{3.26}$$

This quasilocal form of the first law holds regardless of the particular model, the location X_c of the cavity wall, or choice of the dilaton charge. In this regard it contains a great deal more information about the system than the microcanonical expression $dM = TdS$.

3.5 Specific heat at constant dilaton charge and fluctuations

Specific heat at constant dilaton charge is given by

$$C_D = \left.\frac{\partial E_c}{\partial T_c}\right|_{D_c} = T_c \left.\frac{\partial S}{\partial T_c}\right|_{D_c} \;. \tag{3.27}$$

Using (3.18), (D.1) and (D.3) we find

$$C_D = \frac{4\pi\, w_h'(w_c - 2M)}{2w_h''(w_c - 2M) + w_h'^2}\;. \tag{3.28}$$

If $w_h'' \neq 0$ then the asymptotic limit of D_D reduces to the microcanonical specific heat

$$\lim_{w_c \to \infty} C_D = 2\pi\, \frac{w_h'}{w_h''} = -\frac{\left(\frac{dS}{dM}\right)^2}{\frac{d^2S}{dM^2}}\;. \tag{3.29}$$

We have implicitly assumed in the previous subsections that the specific heat is positive. This assumption will be examined in more detail in subsection 3.10, but we can offer a partial justification by considering the near horizon limit of (3.28). If the classical solution determined by (3.3) satisfies $w_c = w_h + \epsilon$, with $0 < \epsilon \ll 1$, then the specific heat simplifies to

$$C_D = \frac{\epsilon}{T}\;. \tag{3.30}$$

Remarkably, it is always positive and vanishes as the cut-off approaches the horizon. This statement is model independent.

If the specific heat is positive and the temperature is large enough then one can take into account the effect of thermal fluctuations of the internal energy. The temperature is 'large enough' if quantum fluctuations are considerably smaller than thermal fluctuations. Quantum fluctuations are typically of the order $\Delta_q E_c \approx 1$, whereas thermal fluctuations, $\Delta_t E_c = \sqrt{\langle E_c^2 \rangle - \langle E_c \rangle^2} = \sqrt{C_D} T_c$, depend on specific heat at constant dilaton charge and on temperature. Thermal fluctuations dominate over quantum fluctuations if $C_D T_c^2 \gg 1$. The canonical entropy, which takes into account thermal fluctuations, is given by the well-known result

$$S_{\text{fluct}} = S + \frac{1}{2} \ln C_D T_c^2 + \dots \tag{3.31}$$

Despite the fact that $T_c \to \infty$ at the horizon, the fluctuations remain finite there

$$\lim_{w_c \to 2M} C_D T_c^2 = e^{-Q_h} T. \tag{3.32}$$

For some models (including the Schwarzschild BH) (3.32) does not satisfy $C_D T_c^2 \gg 1$, even for large BH masses. In these cases thermal fluctuations do not dominate over quantum fluctuations near the horizon.

3.6 Free enthalpy and Gibbs-Duhem violation

With the exception of the dilaton chemical potential, the thermodynamical quantities obtained in the previous subsections do not depend on the specific choice of the dilaton charge. This is not the case for quantities calculated in this and the following subsection. Quantities without a bar on top either refer to the dilaton charge (3.1) or are independent from it; quantities with a bar refer to the dilaton charge (3.2).

We start with the free enthalpy (Gibbs function),

$$G_c(T_c, \psi_c) = F(T_c, D_c) + \psi_c D_c = E_c(S, D_c) - T_c S + \psi_c D_c. \tag{3.33}$$

Even though we have Legendre-transformed the internal energy with respect to the quantities that usually are considered extensive (entropy and "volume", viz., dilaton charge) the result is, in general, not zero. This can be thought of as a 'violation' of the standard form of the Gibbs-Duhem relation.[4] There are, however, a few notable exceptions which we shall discuss in sections 4 and 5. This shows that extensivity properties are not necessarily as in standard thermodynamics. The physical reason underlying these observations is the existence of gravitational binding energy, as in (3.25). Notably, free enthalpy is finite at the horizon, despite the fact that free energy diverges there:

$$G_h = \frac{1}{\sqrt{\xi_0}} \left(2M - TS + \frac{1}{2\pi} MSU_c \right) \bigg|_{D_c = X_h}. \tag{3.34}$$

For comparison it is instructive to calculate

$$\bar{G}_c = \frac{1}{\sqrt{\xi_c}} (2\eta M - TS), \tag{3.35}$$

---

[4]For a discussion of alternative forms of this relation relevant for BHs, see [36].

with η as defined in (3.15). For constant η the quantity $\sqrt{\xi_c}\bar{G}_c$ exclusively depends on properties of the horizon and the Gibbs-Duhem relation holds if $2\eta M = TS$. The result (3.35) is not only considerably simpler than the expression for G_c (which we omitted to present) but it is also divergent at the horizon because of the Tolman factor. This is a clear demonstration of the fact that free enthalpy crucially depends on the choice of the dilaton charge.

The enthalpy

$$\bar{H}_c(S, \psi_c) = E_c(S, D_c) + \bar{\psi}_c \bar{D}_c = \bar{G}_c(T_c, \psi_c) + T_c S \tag{3.36}$$

can be calculated from free enthalpy (3.35) with the result

$$\bar{H}_c = 2\eta \frac{M}{\sqrt{\xi_c}} . \tag{3.37}$$

It is positive if $\eta > 0$ and also scales with the Tolman factor. The quantity H_c is much more complicated so we do not present it here.

3.7 Miscellaneous thermodynamical quantities

The isothermal compressibility

$$\kappa_T = -\frac{1}{D_c} \left. \frac{\partial D_c}{\partial \psi_c} \right|_{T_c} \tag{3.38}$$

requires the calculation of $\partial \psi_c / \partial D_c|_{T_c}$ performed in appendix D, cf. (D.10). Although the dilaton chemical potential diverges at the horizon isothermal compressibility in general remains finite there,

$$\lim_{w_c \to 2M} \kappa_T = \frac{2\pi}{S} \cdot \frac{(2M)^{3/2} e^{Q_h/2}}{w_h'' M - 4\pi^2 T^2 - 4\pi T M U_h - 2M^2 U_h' + M^2 U_h^2} . \tag{3.39}$$

We do not include the result for $\bar{\kappa}_T$ here, but it may be deduced from the results for C_D and $\bar{C}_\psi$.

Specific heat at constant dilaton chemical potential is defined by

$$C_\psi = \left. \frac{\partial H_c}{\partial T_c} \right|_{\psi_c} = T_c \left. \frac{\partial S}{\partial T_c} \right|_{\psi_c} . \tag{3.40}$$

Using (D.11) the difference between the specific heats,

$$C_\psi - C_D = \kappa_T D_c T_c \left(\left. \frac{\partial \psi_c}{\partial T_c} \right|_{D_c} \right)^2 , \tag{3.41}$$

turns out to be positive if and only if isothermal compressibility is positive. The explicit calculation of C_ψ is somewhat lengthy. A simpler result is obtained starting with the dilaton charge (3.2):

$$\bar{C}_\psi = 2\pi \frac{w_h'}{w_h''} . \tag{3.42}$$

This coincides with the asymptotic limit (3.29) and with the microcanonical specific heat. The difference between the specific heats is

$$\bar{C}_\psi - C_D = 2\pi \frac{w'_h}{w''_h} \cdot \frac{w'^2_h}{w'^2_h + 2w''_h(w_c - 2M)} \,. \tag{3.43}$$

For generic thermodynamical systems mechanical stability requires positive isothermal compressibility or, equivalently [cf. (3.41)], $C_\psi > C_D > 0$. However, as we have demonstrated the quantity C_ψ depends crucially on the definition of the dilaton charge. Thus, unless there is a good physical reason to prefer one particular dilaton charge over all other possible choices, we cannot provide an unambiguous answer to the question of mechanical stability. It is intriguing that, for the dilaton charge (3.2), thermodynamic stability implies mechanical stability: $\bar{\kappa}_T > 0$. This is because in a region where both C_D and $\bar{C}_\psi$ are positive their difference $\bar{C}_\psi - C_D$ is also positive, as evident from (3.43).

Additional thermodynamical quantities can be deduced from the ones we have presented here. One can basically exploit the results of standard thermodynamics by identifying the dilaton charge with "volume", and the dilaton chemical potential with "pressure". A few examples will be provided in section 4. We conclude with a supplementary observation regarding an interesting geometric description of the thermodynamics of equilibrium systems due to Ruppeiner [37]. The main idea is to introduce a metric on the thermodynamic state space,

$$ds^2_R := G^R_{IJ} dX^I dX^J \,, \qquad G_{IJ} = -\partial_I \partial_J S(X) \,, \qquad X^I = (E_c, D_c, \dots) \,, \tag{3.44}$$

and to relate its geometric properties to thermodynamic quantities. The entropy S here is expressed as a function of extensive variables X^I, which in our case comprise internal energy (3.18), the dilaton charge (3.1), and any additional extensive quantities associated with generalizations of the metric-dilaton system. This formalism was applied to various BH systems and led to some intriguing results [38]. It is worth mentioning that even in the absence of $U(1)$-charges the Ruppeiner metric is two-dimensional, which differs from the microcanonical analysis in [39], where the Ruppeiner metric was two-dimensional for charged BHs only.

3.8 Charged black holes

Gauge fields in two dimensions have no propagating physical degrees of freedom [40], which makes the generalization of our results to charged BHs straightforward. We consider the addition of a single abelian gauge field, but the extension to several abelian or non-abelian gauge fields is straightforward. The improved action for dilaton gravity coupled to a gauge field is

$$\Gamma = -\frac{1}{2} \int_\mathcal{M} d^2x \sqrt{g} \left[X\,R - U(X)\,(\nabla X)^2 - 2\,V(X) \right] + \int_\mathcal{M} d^2x \sqrt{g}\, f(X)\, F_{\mu\nu} F^{\mu\nu}$$
$$- \int_{\partial\mathcal{M}} dx \sqrt{\gamma}\, X\, K + \int_{\partial\mathcal{M}} dx \sqrt{\gamma}\, e^{-Q(X)} \sqrt{\xi_0(X)} \,. \tag{3.45}$$

The abelian field strength $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ is coupled to the dilaton field via the function $f(X)$. If this function is constant then we say that the gauge field is minimally coupled, and non-minimally coupled otherwise. We emphasize that the gauge field A_μ cannot contribute to the boundary counterterm because we require gauge- and diffeomorphism-invariance. The only gauge-invariant scalar that is linear in A_μ or its derivatives comes from the contraction of the field strength with the ϵ-tensor: $F_{\mu\nu}\epsilon^{\mu\nu}$. However, this term contains a derivative normal to the boundary and is not allowed in a counter-term of the form (2.22). All other gauge-invariant scalars constructed from A_μ, like $F_{\mu\nu}F^{\mu\nu}$, necessarily contain such derivatives and are also ruled out.

The solution of the Maxwell equation $\nabla_\mu(f(X)F^{\mu\nu}) = 0$ is

$$F^{\mu\nu} = \frac{q}{4f(X)}\,\epsilon^{\mu\nu}\,, \tag{3.46}$$

where q is the electric charge, and the factor of $\frac{1}{4}$ has been introduced for convenience. When working in the diagonal gauge (2.3) we use the convention $\epsilon^{\tau r} = -\epsilon^{r\tau} = +1$. The equations of motion for the metric and dilaton can be solved in a manner similar to uncharged BHs. The metric and dilaton are again given by (2.3) and (2.4), but the square of the Killing norm is now

$$\xi(X) = e^{Q(X)}\left(w(X) - 2M + \frac{1}{4}\,q^2\,h(X)\right) \tag{3.47}$$

The 'ground state' is defined as the solution with vanishing mass and charge, $M = q = 0$, leading to the same Killing norm $\sqrt{\xi_0}$ as in (2.8). The function $h(X)$ is defined as

$$h(X) := \int^X d\tilde{X}\,\frac{e^{Q(\tilde{X})}}{f(\tilde{X})}\,, \tag{3.48}$$

where the integral is evaluated at X and the integration constant is absorbed into the definition of w. The result (3.47) could also have been obtained from our earlier solution for the uncharged BH, by replacing $V(X)$ in (2.1) and (2.2) with an effective potential of the form

$$V^{\text{eff}}(X) = V(X) - \frac{q^2}{8f(X)}\,. \tag{3.49}$$

We shall return to this point at the end of this subsection.

To obtain the on-shell action we use

$$\int_{\mathcal{M}} d^2x\sqrt{g}\,fF^{\mu\nu}F_{\mu\nu} = 2\int_{\mathcal{M}} d^2x\sqrt{g}\,\nabla_\mu(fF^{\mu\nu}A_\nu) - 2\int_{\mathcal{M}} d^2x\sqrt{g}\,A_\nu\nabla_\mu(fF^{\mu\nu}) \tag{3.50}$$

and note that the second term vanishes on-shell. The first term, a total derivative, can also be interpreted as one-dimensional Chern-Simons terms with support at both the boundary and the horizon

$$2\int_{\mathcal{M}} d^2x\sqrt{g}\,\nabla_\mu(fF^{\mu\nu}A_\nu) = -\frac{q}{2}\int A_\tau d\tau\Big|_{X_h}^{X_c} = -\frac{q}{2}\left(CS_1(X_c) - CS_1(X_h)\right)$$
$$= -\frac{q}{2}\,\beta\left(A_\tau(X_c) - A_\tau(X_h)\right)\,. \tag{3.51}$$

It is convenient to fix the gauge so that $A_r = 0$. In that case the τ component of the gauge field is given by

$$A_\tau(X) = -\frac{q}{4}\left(h(X) - h(X_h)\right) + A_\tau(X_h).$$

(3.52)

This is just the potential difference between the cavity wall and the horizon, plus an arbitrary constant. The proper electrostatic potential between the cavity wall and the horizon is related to A_τ by

$$\Phi(X) := \frac{A_\tau(X) - A_\tau(X_h)}{\sqrt{\xi(X)}}\ .$$

(3.53)

Note that it contains a Tolman factor, like the proper temperature.

We are now ready to present the generalization of (2.38) to the case of charged BHs,

$$\Gamma_c = \beta_c\left(e^{-Q_c}\left(\sqrt{\xi_0} - \sqrt{\xi_c}\right) - 2\pi X_h T_c - q\Phi_c\right)\ .$$

(3.54)

As pointed out in [10] the logarithm of the partition function,

$$Y_c(T_c, D_c, \Phi_c) = T_c\,\Gamma_c = -2\pi\,X_h\,T_c + e^{-Q_c}\left(\sqrt{\xi_0} - \sqrt{\xi_c}\right) - q\,\Phi_c,$$

(3.55)

is not the Helmholtz free energy. Rather, Y_c is the Legendre transform of F_c with respect to the proper electrostatic potential Φ_c. This is because the Maxwell equation is obtained by varying the action with respect to the gauge field, so it is the gauge field (and not the field strength) that obeys Dirichlet conditions at $\partial\mathcal{M}$. The on-shell action is therefore a function of A_μ evaluated at the boundary. The derivative of Y_c with respect to Φ_c, holding T_c and D_c fixed, gives the conserved electric charge

$$-\left.\frac{\partial Y_c}{\partial \Phi_c}\right|_{D_c, T_c} = q\ .$$

(3.56)

The Legendre transformation

$$F_c(T_c, D_c, q) = Y_c(T_c, D_c, \Phi_c) + q\,\Phi_c$$

(3.57)

leads to the same expressions for the Helmholtz free energy (3.5) and consequently also for entropy (3.9) and internal energy (3.18).

For minimally coupled Maxwell fields the electrostatic potential $\Phi \propto (X_c - X_h)$ is proportional to the separation between the cavity wall and the horizon. If the Killing norm asymptotes to unity then the proper electrostatic potential Φ_c diverges in the $X_c \to \infty$ limit. As a result, the improved action for these models diverges in the limit where the cavity wall is removed from the system. We emphasize that this divergence reflects the physics of the linear electrostatic potential and therefore cannot be removed from the action. The same rule that was mentioned at the end of subsection 2.2 applies here, as well. One does not obtain the correct improved action by simply dropping divergent terms. In the present case, the action depends on the potential difference between the wall and the horizon, which is gauge invariant. Dropping the divergent term from the action leads to a violation of gauge invariance. An example of a charged BH with minimal coupling

is provided by 2D type 0A string theory [10], where $e^{-Q} = w = \lambda X$ and $f = \pi\alpha'$. As an example of non-minimal coupling we mention the Reissner-Nordström BH, $e^{-Q} = w = 2\sqrt{X}$ and $f(X) = X$ (with convenient normalizations for e^Q and f). Our formula (3.48) gives $h = -1/\sqrt{X} = -1/r$ in this case, which is the correct Coulomb-potential. Thus, the typical linear (confining) behavior of the electrostatic potential in two dimensions can be modified by a non-minimal coupling $f(X)$.

Maxwell fields open up an interesting possibility. Consider a theory without Maxwell fields whose potential V takes the form

$$V(X) = \mathcal{V}(X) - \frac{q^2}{8f(X)} \ . \tag{3.58}$$

We can also think of this potential as the result of integrating out a gauge field with coupling function $f(X)$, in a theory with dilaton potential $\mathcal{V}(X)$. We simply re-interpret (3.58) as (3.49). These two points of view will generally lead to different conclusions about the thermodynamics. In particular, the "natural" choice of ground state is different in the two formulations. If we treat the system as dilaton gravity with a Maxwell field then the natural ground state is given by $M = q = 0$. But in the original description, with no Maxwell fields, the ground state is associated with $M = 0$. For instance, the ground state $M = q = 0$ is Minkowski space for the Reissner-Nordström BH, whereas in the alternative description without Maxwell field the natural ground state $M = 0$ is a BPS solution. It is straightforward to generalize these considerations to potentials of the form

$$V = \sum_{n=-\infty}^{\infty} \lambda_n X^{n/a} \,, \tag{3.59}$$

where a is some natural number. This applies to almost all models in the literature. Now we may introduce Maxwell fields to convert all but one of the terms in the generalized Laurent series into Maxwell terms. This leads to a rather simple thermodynamical description and may be useful in various contexts. We shall provide an example in 6.3 when we consider the toroidal reduction of the BTZ BH.

3.9 Conformal properties and microcanonical thermodynamics

In appendix B we discuss some consequences of dilaton dependent Weyl rescalings (B.1). If interpreted as a conformal transformation,[5] rather than as a change of variables, a new action (1.1) is obtained with potentials given by (B.3). We say that these models are conformally related to each other and sometimes refer to them as representing different 'conformal frames' within an equivalence class of models. We consider now the transformation behavior of thermodynamical quantities. To this end it is very helpful to note (B.4) that w is conformally invariant. Therefore, quantities which depend only on w and its derivatives are conformally invariant, but quantities that depend on Q are not.

The conformally invariant thermodynamical quantities comprise the on-shell action Γ_c (2.38), surface gravity T (2.13), mass parameter M (2.11) and entropy S (3.9). It is

---

[5] We use the term 'conformal transformation' in this section, as commonly done in the literature, though (B.1) should really be referred to as a Weyl transformation.

interesting to note that the specific heat at constant dilaton charge C_D (3.28) is conformally invariant, despite the fact that both E_c and T_c depend on the conformal frame. This is because the temperature T_c (2.14), free energy F_c (3.5) and internal energy E_c (3.18) are conformally covariant, with a conformal weight of -1 due to the Tolman factor $1/\sqrt{\xi_c}$. With this convention the metric is defined to have weight $+2$. Quantities that are neither invariant nor covariant are the dilaton chemical potential $\bar{\psi}_c$ (3.14), free enthalpy $\bar{G}_c$ (3.35), enthalpy $\bar{H}_c$ (3.37) and isothermal compressibility κ_T (3.38). These statements are independent from the choice of the dilaton charge. An interesting issue arises for the specific heat at constant dilaton chemical potential: while the result for $\bar{C}_\psi$ (3.42) is conformally invariant, this is not the case for C_ψ with a generic choice of dilaton charge D_c. This provides another rationale for working with the charge $\bar{D}_c$, as defined in (3.2).

As a simplification one can consider microcanonical thermodynamics (or "horizon thermodynamics") without referring to an actual observer [33]. If one is interested only in the microcanonical entropy S, surface gravity T, and their relation to the mass parameter M, then one can use a conformal transformation to set $U = 0$. In this scenario the two relevant formulas, besides the result for entropy (3.9), are the Smarr formula for two-dimensional dilaton gravity,

$$M = \frac{w(S/2\pi)}{2} \tag{3.60}$$

and the microcanonical first law

$$dM = TdS. \tag{3.61}$$

It is important to realize that the microcanonical thermodynamics is not equivalent to canonical thermodynamics in a limit where the dilaton charge approaches some specific value, with rare exceptions. The two pictures are equivalent only if there is a value of the dilaton charge where the dilaton chemical potential vanishes, $\psi_c = 0$, the proper temperature coincides the with Hawking temperature, $T = T_c$, and the canonical ensemble is well-defined. In that case the canonical first law (3.26) coincides with the microcanonical one (3.61), and the mass M coincides with the internal energy E_c. The first two conditions are satisfied in the asymptotic region of models with a Minkowskian ground state. But the third condition is typically violated: in these models the canonical ensemble quite often is not defined in the $X_c \to \infty$ limit. We show now why this is the case and under which conditions the cut-off can be removed.

3.10 Stability and phase transitions

The most important issue related to stability in the canonical ensemble is the effect of thermal fluctuations. If the specific heat at constant dilaton charge is negative, $C_D < 0$, then thermal fluctuations in the internal energy have an imaginary component and the system is unstable. From the point of view of the path integral, $C_D < 0$ implies that the Gaussian integral around a particular extremum of the action does not converge. Fortunately, the assumptions $w_c > w_h$ and $w_h' > 0$ guarantee that there is an open region around the horizon where the specific heat (3.30) is positive. This result is model-independent. Thus, if the horizon is located sufficiently close to the cavity wall then the canonical ensemble is

always well-defined. Of course, a quantitative definition of 'sufficiently close' will depend on the features of the particular model.

In most models a classical solution is only stable when the cavity wall is located at an X_c that is less than some finite, critical value of the dilaton. Consider a BH associated with a solution M of the condition (3.3). We assume that this system has a positive specific heat at constant dilaton charge. Now we want to change the boundary conditions X_c and β_c while holding M fixed. As described at the beginning of section 3, this is accomplished by moving X_c away from the horizon while simultaneously relaxing β_c according to (3.3). The equation (3.28) allows us to make a quantitative statement about the point at which this process destabilizes the system. Suppose that w is a monotonic function. Then a solution $X_{\text{crit}} > X_h$ of

$$w(X_{\text{crit}}) = 2M - \frac{w_h'^2}{2w_h''} \tag{3.62}$$

represents a critical value of the dilaton where the specific heat diverges and changes sign. A well-defined canonical ensemble containing the classical solution exists for any value $X_h < X_c < X_{\text{crit}}$, but not for $X_c > X_{\text{crit}}$. A similar argument can be made for w that are not monotonic. Notice that in either case the specific heat approaches zero at $X_c = X_h$. This reflects the fact that the system is only defined for a cavity wall that is external to the horizon.

In general, the condition (3.3) identifies multiple classical solutions consistent with the boundary conditions [11]. We refer to each solution as a 'branch' of (3.3), denoted by $M_i(X_c, \beta_c)$. The thermodynamic stability of each branch can be determined by looking for a solution $X_{\text{crit}}^i > X_h$ of (3.62). There are three possibilities:

i) There is no solution $X_{\text{crit}}^i > X_h$. This implies that C_D is positive, and the branch is thermodynamically stable.

ii) There is a solution $X_{\text{crit}}^i > X_c > X_h$. Again, the specific heat is positive and the branch is thermodynamically stable.

iii) There is a solution such that $X_c > X_{\text{crit}}^i > X_h$. In this case the specific heat is negative, and the branch is not thermodynamically stable.

In the third case the solution $M_i(X_c, \beta_c)$ may decay into another field configuration with lower free energy via a Hawking-Page phase transition [41]. Because these considerations involve only w the emergence of a phase transition happens in any conformal frame.

We are now able to address the issue of the $X_c \to \infty$ limit in more detail. Given a branch M_i of (3.3) that is thermodynamically stable for some initial values of the boundary conditions, we are interested in the process described by

$$X_c \to \infty \qquad \beta_c \to \lim_{X_c \to \infty} \sqrt{\xi(D_c, M_i)}\, \beta(M_i) \tag{3.63}$$

while holding M_i fixed. This removes the cavity wall, leaving an isolated BH. The process is only defined if β_c remains real and non-negative for all values of X_c. Furthermore, it may be defined on some branches but not others. Finally, even if the process (3.63) is defined

its endpoint may not be thermodynamically stable. The existence of a thermodynamically stable limit (3.63) is far from being a generic feature in the models we consider. For instance, the Schwarzschild BH has $\lim_{X_c \to \infty} C_D < 0$. In models where the endpoint might be stable, the equation (3.62) can be used to group solutions into two sectors. If (3.62) admits a finite solution $X_{\text{crit}} > X_h$ for a particular value of M, then that solution is unstable. But if there is no $X_{\text{crit}} > X_h$ for that value of M then the solution is stable. In higher dimensions the best known example of this phenomenon is the Schwarzschild-AdS solution, where the presence of a cosmological constant stabilizes very massive BHs without the need for an external reservoir. An analogous two-dimensional dilaton gravity model has $w_{SAdS} = 2\sqrt{X}(1 + X/\ell^2)$ with $X = r^2$. For large enough masses, such that $X_h > \ell^2/3$, the quantity w_h'' is positive and the system is stable, while for $X_h < \ell^2/3$ the quantity w_h'' becomes negative and the system is unstable. This can be translated into a critical temperature, where the specific heat diverges. It signals a phase transition [41] between a BH phase for large masses (small temperatures) and a thermal AdS phase for small masses (large temperatures). We shall encounter similar examples in subsection 6.2.

3.11 Tunneling and constant dilaton vacua

In addition to the generic family of solutions (2.3), certain models may admit isolated solutions known as constant dilaton vacua (CDV). Namely, for each solution $X = X_0$ of $V(X) = 0$ there is a CDV with a line element (2.3) whose Killing norm $\sqrt{\xi}$ is given by

$$\xi = c + ar - \frac{1}{2}\lambda r^2. \tag{3.64}$$

Here c and a are constants of motion. The constant λ is determined by

$$R^{CDV} = -\partial_r^2 \xi = \lambda = 2V'(X_0). \tag{3.65}$$

Because they have a different number of constants of motion CDVs and the generic solutions belong to different 'superselection sectors'. Therefore there is no perturbative way in which a generic solution could decay into a CDV or vice versa. This is consistent with the result [42] that fully supersymmetric solutions necessarily are CDVs, while half BPS solutions can never be CDVs. If a supersymmetric CDV exists then the natural ground state solution in the generic sector is a half BPS solution, and both kinds of solutions are stable. However, once matter is switched on there could be non-perturbative interactions which allow tunneling from one sector into the other. We would like to determine whether such a tunneling between a generic solution and a CDV is possible, and if so, which of the two configurations is favored.

We immediately encounter an obstacle: CDVs are isolated in the sense that for an open interval of boundary conditions, $X_c \in (X_<, X_>)$, there is only a countable number of values of the dilaton consistent with the equations of motion. Typically this number is either zero or of order unity. Thus, at first glance CDVs are irrelevant because they seem to be a set of measure zero in the partition function (1.2). Only for infinite finetuning, $X_c = X_0$, there may be a contribution from CDVs. In order to decide whether or not CDVs may be disregarded we consider an open interval of boundary conditions for the dilaton and

the metric. This interval may be arbitrarily small, but it should include exactly one value $X_c = X_0$ that leads to a CDV.[6] In order to get the number of classical solutions determined by this set of boundary data we take the product of the subset of initial conditions consistent with the classical equations of motion times the space of continuous constants of motion that may be adjusted freely. We call the dimension of this space "weight" and want to compare the weight of generic solutions with the weight of CDVs. Let us work in fixed physical units. Then, for generic solutions (2.3) there is only a discrete set of values for M consistent with the boundary conditions on X_c and ξ_c. So the weight is 2 (from the set of boundary conditions consistent with the equations of motion) plus 0 (from the space of continuous constants of motion that may be adjusted freely) and thus equals to 2. For CDVs the space of allowed boundary conditions is just one-dimensional since the dilaton is fixed to a certain value, but now there are two constants of motion, c and a. Therefore, the space of constants of motion that may be adjusted freely is one-dimensional. We conclude that also for CDVs the weight is 2 and they should not be regarded as a set of measure zero. Below we use again sharp boundary conditions and assume $X_c = X_0$ for simplicity.

Now that we have convinced ourselves that CDVs are not a set of measure zero we address the question whether tunneling from a generic solution into a CDV is a favorable process or not. To this end we compare the generic result for the 'improved' on-shell action (2.38) with the on-shell action for a CDV,

$$\Gamma^{CDV} = -2\pi X_0 \chi + \beta \sqrt{\xi} e^{-Q(X_0)/2} \sqrt{w(X_0)} \,. \tag{3.66}$$

The first term in (3.66) contains the Euler-characteristic χ of the manifold $\mathcal{M}$. The last term is the universal counterterm (2.29) evaluated at $X = X_0$. It has the same structure and, for the same set of boundary data, also the same magnitude as the corresponding term in the 'improved' on-shell action for a generic solution. The periodicity in Euclidean time β is well-defined and positive provided the inequality $a^2 + 2c\lambda > 0$ holds, which we assume henceforth. Minkowski or AdS vacua require a non-vanishing Rindler acceleration a while de Sitter solutions ($c\lambda > 0$) are compatible with $a = 0$. We conclude that tunneling from a generic solution into a CDV is favored if the inequality

$$X_0 \chi - X_h > \frac{2w(X_0) - 2w(X_h)}{w'_h} \tag{3.67}$$

holds. Here $X_h < X_0$ is the dilaton evaluated at the horizon of the generic solution. Since the right hand side is non-negative the Euler characteristic must be positive for this inequality to hold. Therefore we restrict the remaining discussion to the topologies of either a sphere ($\chi = 2$) or a disk ($\chi = 1$). The latter case is consistent with a BH interpretation, while the former case may arise for de Sitter. As an example for a model with enhanced tunneling to its CDV solution we mention $w(X) = c^3 + (X - c)^3$. The CDV solution solution for this model is $X_0 = c > 0$, and the inequality (3.67) holds for any value of M consistent with $X_0 > X_h$. For generic models there are two interesting limits. If X_0 is

---

[6]In fact, in any experimental situation this smearing of boundary data is physically more reasonable than an infinite finetuning.

close to the horizon, $X_0 = X_h + \epsilon$, then for $\chi = 2$ the inequality (3.67) holds and tunneling to a CDV is enhanced, whereas for $\chi = 1$ tunneling is suppressed. If $X_0 \gg 1$ then (3.67) simplifies to $X_0\chi > 2w(X_0)/w'_h + \ldots$ Since we demand that $w \to \infty$ as $X \to \infty$ this inequality establishes a dichotomy in the space of dilaton gravity theories: if the function w diverges asymptotically faster (slower) than linearly tunneling into a CDV is suppressed (enhanced), regardless of whether $\chi = 1$ or $\chi = 2$.

In summary, a transition between generic solutions and CDVs is possible only through tunneling. If the inequality (3.67) holds tunneling from a generic solution into a CDV is enhanced, otherwise it is suppressed. We remark that such a tunneling is far from being a generic feature: first of all, many models do not exhibit CDVs, and second, many of the models that do allow for CDV solutions violate the inequality (3.67) for all possible values of $X_h < X_0$. In fact, none of the models that we are going to discuss explicitly in sections 4.2, 5 and 6 allows for an enhanced tunneling into a CDV.

4. Two dimensional examples

4.1 Minkowski ground state models

If the relation

$$e^Q w = 1 \tag{4.1}$$

holds then $\xi_0 = 1$ and the ground state solution (2.8) is (upon Wick rotation back to Lorentzian signature) Minkowski space-time. These models are of particular interest as they contain asymptotically flat solutions like the Schwarzschild and Witten BHs. As we shall see (4.1) leads to considerable simplifications.

The free energy

$$F_c^{MGS} = E_c^{MGS} - \frac{ST}{\sqrt{1 - \frac{2M}{w_c}}} \tag{4.2}$$

and internal energy

$$E_c^{MGS} = w_c \left(1 - \sqrt{1 - \frac{2M}{w_c}} \right) \tag{4.3}$$

have a finite limit if the cut-off is removed, $\lim_{w_c \to \infty} F_c^{MGS} = M - TS$ and $\lim_{w_c \to \infty} E_c^{MGS} = M$. The asymptotic version of the first law (3.26) simplifies to the microcanonical $dM = TdS$. The chemical potential conjugate to the dilaton charge $\bar{D}_c^{MGS} = w_c$ (3.2),

$$\bar{\psi}_c^{MGS} = \frac{1}{2} \left(1 - \frac{2M}{w_c} \right)^{1/2} + \frac{1}{2} \left(1 - \frac{2M}{w_c} \right)^{-1/2} - 1 \geq 0 \,, \tag{4.4}$$

is always non-negative and vanishes for the ground state solution only. It is useful to notice that the quantity defined in (3.15) becomes a universal constant, $\eta = 1/2$. For the formulation based upon the dilaton charge (3.1) we get $\psi_c^{MGS} = \bar{\psi}_c^{MGS} w'_c$, which is also

non-negative provided $w_c' \geq 0$. Asymptotically (4.4) implies $\lim_{w_c \to \infty} \bar{\psi}_c^{MGS} = M^2/(2w_c^2)$. From (3.35) we obtain the free enthalpy

$$\bar{G}_c^{MGS} = \frac{M - TS}{\sqrt{1 - \frac{2M}{w_c}}} \,. \tag{4.5}$$

The relevant quantities for isothermal compressibility simplify to $f_c^{MGS} = w_c''$, $g_c^{MGS} = -w_c''/2 - w_c'N_c/2$, $h_c^{MGS} = -w_c''/2 + w_c'N_c/2$ and $N_c = (2Mw_c'w_h'')/[w_c(w_h'^2 + 2w_h''(w_c - 2M))]$. Evaluation at the horizon yields $\lim_{w_c \to 2M} \kappa_T^{MGS} = 2\pi/(Sw_h'')$. In the alternative formulation the result is

$$\bar{\kappa_T}^{MGS} = w_c \sqrt{1 - \frac{2M}{w_c}} \cdot \frac{2w_c - 4M + w_h'^2/w_h''}{4M^2} \,. \tag{4.6}$$

The zeros $X_c > X_h$ of $\bar{\kappa_T}^{MGS}$ coincide with the poles of C_D^{MGS} (3.28). We do not investigate these models in more detail here but shall encounter them again in this and the next section.

4.2 Models in the a-b family

The so-called ab-family of models [43],

$$U = -\frac{a}{X} \,, \qquad V = -\frac{B}{2}X^{a+b} \,. \tag{4.7}$$

includes the dimensional reduction of spherically symmetric BHs in $d + 1$ space-time dimensions $[a = (d-2)/(d-1), b = -1/(d-1)]$, the Witten BH $(a = 1, b = 0)$ [1], the Jackiw-Teitelboim (JT) model $(a = 0, b = 1)$ [44], and many others. In particular, the Schwarzschild BH arises for $a = -b = 1/2$. Minkowskian ground state models obey the relation $a = 1 + b$. Models with $a = 1 - b$ have an (A)dS ground state and models with $b = 0$ a Rindler ground state. Even more general models often are approximated quite well by the monomial potentials of the ab family. Therefore our discussion here will reveal some generic features of the thermodynamics of dilaton gravity in two dimensions. For the sake of definiteness, and to comply with our working assumption (2.10), we assume $b > -1$ and $B > 0$. In addition, we reduce clutter by imposing $B = b + 1$, without loss of generality. We fix the constants in (2.6) and (2.7) so that

$$w = X^{b+1} \,, \qquad e^Q = X^{-a} \,. \tag{4.8}$$

With these assumptions the BH horizon is located at

$$X_h = (2M)^{1/(b+1)} \,. \tag{4.9}$$

Note that a negative value of M would result in a naked singularity at $X = 0$ if $a < 2$. Surface gravity yields

$$T = \frac{b+1}{4\pi}(2M)^{b/(b+1)} \,, \tag{4.10}$$

which leads to a mass-to-temperature relation $M \propto T^{1+1/b}$. The Killing norm squared and asymptotic Killing norm squared are given by

$$\xi_c = \xi_0 - 2MX_c^{-a} \,, \qquad \xi_0 = X_c^{b-a+1} \geq \xi_c \,. \tag{4.11}$$

If we demand $\xi_0 \neq 0$ for $X_c \to \infty$ the inequality $b - a + 1 \geq 0$ must be met.

4.2.1 Thermodynamics of the a-b family

Entropy (3.9) reads

$$S = 2\pi(2M)^{1/(b+1)} \,. \tag{4.12}$$

The free energy (3.5) simplifies to

$$F_c = X_c^a(\sqrt{\xi_0} - \sqrt{\xi_c}) - \frac{(b+1)M}{\sqrt{\xi_c}} \,. \tag{4.13}$$

For $b < 0$ free energy changes its sign at a particular value of the cut-off,

$$\sqrt{\xi_c} = \frac{1+b}{1-b}\sqrt{\xi_0} \,, \tag{4.14}$$

and the asymptotic free energy is positive. This equation always has exactly one solution under the assumptions $-1 < b < 0$ and $M > 0$. For Schwarzschild this happens at $\sqrt{X} = 9M/4$. With the dilaton charge $\bar{D}_c = X_c^{(a+b+1)/2}$ the dilaton chemical potential (3.14) simplifies to

$$\bar{\psi}_c = \frac{b+1}{a+b+1}\sqrt{\frac{\xi_0}{\xi_c}} + \frac{a}{a+b+1}\sqrt{\frac{\xi_c}{\xi_0}} - 1 \,. \tag{4.15}$$

It vanishes at a particular value of the dilaton given by

$$\sqrt{\xi_c} = \frac{1+b}{a}\sqrt{\xi_0} \,. \tag{4.16}$$

This equation has a solution if $0 < b + 1 < a$. For Minkowskian ground state models the dilaton chemical potential vanishes for $X_c \to \infty$ only. For (A)dS ground state models the dilaton chemical potential vanishes precisely at the value of the cut-off that leads to vanishing free energy. Internal energy (3.18) is given by

$$E_c = X_c^{(a+b+1)/2}\left(1 - \sqrt{1 - 2MX_c^{-1-b}}\right) \,. \tag{4.17}$$

The specific heat at constant dilaton charge (3.28) simplifies to

$$C_D = S\,\frac{X_c^{b+1} - 2M}{bX_c^{b+1} - (b-1)M} \,. \tag{4.18}$$

It changes sign at X_{crit} if the inequality

$$X_{\mathrm{crit}} = \left(M\frac{b-1}{b}\right)^{1/(1+b)} > X_h \tag{4.19}$$

holds, which is possible for $-1 < b < 0$ only. Thermal fluctuations near the horizon are determined by

$$\lim_{w_c \to 2M} C_D T_c^2 \propto M^{(a+b)/(b+1)} \,, \tag{4.20}$$

with a proportionality constant of order of unity. If $b > -a$ ($b < -a$) thermal (quantum) fluctuations dominate for large masses. If $a + b = 0$ both kinds of fluctuations are relevant.

The only Minkowskian ground state model with this property is the Schwarzschild BH. Specific heat at constant dilaton chemical potential (3.42),

$$\bar{C}_\psi = \frac{S}{b} \,, \tag{4.21}$$

is positive for $b > 0$ and coincides with the $X_c \to \infty$ limit of (4.18). The free enthalpy

$$\bar{G}_c = (1 - a - b) \frac{b+1}{a+b+1} \cdot \frac{M}{\sqrt{\xi_c}} \tag{4.22}$$

obeys the Gibbs-Duehm relation $\bar{G}_c = 0$ for (A)dS ground state models only. To get the isothermal compressibility (3.38) we calculate (D.6): $f_c = \frac{1}{4}(a+b+1)(a+b-1)X_c^{(a+b-3)/2}$, $g_c = \frac{a}{2}X_c^{a-2}(1-a-X_cN_c)$, and $h_c = \frac{1}{2}(b+1)X_c^{b-1}(-b+X_cN_c)$, with

$$N_c = -\frac{b}{2X_c} \frac{2aM - (a-b-1)X_c^{1+b}}{(b-1)M - bX_c^{1+b}} \,. \tag{4.23}$$

For any model without (A)dS ground state ($a \neq 1 - b$) the poles of N_c and C_D coincide with each other. For (A)dS ground state models $N_c = b/X_c$ and therefore κ_T vanishes.

For Minkowskian ground state models we can discuss the life-time of a BH without specifying explicitly the degrees of freedom that carry the Hawking quanta. Let us assume that the two-dimensional Stefan-Boltzmann law, asymptotic flux $\propto T^2$, is valid. Then the mass loss per time is determined by $dM/dt \propto -T^2$. The proportionality constant depends on the matter content and on the physical units. With (4.10) it is straightforward to calculate the time scale Δt for the evaporation process

$$\Delta t \propto M^{(1-b)/(1+b)} \left[= M^{d/(d-2)} \right] \,. \tag{4.24}$$

The second expression in (4.24) refers to the Schwarzschild BH in $d + 1$ space-time dimensions. The result (4.24) provides a good estimate of the time-scale associated with BH evaporation, as long as the semi-classical approximation remains valid throughout most of the process. Typically, this requires that the BH initially have large M.

4.2.2 Equation of state and scaling properties

For some purposes it is useful to extract an equation of state $\bar{\psi}_c(T_c, \bar{D}_c)$ from the formulas above. It can be derived from (4.15) upon expressing the right hand side as a function of $\bar{D}_c$ and T_c. To this end one has to find M or T as a function of $\bar{D}_c$ and T_c using (4.10), (4.11) and the defining relation (2.14). As we shall demonstrate, for given T_c and $\bar{D}_c$ there is in general an ambiguity, i.e., the BH mass M or, equivalently, its Hawking temperature T does not follow uniquely from specifying these quantities [cf. the discussion around (3.3)]. A simple exception is provided by Rindler ground state models, where T from (4.10) not only is unique but actually is constant, thus establishing

$$\text{Rindler}: \qquad \bar{\psi}_c(T_c, \bar{D}_c) = \frac{1}{a+1} 4\pi T_c \, \bar{D}_c^{(1-a)/(1+a)} + \frac{a}{a+1} \frac{1}{4\pi T_c} \bar{D}_c^{(a-1)/(a+1)} - 1 \,. \tag{4.25}$$

The dilaton chemical potential becomes minimal at $T_c = \sqrt{a}\bar{D}_c^{(a-1)/(a+1)}/(4\pi)$. For the Witten BH $a = 1$ the dilaton chemical potential depends on T_c only and attains its minimum in the asymptotic region. Also models with $b = 1$ lead to a unique T and to a simple equation of state,

$$\bar{\psi}_c(T_c, \bar{D}_c) = \frac{2}{a+2}\left(1 + 4\pi^2 T_c^2 \bar{D}_c^{-2a/(a+2)}\right)^{1/2} + \frac{a}{a+2}\left(1 + 4\pi^2 T_c^2 \bar{D}_c^{-2a/(a+2)}\right)^{-1/2} - 1\,.$$
(4.26)

For the JT model $a = 0$ the dilaton chemical potential becomes independent from the dilaton charge. The same happens for all (A)dS ground state models

$$(A)dS:\qquad \bar{\psi}_c(T_c) = \frac{1+b}{2}\frac{T_c}{k(T_c)} + \frac{1-b}{2}\frac{k(T_c)}{T_c} - 1\,.$$
(4.27)

However, the function $k(T_c)$ in general has more than one branch and is known implicitly only. For all other models one has to solve

$$T = T_c\,\bar{D}_c^{(1+b-a)/(1+b+a)}\sqrt{1 - \bar{D}_c^{-2(1+b)/(1+b+a)}\left(\frac{4\pi T}{1+b}\right)^{1+1/b}}$$
(4.28)

for T in terms of T_c and $\bar{D}_c$. The solution of (4.28) is not unique in general. For instance, the Schwarzschild BH in 5 space-time dimensions, $a = -2b = 2/3$, leads to two branches $T_\pm = T_c\left(1 \pm \sqrt{1 - 144\pi^2/(T_c^4\bar{D}_c)}\right)^{1/2}/\sqrt{2}$. If we restrict the discussion to an equation of state for the asymptotic region $\bar{D}_c \to \infty$ it suffices to solve (4.28) perturbatively to leading order, in which case there is only one branch. For Minkowski ground state models the result is

$$\text{Minkowski ground state}:\qquad \lim_{\bar{D}_c\to\infty}\bar{\psi}_c(T_c, \bar{D}_c) \propto \frac{T_c^{2+2/b}}{\bar{D}_c^2}$$
(4.29)

while for all other models it reads

$$\text{Others}:\qquad \lim_{\bar{D}_c\to\infty}\bar{\psi}_c(T_c, \bar{D}_c) \propto T_c^{1+1/b}\,\bar{D}_c^{-(1+b)(a+b-1)/(b(a+b+1))}\,.$$
(4.30)

The proportionality constants in these equations are numerical coefficients depending on a, b. Asymptotically the equation of state for an ideal gas, $\bar{\psi}_c \propto T_c/\bar{D}_c$, arises in the limit $b \to \infty$, only. As a cautionary remark we recall that the asymptotic limit is not accessible for $b < 0$ because the asymptotic specific heat (4.21) is negative. By contrast, the near horizon approximation is always accessible. It corresponds to the limit $T_c \to \infty$ and allows to derive a simple equation of state for the branch where T from (4.28) is finite at the horizon,

$$\text{Near horizon}:\qquad \bar{\psi}_c(T_c, \bar{D}_c) \approx \frac{4\pi}{a+b+1}T_c\,\bar{D}_c^{(1-a-b)/(a+b+1)}\,.$$
(4.31)

Extensivity properties of BHs are quite unusual; generically, the temperature is not intensive and depending on the choice of the dilaton charge, which is extensive by definition, either entropy is non-extensive or internal energy or both. For (3.1) the dilaton field $D_c = X_c$ is extensive by definition and thus also entropy is extensive. However, generically no other quantity is extensive or intensive. For (3.2) all thermodynamical potentials are

Quantity	Witten	JT	(A)dS	MGS	Rindler	$a = 0$	generic
$\bar{D}_c, E_c, F_c, \bar{G}_c, \bar{H}_c$	ext.	ext.	ext.	ext.	ext.	ext.	ext.
$\bar{\psi}_c, \bar{\kappa}_T$	int.	int.	int.	int.	int.	int.	int.
$S, C_D, \bar{C}_\psi$	ext.	ext.	ext.	—	—	—	—
T_c	int.	int.	int.	—	—	—	—
$\sqrt{\xi_c}$	int.	ext.	—	int.	—	ext.	—
T	int.	ext.	—	—	int.	—	—
M	ext.	—	—	ext.	—	—	—

Table 1: Extensitivity properties based upon the dilaton charge (3.2)

	Conf.	Scale	Shift	Ext.	Tol.	G_2^{-1}	Sign	Sign$_{GS}$	D_c-dep.
e^Q	**2**	**1**	0	—	0	0	+	+	no
w	0	**1**	**-1**	—	0	0	+	+	no
$\bar{D}_c$	-1	0	—	**1**	0	1	+	+	**yes**
ξ_c	2	2	0	—	**-2**	0	+	+	no
Γ_c	0	0	—	—	0	**1**	?	0	no
M	0	1	1	—	0	0	**+**	**0**	no
T	0	1	0	—	0	0	+	0	no
T_c	-1	0	0	—	1	0	+	0	no
F_c	-1	0	—	1	1	1	?	0	no
S	0	0	0	—	0	1	+	0	no
E_c	-1	0	—	1	0	1	+	0	no
C_D	0	0	0	—	-2	1	?	0	no
$T_c S$	-1	0	0	1	1	1	+	0	no
M/w	0	0	—	0	0	0	+	0	no
$\bar{\psi}_c$	?	0	—	0	1	0	?	0	yes
$\bar{\kappa}_T$	?	0	—	0	-1	0	?	?	yes
$\bar{G}_c$	?	0	—	1	1	1	?	0	yes
$\bar{H}_c$	?	0	—	1	1	1	?	0	yes
$\bar{C}_\psi$	0	0	0	—	0	1	?	0	yes
$\bar{C}_\psi - C_D$	0	0	0	—	0	1	?	0	yes

Table 2: Summary of scaling properties

extensive and the dilaton chemical potential becomes intensive. However, generically entropy is not extensive. We summarize the extensitivity properties with respect to (3.2) of all *ab*-models in table 1. Extensive (intensive) quantities are denoted by "ext." ("int."), while quantities which are neither are denoted by "—".

Besides extensitivity properties there are additional interesting scaling properties of thermodynamical quantities, which are summarized for generic models in table 2. Here is an explanation of all abbreviations: "Conf." refers to the conformal weight (0 means the quantity is conformally invariant and "?" means it transforms inhomogeneously); "Scale"

refers to the scale ambiguity $e^Q \to c e^Q$, $w \to cw$ inherent to the definitions of these functions (0 means again that the quantity is independent from such rescalings); similarly, "Shift" refers to the shift ambiguity $w \to w + w_0$ (quantities which have no well-defined shift-weight are denoted by "—"); "Ext." denotes extensivity properties (1 means extensive and 0 means intensive); "Tol." refers to the scaling with powers of the Tolman factor $1/\sqrt{\xi_c}$ close to the horizon (if it is positive the corresponding quantity diverges at the horizon and if the entry is underlined the whole cut-off dependence comes from a single Tolman factor with corresponding weight); "G_2^{-1}" determines the scaling with the gravitational coupling constant in front of the action; "Sign" is $+$ if the quantity is positive everywhere outside the horizon and "?" if the quantity may change its sign; "Sign$_{GS}$" is the same as "Sign" but for the ground state solution (therefore many quantities vanish); the last entry just keeps track of quantities which depend in an essential way on the definition of the dilaton charge. All scalings act multiplicatively $[s(A) \cdot s(B) = s(AB)]$ except for "Shift" which acts additively $[s(A) + s(B) = s(A + B)]$. The bold entries in this table are definitions rather than results; e.g. the quantity e^Q is the conformal factor in front of the metric and thus by definition has a conformal weight of $+2$.

Finally, we collect various restrictions on the parameters from thermodynamical considerations. We recall that we have assumed $B > 0$ for positive w, $M \geq 0$ for non-negative mass and $b > -1$ for a well-defined (positive) temperature. If we want space-time to be regular in the asymptotic region we encounter the inequalities $a \leq 2$ and $a + b \leq 1$. If we demand that the dilaton chemical potential be non-negative everywhere then we need $0 \leq a \leq 1 + b$. Finally, if we wish the specific heat to be asymptotically positive the inequality $b \geq 0$ must hold. These results are summarized in figure 1. The red (dark gray) region is excluded because $b < -1$ there, thus violating (2.10). The orange (medium gray) region exhibits a curvature singularity in the "asymptotic" region $X_c \to \infty$. In the green (light gray) region the dilaton chemical potential can become negative. In the region below the a-axis the specific heat is asymptotically negative. The white region above the a-axis is free from thermodynamical pathologies regardless of the location of the cavity. The points S (W) [JT] correspond to the Schwarzschild BH (Witten BH) [JT model]. The dashed line denotes (A)dS ground state models and the dotted line Minkowskian ground state models. The a-axis contains all Rindler ground state models.

Two particular models of this class deserve further study. The Schwarzschild BH will be addressed in more detail in section 6 and the Witten BH will be studied in the next section.

5. Examples from string theory

In this section we consider two-dimensional BHs that emerge as solutions, either approximate or exact, of string theory. In [1] it was shown that the Euclidean BHs studied in [45] admit an exact CFT description in terms of an $SL(2, \mathbb{R})/U(1)$ gauged WZW model. In the large k limit, where k is the level of the $SL(2)$ current algebra, the corresponding background takes the form

$$ds^2 = \alpha' k \left(dy^2 + \tanh^2 y \, d\theta^2 \right) \qquad X = X_0 \cosh^2 y \ . \tag{5.1}$$
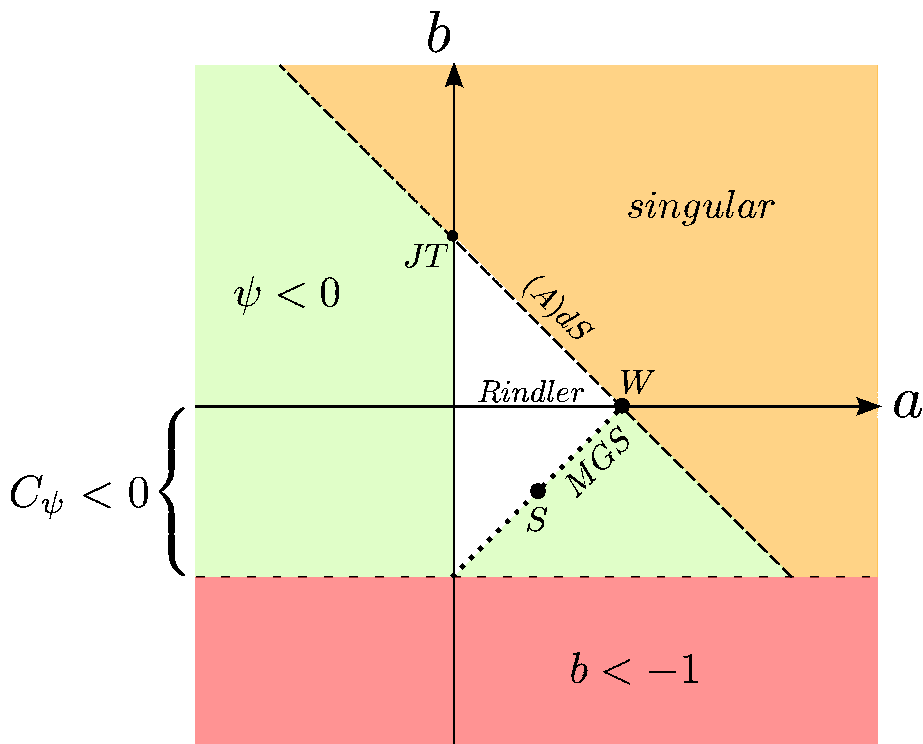
Figure 1: Summary of the ab-family

This solution is commonly referred to as the 'Witten BH'. It is related[7] to the exact background obtained in [2]

$$ds^2 = dx^2 + \frac{\tanh^2(b\,x)}{1 - p\,\tanh^2(b\,x)}\,d\tau^2 \qquad X = X_0\,\cosh^2(b\,x)\,\sqrt{1 - p\,\tanh^2(b\,x)} \qquad (5.2)$$

where the parameters b and p are

$$\alpha'\,b^2 = \frac{1}{k-2} \qquad p := \frac{2\,\alpha'\,b^2}{1 + 2\,\alpha'\,b^2} = \frac{2}{k}\;. \qquad (5.3)$$

The background (5.2) can be expressed in the form (2.3) by means of coordinate transformations described in [46].

In the following subsections we wish to consider BHs for the full range $k \in (2, \infty)$ allowed by the CFT. However, in order for the background (5.2) to be a solution of string theory it must satisfy the condition

$$\mathrm{Dim} - 26 + 6\,\alpha'\,b^2 = 0\;. \qquad (5.4)$$

Because the target space here is two-dimensional, $\mathrm{Dim} = 2$, requiring the correct central charge fixes the level at the critical value $k_{\mathrm{crit}} = 9/4$. Following [47], we vary k by allowing for additional matter fields that contribute to the total central charge, modifying the condition (5.4) so that $k \neq 9/4$.

---

[7]The precise relationship is as follows: in the limit $k \to \infty$ (or, equivalently, $p \to 0$) upon identifying $y = bx$ and $\tau = \sqrt{\alpha' k}\,\theta$ the line-element and dilaton (5.1) emerge from (5.2).

5.1 Witten black hole

The Witten BH has the most intriguing thermodynamics of all the models we have examined so far [12, 32, 10]. We focus on the background (5.1) as a solution of the string theory β-functions at lowest order in α'. The β-functions can be derived from a target-space action of the form (1.1) with

$$U = -\frac{1}{X} \qquad V = -\frac{\lambda^2}{2} X \tag{5.5}$$

where $\lambda = 4/\sqrt{\alpha'}$. With these potentials and convenient choices for the integrations constants in (2.6) and (2.7) the functions Q and w are given by

$$e^Q w = 1, \qquad w = \lambda X. \tag{5.6}$$

Because w is linear in X, the surface gravity and asymptotic temperature $T = \lambda/(4\pi)$ do not depend on the mass M. As a result, the condition (3.3) identifies a single solution consistent with the boundary conditions β_c and D_c. The relation $2M = \lambda X_h$ then implies a simple proportionality between mass and entropy

$$M = TS. \tag{5.7}$$

An immediate consequence is that the Witten BH satisfies the Gibbs-Duhem relationship: $G_c^W = \bar{G}_c^W = 0$.

The two definitions for the dilaton charge (3.1) and (3.2) are both linear in X_c for the Witten BH. They are related by $\bar{D}_c = \lambda D_c$. Therefore, the dilaton chemical potential associated with $\bar{D}_c$,

$$\bar{\psi}_c^W = \frac{1}{2}\left(1 - \frac{X_h}{X_c}\right)^{1/2} + \frac{1}{2}\left(1 - \frac{X_h}{X_c}\right)^{-1/2} - 1, \tag{5.8}$$

is related to ψ_c^W from (3.13) by the constant rescaling $\psi_c^W = \lambda \bar{\psi}_c^W$. Both the internal energy and the Helmholtz free energy

$$E_c^W = \lambda X_c \left(1 - \sqrt{1 - \frac{X_h}{X_c}}\right) \qquad F_c^W = E_c^W - T_c^W S \tag{5.9}$$

have finite asymptotic limits, $E^W \to M$ and $F^W \to 0$, respectively. The extensivity properties of the Witten BH are comparable to those of the standard thermodynamics of ideal gases, as can be seen from table 1 on page 33.

There are various reasons to expect that a semi-classical analysis of the Witten BH might encounter problems. For instance, the curvature is of order $1/\alpha'$ near the horizon, indicating that corrections have to be taken into account. It is not surprising that this system also exhibits some peculiar thermodynamic properties. The most obvious is the behavior of the the specific heat, which is given by

$$C_D^W = 4\pi(X_c - X_h). \tag{5.10}$$

The specific heat is positive for all $X_c > X_h$, but it diverges linearly if we try to remove the cavity wall to infinity. In addition, the quantities C_ψ and κ_T are not well-defined.

However, since we know that the Witten BH is related to the exact solution (5.2), it is reasonable to expect that these problems might be absent in a more complete treatment. As we shall demonstrate in the next subsection, α' corrections have a significant impact on the thermodynamics.

5.2 String theory is its own reservoir: exact string black hole

The discussion of the canonical ensemble in section 3 required the addition of a cavity wall at X_c that couples the system to a thermal reservoir. In most cases thermodynamic stability is only possible for finite values of X_c. We have not questioned the nature of this reservoir, or the form of the coupling between the two systems. In previous examples we simply assumed that it was possible to couple a given model to some other degrees of freedom that functioned as the reservoir. String theory is fundamentally different in this regard. It is not possible to couple arbitrary degrees of freedom to the theory, or place an abrupt cut-off on the space-time fields. Therefore, if we apply our methods to a non-compact solution of string theory any cut-offs introduced during a calculation *must* be removed.

As shown in the previous subsection, the specific heat of the Witten BH diverges as the cut-off is removed to infinity. Remarkably, this and other issues are resolved when α' corrections are taken into account. The metric and dilaton for the ESBH were given in (5.2). A target space action of the form (1.1) was constructed for the Exact String BH in [46]. In our current notation the relevant functions determining the model are given by[8]

$$e^Q w = 1 \,, \qquad w(X) = 2b \left(\sqrt{\rho^2 + 1} + 1 \right) \,, \tag{5.11}$$

where b was defined in (5.3), and the canonical dilaton X is related to a new field ρ by

$$X = \rho + \operatorname{arcsinh} \rho \,. \tag{5.12}$$

It will generally be more convenient to work with the field ρ, rather than X. The relation (5.12) between the two fields is globally invertible, and the limits $X \to \infty$ and $X \to 0$ coincide with the corresponding limits for ρ. Therefore, ρ can be employed not only locally but globally. Expressed in terms of ρ, the square of the Killing norm is given by

$$\xi = 1 - \frac{k}{\sqrt{\rho^2 + 1} + 1} \,. \tag{5.13}$$

This vanishes at

$$\rho_h = \sqrt{k(k-2)} \,. \tag{5.14}$$

which establishes the location of the horizon in terms of the level k.

As in [46], the integration constant in $w(X)$ has been chosen assuming a Minkowski ground state, i.e. $e^Q w = 1$. Comparing the general form of the Killing norm (2.5)

---

[8]Winding/momentum mode duality relates the Exact String BHs to a different set of target space geometries, namely the Exact String Naked Singularities. They are also described by (5.11) and (5.12), but with sign changes in front of the square root and in front of the arcsinh .

with (5.13) shows that the mass parameter is then proportional to the level: $M^{MGS} = bk$. This means that the choice of w_0 in (5.11) identifies the ground state with $k = 0$. However, the CFT description requires $k > 2$. This can also be seen from the Hawking temperature [47],

$$T^{ESBH} = \frac{b}{2\pi}\sqrt{1 - \frac{2}{k}}\,, \tag{5.15}$$

which is real and positive only if $k > 2$. Therefore, we must shift the constant w_0 so that the ground state corresponds to level $k = 2$. This is accomplished by replacing w with $w - 4b$ in (5.11). Notice that this is the largest shift by a negative constant that is consistent with $w > 0$. The result is

$$e^{Q^{ESBH}(X)} = \frac{1}{2b}\frac{1}{\sqrt{\rho^2 + 1} + 1}\,, \qquad w^{ESBH}(X) = 2b\left(\sqrt{\rho^2 + 1} - 1\right)\,. \tag{5.16}$$

Comparing the resulting expression for the Killing norm with (5.13) gives the mass parameter in terms of b and k

$$M^{ESBH} = b(k - 2)\,. \tag{5.17}$$

With (5.16) the dilaton charge (3.2) reads

$$\bar{D}_c^{ESBH} = 2b\,\rho_c\,. \tag{5.18}$$

For given β_c and D_c the equation (3.3) has a unique solution with level $k > 2$. In the asymptotic region $D_c \to \infty$ the temperature is bounded from above and below for $k \in (2, \infty)$ according to (5.15). As we shall show in the next subsection, the thermal partition function diverges in the $k \to \infty$ limit, which identifies the maximal temperature $T_H = b/(2\pi)$ as the Hagedorn temperature. As might have been anticipated on general grounds the asymptotic temperature has to be positive and below the Hagedorn temperature for the existence of the ESBH solution.

5.2.1 Thermodynamics of the exact string black hole

With (5.16) as a starting point we can calculate all thermodynamical quantities of interest for the ESBH by applying the general results of section 3. The cut-off is always removed at the end of a calculation, so we drop the subscript 'c' from the following expressions. The result for the entropy,

$$S^{ESBH} = 2\pi\left(\sqrt{k(k - 2)} + \operatorname{arcsinh}\sqrt{k(k - 2)}\right)\,, \tag{5.19}$$

agrees with equation (4.17) found in [46] (In that reference $M = k/2$.). For $k = 9/4$ we obtain $S^{ESBH} = 2\pi(3/4 + \ln 2)$. In the large k limit the entropy $S^{ESBH} = 2\pi k + \mathcal{O}(\ln k)$ asymptotes to the result for the Witten BH with logarithmic α' corrections [48]. The free energy,

$$F^{ESBH} = -b\sqrt{1 - \frac{2}{k}}\operatorname{arcsinh}\sqrt{k(k - 2)} \tag{5.20}$$

is manifestly non-positive.[9] This would not have been the case with the original choice of w_0 that lead to (5.11). The result (5.20) coincides with (5.7) in [46]. The internal energy is given by the mass, $E^{ESBH} = M^{ESBH} = b(k-2)$. The specific heat

$$C_D = \bar{C}_\psi = 2\,\pi\,k\,\sqrt{k(k-2)} \tag{5.21}$$

is positive and finite. Thus the problem encountered for the Witten BH is absent: removing the cut-off is possible for the Exact String BH and leads to a positive and finite specific heat. It is worth mentioning that the specific heat and entropy both vanish in the limit $T \to 0$, so the Exact String BH is one of the few examples for which the third law of thermodynamics holds.

Even though the Minkowski ground state condition is not satisfied by (5.16), the Killing norm still approaches unity as the cut-off is removed. Therefore the asymptotic limit of (3.15) is constant, $\eta^{ESBH} = 1/2$. As a consequence the dilaton-chemical potential vanishes as X_c^{-2} in the limit $X_c \to \infty$. The Gibbs-Duhem relation

$$\bar{G}^{ESBH} = F^{ESBH} = M^{ESBH} - T^{ESBH}S^{ESBH} \neq 0 \tag{5.22}$$

is violated even in the large k limit,

$$\lim_{k\to\infty} \bar{G}^{ESBH} = \lim_{k\to\infty} F^{ESBH} = -b\ln 2k + \mathcal{O}\left(\frac{1}{k}\ln k\right). \tag{5.23}$$

Thus, comparing the thermodynamic behavior of the exact String BH with that of the Witten BH demonstrates the importance of α' corrections, even in the large k limit. In the limit $k \to 2$ the Gibbs-Duhem relation,

$$\lim_{k\to 2} \bar{G}^{ESBH} = \lim_{k\to 2} F^{ESBH} = -b(k-2) + \mathcal{O}(k-2)^2\,, \tag{5.24}$$

is restored.

It is interesting to observe how the free energy scales with the level k in various limits. For large values of k the free energy scales logarithmically, while for k close to the ground state value it is linear in $k - 2$. This implies corresponding behavior of the thermal partition function (2.31). Suppose that we fix the factor in front of the action as $\sigma/(2\pi)$, where σ is some positive parameter, so that the curvature term in the action 1.1 is given by $-\sigma X R/(4\pi)$. Then the thermal partition function in the large k limit is given by

$$\left. \mathcal{Z}^{ESBH} \right|_{k\gg 2} = (2k)^\sigma + \cdots \sim \left(S^{ESBH}\right)^\sigma + \ldots \tag{5.25}$$

In this limit the partition function grows as a positive power of k. It diverges in the $k \to \infty$ limit, as the temperature approaches the asymptotic value $T_H = b/(2\pi)$. In this sense, T_H represents the Hagedorn temperature. Similarly, for $k \to 2$ we obtain

$$\left. \mathcal{Z}^{ESBH} \right|_{k\to 2} = e^{\sigma\sqrt{2k-4}} + \cdots \sim e^{\frac{\sigma}{4\pi}S^{ESBH}} + \ldots \tag{5.26}$$

---

[9]It is worth mentioning that there is also a CDV solution with $X = 0$ and AdS geometry with curvature proportional to b^2. It has vanishing free energy. Therefore the Exact String BH is stable against tunneling (cf. subsection 3.11).

Thus, in the weak coupling limit the partition function grows only with a power of entropy, while in the strong coupling limit it grows exponentially with entropy. This is consistent with fermionization in the strong coupling regime.[10] For arbitrary values of k we obtain

$$\mathcal{Z}^{ESBH}(k) = e^{\sigma \operatorname{arcsinh} \sqrt{k(k-2)}}. \tag{5.27}$$

Because $\exp\left[\operatorname{arcsinh}(3/4)\right] = 2$ we obtain the interesting result

$$\mathcal{Z}^{ESBH}(9/4) = 2^{\sigma}. \tag{5.28}$$

As a caveat, we remind the reader that these results apply to the semi-classical approximation of the thermal partition function, based on the action for the ESBH given in [46]. They do not take into account contributions from the tachyon that appears in the bosonic string spectrum, or from any matter that must be coupled to the ESBH to maintain the correct central charge (5.4). Furthermore, while we have focused on the effects of α' corrections, we have not taken into account contributions from world-sheet topologies other than the leading term.

5.2.2 Thermodynamical derivation of the exact string black hole

It is interesting to note that the potentials (5.11) can be derived purely from thermodynamical considerations. Of course, this requires some ad-hoc assumption: we look for a model which may be shifted to a Minkowski ground state model and whose BH solutions have a specific heat whose asymptotic behavior is

$$C_D = \bar{C}_\psi \propto M^2 T \tag{5.29}$$

for all values of M and T. The rationale behind this assumption comes from the consideration of the weak ($k \to \infty$) and strong ($k \to 2$) coupling limit. In the former case we have seen that the ESBH is approximated by the Witten BH which has a divergent asymptotic specific heat. Taking into account corrections from fluctuations at 1-loop level leads to a specific heat $C_D \propto M^2 T$ [50]. This is a non-stringy calculation, but as long as we are not interested in the precise proportionality factor it should not matter which kind of fluctuations we consider perturbatively: thermal fluctuations, α' corrections and quantum fluctuations of some additional massless matter fields are expected to yield the same qualitative behavior. In the strong coupling limit the system can be fermionized (cf. footnote 10). A free Fermi gas has a specific heat of the form $C \propto T$. The proportionality constant can be rescaled such that it includes a factor M^2. Then the weak and

---

[10] In the WZW formulation the stress tensor is given by

$$T(z) = \frac{1}{k-2}\eta_{ab}J^a J^b + \frac{k}{4}(\partial\phi)^2 + b\partial c.$$

Here η_{ab} is the metric on the algebra $sl(2,\mathbb{R})$ and the J^a are given by equation (2.11) in [2]. For $k \to 2$ the first term dominates and the system behaves like a free Fermi gas, cf. e.g. [49]. This suggests that an appropriate interpretation of the Exact String BH in that limit is not in terms of some effective geometry but rather in terms of free fermions.

strong coupling limits qualitatively both yield the same result for the specific heat, which motivates our Ansatz (5.29).

With $C_D \propto w_h'/w_h''$, $T \propto w_H'$ and $2M = w_h$ the most general model (1.1) obeying (5.29) has to fulfill the autonomous differential equation

$$w'' \propto \frac{1}{w^2} \,. \tag{5.30}$$

Without loss of generality we fix the proportionality constant in (5.30) to unity. Its general solution is

$$\int \frac{dw}{\sqrt{c_1 - 2/w}} = c_0 \pm X \,. \tag{5.31}$$

The integration constant c_1 may be adjusted by rescaling w, c_0 and X, so we fix it to unity. The integration constant c_0 may be adjusted by shifting the origin of X, so we fix it to zero. The sign ambiguity is resolved by requiring X to be positive. Thus we obtain

$$\sqrt{w(w-2)} + \ln\left[w - 1 + \sqrt{w(w-2)}\right] = X \,. \tag{5.32}$$

Defining $\rho = \sqrt{w(w-2)}$ reveals that (5.32) actually is equivalent to (5.12). Solving w in terms of ρ and taking the branch of positive w leads to

$$w(X) = \sqrt{\rho^2 + 1} + 1 \,, \qquad X = \rho + \operatorname{arcsinh} \rho \,. \tag{5.33}$$

This is the same as $w(X)$ in (5.11) for $b = 1/2$. If we had chosen a different proportionality constant in (5.30) we would obtain a different value for b but otherwise the same result for w. The function Q is fixed by the Minkowski ground state property (4.1). Thus, the Exact String BH in the Minkowski ground state representation (5.11) is the most general two-dimensional dilaton gravity model with a Minkowskian ground state that has an asymptotic specific heat $C_D = \bar{C}_\psi \propto M^2 T$.

6. Spherically symmetric black holes in higher dimensions

Up to this point we have focused on theories in two dimensions, but the results of section 3 can also be applied to solutions of gravitational theories in higher dimensions, as long as the on-shell action reduces to the form (2.38) or (3.54). With a few caveats, this leads to a simple description of the thermodynamics of a class of spherically symmetric BHs in $d + 1 > 2$ dimensions, with or without a cosmological constant.

Pure gravity with a cosmological constant in $d + 1 > 2$ dimensions is described by the action

$$I_{d+1} = -\frac{1}{16\pi G_{d+1}} \int_{\mathcal{M}} d^{d+1}x \sqrt{g_{d+1}} \left(R_{d+1} - 2\Lambda\right) - \frac{1}{8\pi G_{d+1}} \int_{\partial \mathcal{M}} d^d x \sqrt{\gamma_d} \left(K_d + \ldots\right) \,. \tag{6.1}$$

Subscripts '$d + 1$' and 'd' have been used to distinguish quantities from their two- and one-dimensional analogs. The '$\ldots$' in the boundary integral indicates boundary counterterms, whose precise definition depends on the space-times asymptotics [8, 19]. We are interested

in solutions of this theory that can be expressed as a direct product of the two-dimensional metric (2.3) and the metric on a round $d-1$ sphere

$$ds^2 = \xi(r)\, d\tau^2 + \frac{1}{\xi(r)}\, dr^2 + (G_{d+1})^{\frac{2}{d-1}}\, \varphi(r)^2\, d\Omega_{d-1}^2\ . \tag{6.2}$$

Explicit factors of the Newton's constant have been included in this expression so that the field $\varphi(r)$ is dimensionless. Ignoring for a moment the contributions from the boundary counterterms, the on-shell reduction of (6.1) on the $d-1$ sphere leads to a two-dimensional expression of the form (1.1). The dilaton X in this action is related to the scalar field φ by

$$X(r) = \Upsilon\, G_{d+1}\, \varphi(r)^{d-1} \qquad \Upsilon := \frac{A_{d-1}}{8\,\pi\, G_{d+1}} \tag{6.3}$$

where A_{d-1} is the solid angle subtended by the $d-1$ sphere, and the constant Υ has been defined for convenience. Thus, the dilaton $X(r)$ is the proper area of a sphere with coordinate radius r, in $d+1$ dimensional Planck units. The kinetic and potential functions for the dilaton are

$$U(X) = -\left(\frac{d-2}{d-1}\right) \frac{1}{X} \tag{6.4}$$

$$V(X) = -\frac{1}{2}\,(d-1)(d-2)\,\Upsilon^{\frac{2}{d-1}}\, X^{\frac{d-3}{d-1}} + e\,\frac{d(d-1)}{2\,\ell^2}\, X\ . \tag{6.5}$$

The last term in V comes from the cosmological constant, which we parameterize in terms of a length scale ℓ by

$$\Lambda := e\frac{d(d-1)}{2\,\ell^2} \qquad e = \pm 1, 0\ . \tag{6.6}$$

The coefficient of the reduced action explicitly sets the two-dimensional Newton's constant to $8\pi G_2 = 1$. From (6.4) and (6.5) we obtain

$$e^{Q(X)} = \frac{\Upsilon^{\frac{1}{1-d}}}{d-1}\, X^{\frac{2-d}{d-1}}\ , \tag{6.7}$$

$$w(X) = (d-1)\,\Upsilon^{\frac{1}{d-1}}\, X^{\frac{d-2}{d-1}} \left(1 - \frac{e}{\ell^2}\,\Upsilon^{-\frac{2}{d-1}}\, X^{\frac{2}{d-1}}\right)\ , \tag{6.8}$$

where we have fixed the free constants in the definitions (2.6) and (2.7) conveniently.

The results of section 3 can be applied to these solutions of the $d+1$ dimensional theory. This is because the on-shell action, after integrating over the sphere, takes the form (2.38). Of course, the contributions from the boundary counterterms, alluded to by the '...' in (6.1), must be properly accounted for. Otherwise the reduced on-shell action would diverge, as in (2.15). This raises an interesting technical problem. The formulation of the boundary counterterms in the $d+1$ dimensional theory depends on the asymptotics of the space-time [51, 8]. This would seem to imply that the reduction of theories with $\Lambda = 0$ and $\Lambda \neq 0$ must be studied separately. We take a different approach. Rather than working out the reduction of the appropriate boundary counterterms in (6.1), we assume that the two-dimensional boundary counterterm derived in subsection 2.4 gives a

sensible renormalization of the reduced on-shell action. For the solution described above, the boundary counterterm is

$$I_{CT} = -\beta \, w(X) \sqrt{1 - \frac{2\,M}{w(X)}} \; . \tag{6.9}$$

This counterterm gives the 'correct' on-shell action in the sense that the analysis of section 3 recovers the standard results associated with the thermodynamics of BHs whose metrics take the form (6.2). An immediate consequence, due to (6.3), is

$$S = 2\pi X_h = \frac{A_h}{4G_{d+1}} \; . \tag{6.10}$$

The entropy of these BHs is *always* given by one-quarter of the proper area of the horizon, in $d+1$ dimensional Planck units.

There is no reason to believe that the counterterm (6.9) can be lifted to higher dimension, to recover the counterterms of the original theory. The combination of the two-dimensional metric and scalar field are simply not sufficient to reproduce the appropriate functional dependence on the higher-dimensional metric. But despite this caveat, it seems remarkable that the quasi-local thermodynamics of spherically symmetric BHs with different space-time asymptotics can be recovered from the reduced on-shell action, renormalized by the two-dimensional counterterm (2.29).

6.1 Schwarzschild black hole in $d+1 \geq 4$ dimensions

As a first example, consider the on-shell reduction of (6.1) with $\Lambda = 0$. Setting $\varphi \sim r$, up to factors of the Newton's constant, gives the standard form of the Schwarzschild solution

$$\xi(r) = 1 - \frac{2\widetilde{M}}{r^{d-2}} \tag{6.11}$$

$$X(r) = \Upsilon \, r^{d-1} \tag{6.12}$$

where the mass parameter $\widetilde{M}$ is related to the constant M by

$$\widetilde{M} := \frac{M}{(d-1)\Upsilon} \; . \tag{6.13}$$

In section 3 the dilaton charge $D_c = X_c$ and proper periodicity β_c were held fixed at $r = r_c$. This is precisely the canonical ensemble studied in [52, 35], where the BH is enclosed inside a cavity of radius r_c whose area and proper temperature are held fixed by means of an external thermal reservoir. The relation (6.3) identifies the dilaton charge $D_c = X_c$ as the area of the cavity, so the dilaton chemical potential (3.13) is the corresponding pressure

$$p_c = \frac{1}{2}(d-2)\, r_c^{-1} \left(\sqrt{\xi(r_c)} + \frac{1}{\sqrt{\xi(r_c)}} - 2 \right) \; . \tag{6.14}$$

The internal energy of the system can be expressed as

$$E_c = (d-1)\, \Upsilon \, r_c^{d-2} \left(1 - \sqrt{1 - \frac{2\widetilde{M}}{r_c^{d-2}}} \right) \; . \tag{6.15}$$
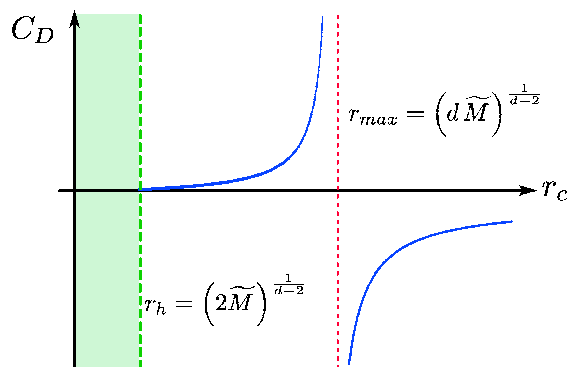
Figure 2: Heat capacity for the Schwarzschild BH in dimension $d + 1 \geq 4$.

The asymptotic limit of this quantity gives the ADM mass M. Furthermore, the internal energy satisfies the quasi-local form of the first law (3.26),

$$dE_c(S, A_c) = T_c\, dS - p_c\, dA_c\,, \tag{6.16}$$

where A_c is the area of the cavity in Planck units ($8\pi G_{d+1} = 1$). Finally, the expression (3.28) for the heat capacity recovers the well-known result for the thermodynamic stability of the Schwarzschild BH enclosed by an isothermal cavity in d+1 dimensions

$$C_D = -4\pi\, M\, r_h\, \frac{r_c^{d-2} - 2\widetilde{M}}{r_c^{d-2} - d\,\widetilde{M}}\,. \tag{6.17}$$

This quantity is positive for $r_h < r_c < r_{max}$, as indicated in figure 2.

6.2 Asymptotically AdS space-times

Now we turn our attention to the case of non-zero cosmological constant. For definiteness, we take the cosmological constant to be negative ($e = -1$) and consider asymptotically AdS space-times.[11] Setting $\varphi \sim r$ leads to the Schwarzschild-AdS solution

$$\xi(r) = 1 + \frac{r^2}{\ell^2} - \frac{2\,\widetilde{M}}{r^{d-2}} \qquad X = \Upsilon\, r^{d-1}\,. \tag{6.18}$$

As in the previous case, holding $D_c = X_c$ and β_c fixed at r_c corresponds to an AdS BH enclosed by an isothermal cavity. The thermodynamics of this system was studied in [53, 54].

Rather than reproduce the full catalog of results regarding the Schwarzschild-AdS solution, we focus on a few interesting properties which are qualitatively different than the previous example. First, the internal energy (3.18) is given by

$$E_c = \frac{(d-1)}{\ell}\, \Upsilon\, r_c^{d-1} \left(\sqrt{1 + \frac{\ell^2}{r_c^2}} - \sqrt{1 + \frac{\ell^2}{r_c^2} - \frac{2\widetilde{M}\ell^2}{r_c^d}} \right)\,. \tag{6.19}$$

---

[11]Asymptotically de Sitter space-times with $\Lambda > 0$ can also be addressed in this framework, but many of the working assumptions made in sections 2.1 and 3 must be changed. We shall discuss this briefly in section 7.

Unlike the result (6.15), the asymptotic limit of the internal energy vanishes due to red-shifting

$$\lim_{r_c \to \infty} E_c = 0 \ . \tag{6.20}$$

However, as in section 3.4, the conserved charge is [35]

$$Q_{\partial_\tau} := \lim_{r_c \to \infty} \sqrt{\xi(r_c)}\, E_c = M \ . \tag{6.21}$$

This is the same result obtained in higher dimensions using any one of a number of methods [55]. But, notably, it does not agree with the standard AdS/CFT result when the spacetime dimension $d+1$ is odd. In that case, global AdS (the ground state) has $Q_{\partial_\tau} = M_{cas}$, where M_{cas} is a Casimir energy associated with the dual field theory on $S^1 \times S^{d-1}$ [22, 23]. Reproducing the Casimir term is possible in the present context: an appropriate shift in w_0 can be used to shift the value of (6.21) by an amount M_{cas}, for all solutions. But from the two-dimensional point of view the shift is arbitrary. The counterterm (6.9) is completely ignorant of the physics attached to this issue in the higher-dimensional theory.

Another interesting result is the structure of the heat capacity for an asymptotically AdS BH. The result is, in general, very complicated for finite r_c [54]. But the presence of a cosmological constant can stabilize the BH even when the cavity wall is removed to infinity. Using (3.29), the $r_c \to \infty$ limit of the heat capacity is

$$\lim_{r_c \to \infty} C_D = (d-1)\, 2\pi X_h \left(\frac{d\, X_h^{\frac{2}{d-1}} + (d-2)\, \ell^2\, \Upsilon^{\frac{2}{d-1}}}{d\, X_h^{\frac{2}{d-1}} - (d-2)\, \ell^2\, \Upsilon^{\frac{2}{d-1}}} \right) \tag{6.22}$$

Unlike the asymptotically flat solutions studied in the previous subsection, the heat capacity may be positive for a large enough BH. The relation between the heat capacity and the size of the BH is shown in figure 3. The condition $C_D > 0$ is satisfied if the horizon radius is larger than a critical size set by the dimension and the AdS length scale

$$r_{crit} = \sqrt{\frac{d-2}{d}}\, \ell \ . \tag{6.23}$$

This corresponds to a critical periodicity for the Euclidean time given by

$$\beta_{crit} = \frac{\sqrt{d(d-2)}}{2\pi\ell} \ . \tag{6.24}$$

The possibility of a phase transition, as studied by Hawking and Page [41], is encountered when r_h drops below the critical value given by (6.23). This is consistent with our general discussion in subsection 3.10.

6.3 Reduction of BTZ

Gravity with a negative cosmological constant in 3 dimensions admits a solution known as the BTZ BH [15]. If $\Lambda = -1/\ell^2$ then (Euclidean) line element of the BTZ BH is

$$ds^2_{BTZ} = \xi(r)d\tau^2 + \frac{1}{\xi(r)}dr^2 + r^2(d\phi - \frac{4G_3 J}{r^2}dt)^2 \tag{6.25}$$
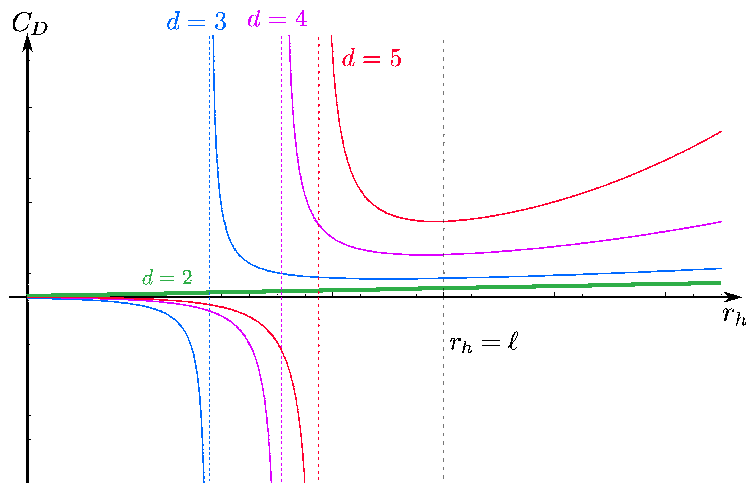
Figure 3: Heat capacity for the Schwarzschild-AdS BH in $d + 1$ dimensions, with the cavity wall sent to infinity. The horizontal axis is the horizon radius r_h, and the vertical axis is the $X_c \to \infty$ limit of C_D.

where

$$\xi(r) = \frac{r^2}{\ell^2} - 8G_3 M + \frac{16G_3^2 J^2}{r^2} \, . \tag{6.26}$$

If the inequality $|J| \leq M\ell$ holds then two horizons exist,

$$r_\pm = \ell \sqrt{4G_3 M} \sqrt{1 \pm \sqrt{1 - \frac{J^2}{M^2\ell^2}}} \, . \tag{6.27}$$

The normalization chosen in the rest of this work would require $G_3 = 1/4$, but we set $G_3 = 1/8$ to obtain the same normalization of the action used in [15, 56]. Some useful formulas are

$$\frac{r_+^2 + r_-^2}{\ell^2} = M \, , \qquad \frac{2r_+ r_-}{\ell} = J \, , \qquad \frac{r_+^2 - r_-^2}{\ell^2} = M\sqrt{1 - \frac{J^2}{M^2\ell^2}} \, . \tag{6.28}$$

Achucarro and Ortiz performed a Kaluza-Klein reduction

$$ds_{BTZ}^2 = g_{\alpha\beta} dx^\alpha dx^\beta + X^2 \left(d\phi + A_\alpha dx^\alpha \right)^2 \tag{6.29}$$

to two dimensions. The dilaton field is the surface radius, $X = r$, and the two-dimensional line element $g_{\alpha\beta} dx^\alpha dx^\beta$ corresponds to the τr-part of (6.25). They integrated out the gauge field A_α and obtained an action of the type (1.1) with [56]

$$U^{AO} = 0 \, , \qquad V^{AO} = -\frac{X}{\ell^2} + \frac{J^2}{4X^3} \, . \tag{6.30}$$

We point out an important deficiency of (6.30): the angular momentum J enters here as a parameter in the action rather than emerging as a constant of motion. Consequently, the thermodynamics is not reproduced correctly from the two-dimensional model (1.1) with the potentials (6.30). But there is a simple solution to this problem: we have demonstrated

in subsection 3.8 how to treat charged BHs and noted that upon introducing Maxwell fields a parameter in the action may be converted into a constant of motion, namely a conserved charge. In fact, the Kaluza-Klein Ansatz (6.29) already contains such a Maxwell field and all one has to do is to refrain from integrating it out. Therefore, we study an action (3.45) [multiplied by 2 to get the normalizations above] with the functions

$$U^{BTZ}(X) = 0 \,, \tag{6.31}$$

$$V^{BTZ}(X) = -\frac{X}{\ell^2} \,, \tag{6.32}$$

$$f^{BTZ}(X) = -\frac{1}{8} X^3 \,. \tag{6.33}$$

The potentials (6.31) and (6.32) are consistent with (6.4) and (6.5), respectively, for $d+1 = 3$. The Maxwell field with coupling function (6.33) arises because we are treating a spinning BH rather than a spherically symmetric one. The conserved $U(1)$ charge therefore has to be identified with the angular momentum, $q \propto J$. The proportionality factor depends on the conventions. With our choices the correct relation is $q = J/2$. For further calculations we need the functions

$$Q^{BTZ}(X) = 0 \,, \qquad w^{BTZ}(X) = \frac{X^2}{\ell^2} \,, \qquad h^{BTZ}(X) = \frac{4}{X^2} \tag{6.34}$$

and note that the ground state solution $M = J = 0$ is pure AdS$_3$, $\xi_0 = X^2/\ell^2$.

We show now that this effective two-dimensional description leads to the correct thermodynamics for the BTZ BH. To this end we apply the general results of subsection 3.8 to the specific choice (6.34) recalling that $q = J/2$. The square of the Killing norm

$$\xi(X) = \frac{X^2}{\ell^2} - M + \frac{J^2}{4X^2} \tag{6.35}$$

indeed is equivalent to (6.26) for $G_3 = 1/8$ and $X = r$. This together with the gauge potential

$$A_r = 0 \,, \qquad A_\tau = -\frac{J}{2X^2} \,, \tag{6.36}$$

plugged into (6.29) allows to recover the BTZ line-element (6.25). With the proper angular momentum [cf. (3.53)]

$$\Omega_c := \Phi_c = \frac{J}{2\sqrt{\xi_c}} \left(\frac{1}{X_h^2} - \frac{1}{X_c^2} \right) \tag{6.37}$$

the thermodynamic potential (3.55) is given by

$$Y(T_c, X_c, \Omega_c) = -T_c S + E_c - \Omega_c J \,. \tag{6.38}$$

Entropy is determined by the value of the dilaton at the outer horizon, $X_h = r_+$. Taking into account that the normalization of Newton's constant differs in this subsection from the one used in the rest of this work by a factor of 2 we get from (3.9)

$$S = 4\pi r_+ \,. \tag{6.39}$$

The result (6.39) coincides with the entropy of the BTZ calculated by counting of microstates, cf. e.g. [57]. The inverse Euclidean periodicity determines temperature,

$$T = \frac{r_+^2 - r_-^2}{2\pi r_+ \ell^2} = \frac{2M}{S}\sqrt{1 - \frac{J^2}{M^2\ell^2}}\,. \tag{6.40}$$

This coincides with the result for the temperature of the BTZ BH [15]. Obviously, in the limit $|J|/(M\ell) \to 1$ temperature vanishes, which is consistent with the fact that for $|J| = M\ell$ the BTZ BH becomes extremal. If we define the quantity Ω as

$$\Omega := \lim_{X_c \to \infty} \sqrt{\xi_c}\,\Omega_c = \frac{J}{2r_+^2} \tag{6.41}$$

then we obtain the correct first law relating the conserved charges,

$$dM = TdS + \Omega dJ\,. \tag{6.42}$$

The integrated version

$$2M = TS + 2\Omega J \tag{6.43}$$

shows that extensitivity properties differ slightly from standard thermodynamics. However, they are as expected for a BH in AdS: if the dilaton (and thus entropy) is extensive then the square-root of the mass M is extensive, consistently with the extensitivity properties in table 1. This explains the factor of 2 on the left hand side of (6.43). The factor 2 on the right hand side is the standard result for spinning BHs. With the results above, the expression for internal energy [note again the factor of 2 as compared to (3.18) because of the different normalization of the action],

$$E_c(T_c, X_c, J) = 2\sqrt{\xi_0} - 2\sqrt{\xi_c} = \frac{2X_c}{\ell}\left(1 - \sqrt{1 - \frac{M\ell^2}{X_c^2} + \frac{J^2\ell^2}{4X_c^4}}\right)\,, \tag{6.44}$$

and surface pressure

$$\psi_c = \frac{E_c}{\ell\sqrt{\xi_c}} - \frac{J^2}{2X_c^3\sqrt{\xi_c}} \tag{6.45}$$

lead to the quasilocal form of the first law

$$dE_c = T_c dS + \Omega_c dJ - \psi_c dX_c\,. \tag{6.46}$$

All our results are compatible with [53] (cf. also [58]). This shows that the BTZ BH thermodynamics is compatible with a Kaluza-Klein reduction to two dimensions. A generalization to a charged BTZ BH is straightforward and requires the addition of another Maxwell field with a linear coupling to the dilaton. A different generalization involves the coupling to a gravitational Chern-Simons term [59], which yields upon dimensional reduction [60] a model of type (3.45) [61] whose entropy recently has been derived in [62] using Wald's Noether charge technique [63].

7. Discussion

In this paper we have studied black hole thermodynamics for arbitrary models of dilaton gravity in two dimensions. The analysis of these models was complicated by the fact that the path integral defined with respect to the action (1.1) does not have a sensible semi-classical limit. We resolved this problem by constructing an improved action,

$$
\Gamma = -\frac{1}{16\pi G_2} \int_{\mathcal{M}} d^2x \sqrt{g} \left[X R - U(X) (\nabla X)^2 - 2 V(X) \right]
$$
$$
- \frac{1}{8\pi G_2} \int_{\partial\mathcal{M}} dx \sqrt{\gamma} X K + \frac{1}{8\pi G_2} \int_{\partial\mathcal{M}} dx \sqrt{\gamma} \sqrt{w(X) e^{-Q(X)}} , \quad (7.1)
$$

generalizing the techniques first used in [10] to obtain the boundary counterterm (2.29). The counterterm is unique up to a shift ambiguity[12] which we fixed by demanding that it be independent under changes of scale associated with constant shifts in $Q(X)$. Thus, one should consider (7.1), instead of (1.1), as 'the generic dilaton gravity action'. Following the approach of York [11], we proceeded to define the canonical partition function for a black hole enclosed in the 'cavity' $X \leq X_c$. The Helmholtz free energy (3.5) for this system is related to the improved action by $F_c = \beta_c^{-1} \Gamma_c$, where the proper inverse temperature β_c^{-1} is fixed at the cavity wall X_c by an external thermal reservoir. All of the results on thermodynamics in section 3 follow from this identification. After deriving various thermodynamical quantities, we generalized our results to include two dimensional Maxwell fields and their associated conserved charges. We then examined the conditions under which the system is thermodynamically and mechanically stable, and described processes which might destabilize the system. Our results were then applied to a wide range of models in sections 4- 6. This includes backgrounds that appear in two-dimensional solutions of string theory (section 5). Finally, we discovered that our results can be used to describe the thermodynamics of certain higher dimensional black holes (section 6).

An important aspect of our analysis is universality. For instance, we confirmed that the entropy (3.9) for a dilaton gravity black hole in two dimensions is essentially given by X_h, the dilaton evaluated at the horizon. This result is independent of the details that define a particular model. If we define the 'area' of the horizon via the $d \to 1$ limit of a d dimensional sphere and restore Newton's constant, then this result can be expressed as

$$
S = \frac{A_h}{4 G_{\text{eff}}} . \quad (7.2)
$$

The effective Newton's constant is defined as $G_{\text{eff}} = G_2/X_h$. Since G_2 is dimensionless there is no notion of Planck length in two dimensions. However, we can think of G_{eff} as defining an effective resolution at the horizon, in the sense that decreasing G_{eff} increases the number of bits that can be stored by the black hole. Then (7.2) can be interpreted in much the same way as the usual relationship between entropy and area in higher dimensions. In section 3.4 we obtained a simple expression for the (proper) internal energy of the system

$$
E_c = e^{-Q(D_c)} \left(\sqrt{\xi_0} - \sqrt{\xi_c} \right) . \quad (7.3)
$$

---

[12]In addition there is a sign ambiguity which is trivially fixed by demanding finiteness of the improved on-shell action.

This agrees with the calculation of the energy using the quasilocal stress tensor defined by Brown and York [35]. We were able to show that every dilaton gravity black hole satisfies a quasilocal form of the first law of black hole thermodynamics

$$dE_c = T_c \, dS - \psi_c \, dD_c + \Phi_c \, dq \; . \tag{7.4}$$

This expression holds at arbitrary values of X_c, incorporates the dilaton charge D_c and its chemical potential ψ_c, allows for Maxwell fields with proper electrostatic potentials Φ_c and conserved charges q, and accounts for the (non-linear) effects of gravitational binding energy present in E_c. The canonical first law (7.4) contains considerably more information than the microcanonical $dM = T \, dS$, which becomes a triviality if expressed as $dw_h = w'_h dX_h$.

Despite our best attempts to provide a comprehensive study, there are several directions that we have not had a chance to explore. A number of 'working assumptions' were made throughout the paper and it is important to understand which ones can be relaxed. An example is our decision to focus on models where $w \to \infty$ as $X \to \infty$. This is a reasonable assumption, as it seems to be the most common behavior among dilaton gravity models. But there are certainly models for which it does not apply. For instance, one might consider models where w approaches a constant, or $w \to -\infty$, as $X \to \infty$. Our approach should still be applicable, though many of the considerations in section 3 must be modified. In the first case, where w approaches a constant, we can make a constant shift so that $w \to 0$ as $X \to \infty$. Then one can exploit the duality described in [64], which preserves the classical solution for the metric but replaces w with $\tilde{w} = 1/w$. This implies $\tilde{w} \to \infty$ as $X \to \infty$, which can be analyzed under our initial assumptions. The second case, $w \to -\infty$ as $X \to \infty$, is perhaps more interesting as it includes the dimensional reduction of the static patch of de Sitter space. We now have to consider the region $X \leq X_{dS}$, where X_{dS} corresponds to the observer-dependent cosmological horizon in the static patch. Although we have not performed this analysis, we expect that our approach should reproduce the standard results for the thermodynamics of cosmological horizons. An additional complication emerges when one considers a second horizon, $X_h < X_{dS}$, associated with a black hole. The analysis of that system is beyond the scope of this paper.

It would be interesting to study the impact of the boundary counterterm in (7.1) on the exact path integral quantization procedure in the first order formulation of dilaton gravity, cf. e.g. [65, 3]. While the bulk term and the Gibbons-Hawking-York boundary term look quite different from (1.1) in the first order formulation, we expect that the boundary counterterm (2.29) can be translated easily to its first order counterpart, because it depends solely on the dilaton field and the induced volume form at the boundary.

Another idea that deserves to be studied in more detail was mentioned at the end of subsection 3.8. Consider a dilaton gravity model with a potential of the form (3.59). The same model can be obtained by integrating out a number of Maxwell fields whose coupling functions $f(X)$ are related to terms in the potential. This means that various coupling constants in the potential $V(X)$ can be converted to conserved $U(1)$ charges associated with Maxwell fields. This technique allowed us to study the thermodynamics of the BTZ BH in subsection 6.3. The first law for the rotating, three dimensional solution was recovered by 'integrating in' a gauge field that Achucarro and Ortiz had integrated out in their toroidal

reduction [56]. It would be interesting to study the extent to which the thermodynamics of higher dimensional solutions can be reproduced in this manner.

The most pressing generalization of our work involves coupling two dimensional dilaton gravity to matter, such as a scalar field. This is important for several reasons. As mentioned in the introduction, we have assumed that these models will be coupled to some form of propagating matter, and scalar fields are a prime candidate. We have also considered backgrounds that arise in two dimensional solutions of string theory, and any further analysis of these models should try to incorporate the tachyon. In general, coupling additional fields to the dilaton gravity model will introduce complications at the level of the Hamilton-Jacobi equation. However, even if this equation can no longer be solved exactly, it can still be addressed perturbatively. This approach has been used successfully for gravity coupled to a scalar field in higher dimensions [66 – 69]. We hope to address this in a future publication.

Acknowledgments

A. Conventions, dimensions, and a dilaton gravity bestiary

Let $\mathcal{M}$ be a two-dimensional manifold with boundary $\partial\mathcal{M}$. We consider a metric $g_{\mu\nu}$ on $\mathcal{M}$ that induces a metric γ_{ab} on $\partial\mathcal{M}$. The (one-dimensional) boundary indices are useful for keeping track of factors of γ and its inverse.

The Riemann tensor on $(\mathcal{M}, g)$ and its contractions are defined so that spheres have positive curvature and hyperbolic spaces have negative curvature

$$R^{\lambda}{}_{\mu\rho\nu} = \partial_{\rho}\Gamma^{\lambda}_{\mu\nu} + \Gamma^{\kappa}_{\mu\nu}\Gamma^{\lambda}_{\kappa\rho} - (\nu \leftrightarrow \rho) \tag{A.1}$$

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu} . \tag{A.2}$$

The extrinsic curvature associated with the embedding of $(\partial\mathcal{M}, \gamma)$ in $(\mathcal{M}, g)$ is given by

$$K_{ab} = \frac{1}{2} n^{\mu}\partial_{\mu}\gamma_{ab} \tag{A.3}$$

where n^{μ} is an outward-pointing unit vector normal to $\partial\mathcal{M}$. The trace of the extrinsic curvature is $K = \gamma^{ab}K_{ab}$. Because the boundary is one-dimensional, it follows that

$$K_{ab} - \gamma_{ab} K = 0 . \tag{A.4}$$

This implies that π^{ab} would vanish if not for the coupling between X and the Ricci scalar in (1.1).

Throughout the paper we work with Euclidean metrics of the form

$$ds^2 = \xi(r)\,d\tau^2 + \frac{1}{\xi(r)}\,dr^2 \ . \tag{A.5}$$

Calculating the action and its variation requires the Ricci scalar and the extrinsic curvature of the boundary. The Ricci scalar is given by

$$R = -\frac{\partial^2 \xi}{\partial r^2} = -e^{-Q}\left[w'' + Uw' + U'(w - 2M)\right] \ . \tag{A.6}$$

The last expression, which follows from (2.4) and (2.5), is coordinate independent and relates the Ricci scalar to the dilaton field. For models obeying the Minkowskian ground state condition (4.1) it simplifies to $R = 2M(w'^2/w - w'')$. Boundary quantities are obtained from the $X \to \infty$ limit of a constant X surface, as described in subsection 2.2. The induced metric on the surface is $\gamma_{ab} = \xi(X)$, and the trace of its extrinsic curvature is

$$K = \frac{1}{2\sqrt{\xi}}\frac{\partial \xi}{\partial r} = e^{-Q/2}\frac{(w - 2M)U + w'}{2\sqrt{w - 2M}} \ . \tag{A.7}$$

For models obeying the Minkowskian ground state condition (4.1) it simplifies to $K = Mw'/(w\sqrt{\xi})$. The Euler characteristic for the space-time is related to (A.6) and (A.7) by

$$4\pi\,\chi = \int_{\mathcal{M}} d^2x\sqrt{g}\,R + 2\int_{\partial\mathcal{M}} dx\sqrt{\gamma}\,K \ . \tag{A.8}$$

One can verify that the BH solutions of discussed in 2.2 have the topology of a disk, $\chi = 1$.

The dimensions of various quantities are assigned conveniently, but at times differently, in the examples we consider. Here we clarify the role of physical dimensions, how to fix them, and how to rescale them. Unless stated otherwise we always work in natural units $\hbar = c = k_{\rm B} = 1$ and choose the dilaton field X to be dimensionless. This means that the potential $U(X)$ is dimensionless, as well. If the coordinates are taken to have the natural dimension of length then the relation (2.4) implies that e^Q must also have dimensions of length. This can be accomplished by defining $e^Q = \lambda \exp \int^X d\tilde{X} U(\tilde{X})$, where λ is a dimensionful constant with units of length that includes the factor of e^{Q_0}. In that case, both w and M must have units of inverse-length, so that (2.5) is dimensionless. Similarly, coordinates with units of length imply that the function V has dimensions of one over length squared. Units of length can then be changed by appropriate rescalings of the dimensionful constants in V, as well as λ and M. Alternatively, if the coordinates are taken to be dimensionless then U, e^Q, V, w, and M are all dimensionless, as well.

After these general observations we focus on the choices of relevance for our work. With one exception (mentioned below) we demand that the dilaton field X be dimensionless. In some of our examples we do not exhibit the dimensionful constants, for simplicity. Instead we treat all quantities as dimensionless, including the line-element, and restore appropriate physical dimensions when necessary. This applies to the discussion of the

Model	$U(X)$	$V(X)$	$e^{Q(X)}$	$w(X)$	Reference
Schwarzschild	$-\frac{1}{2X}$	$-\frac{1}{2G_4}$	$\sqrt{\frac{G_4}{2X}}$	$\sqrt{\frac{2X}{G_4}}$	[70]
Jackiw-Teitelboim	0	$-\Lambda X$	1	ΛX^2	[44]
Witten BH	$-\frac{1}{X}$	$-\frac{\lambda^2}{2}X$	$\frac{1}{\lambda X}$	λX	[1]
CGHS	0	$-\frac{\lambda}{2}$	1	λX	[71]
(A)dS$_2$ Ground State	$-\frac{a}{X}$	$-\frac{1}{2}BX$	$\frac{1}{X^a}$	$\frac{B}{2-a}X^{2-a}$	[72]
Rindler Ground State	$-\frac{a}{X}$	$-\frac{1}{2}BX^a$	$\frac{1}{X^a}$	BX	[73]
BH Attractor	0	$-\frac{B}{2X}$	1	$B\ln X$	[74]
ab-Family	$-\frac{a}{X}$	$-\frac{B}{2}X^{a+b}$	$\frac{1}{X^a}$	$\frac{B}{b+1}X^{b+1}$	[43]
Liouville Gravity	a	$be^{\alpha X}$	e^{aX}	$-\frac{2b}{a+\alpha}e^{(a+\alpha)X}$	[75]
Exact String BH	[46]	[46]	(5.16)	(5.16), (5.12)	[2]
Schwarzschild-(A)dS	$-\frac{1}{2X}$	$-\frac{1}{2G_4}-\frac{3}{\ell^2}X$	$\sqrt{\frac{G_4}{2X}}$	$\sqrt{\frac{2X}{G_4}}\left(1+\frac{2G_4}{\ell^2}X\right)$	[41]
Katanaev-Volovich	α	$\beta X^2-\lambda$	$e^{\alpha X}$	$\frac{2}{\alpha}e^{\alpha X}\left(\lambda-\beta X^2+\right.$ $\left.+2\frac{\beta}{\alpha}X-2\frac{\beta}{\alpha^2}\right)$	[76]
KK Reduced CS	0	$\frac{1}{2}X\left(c-X^2\right)$	1	$\frac{1}{4}\left(c-X^2\right)^2-\frac{1}{4}c^2$	[60, 61]

Table 3: Summary of Dilaton Gravity models, adapted from [34].

Model	$U(X)$	$V(X)$	$f(X)$	$e^{Q(X)}$	$w(X)$	$h(X)$	Reference
Reissner-Nordström	$-\frac{1}{2X}$	$-\frac{1}{2G_4}$	X	$\sqrt{\frac{G_4}{2X}}$	$\sqrt{\frac{2X}{G_4}}$	$-\sqrt{\frac{2G_4}{X}}$	[77]
Achucarro-Ortiz	0	$-\frac{X}{\ell^2}$	$-\frac{X^3}{8}$	1	$\frac{X^2}{\ell^2}$	$\frac{4}{X^2}$	[56]
2D Type 0A/0B	$-\frac{1}{X}$	$-\frac{\lambda^2}{2}X$	$\pi\alpha'$	$\frac{1}{\lambda X}$	λX	$\frac{\ln X}{\lambda\pi\alpha'}$	[78]

Table 4: Dilaton Gravity models with Maxwell fields.

ab-family in subsection (4.2), and to the discussion of the BTZ BH in subsection (6.3) if X, ℓ, J, M are considered to be dimensionless. In other examples we keep track of the physical dimensions for easier comparison with the literature. This applies to sections 5, 6.1, and 6.2. Alternatively, in subsection 6.3 the quantities X, ℓ, J can be thought of as having dimensions of length, with M dimensionless.

The analysis of sections 2 and 3 is carried out for a general model of two-dimensional dilaton gravity. As a convenience, the properties of several models are summarized in tables 3 and 4. Table 3 lists models that contain only a metric and a dilaton, while table 4 lists models that also contain Maxwell fields. The last three models in table 3 can be converted into models belonging to table 4 as explained at the end of subsection 3.8. In most entries in these tables we have chosen the scaling and shift ambiguities inherent to Q and w as to achieve coincidence with the choices in the main text. In some brief examples

in our paper the four-dimensional Newton constant is fixed as $G_4 = 1/2$, thus simplifying the first entries in both tables, as well as the Schwarzschild-(A)dS entry in table 3. In subsection 4.2 we set $B = b + 1$, which simplifies the corresponding entry 'ab-Family' in table 3.

B. Weyl rescaled metric

In subsections 2.3 and 2.5 we employ a change of variables that involves a Weyl rescaling of the metric. Then, in subsection 3.9, we consider the effect of a conformal transformation on a model. In this appendix we work out some of the relevant details.

Consider a new metric related to $g_{\mu\nu}$ by

$$g_{\mu\nu} = e^{2\sigma(X)} \, \hat{g}_{\mu\nu} \tag{B.1}$$

where $\sigma(X)$ is an arbitrary function of the dilaton. If we express the action (1.1) in terms of this new metric we find

$$I[X, g_{\mu\nu}] = -\frac{1}{2} \int_{\mathcal{M}} d^2x \sqrt{\hat{g}} \left[X\hat{R} - \hat{U}(X)(\hat{\nabla}X)^2 - 2\hat{V}(X) \right] - \int_{\partial\mathcal{M}} dx \sqrt{\hat{\gamma}} X\hat{K} . \tag{B.2}$$

This has the same functional form as (1.1), with dilaton potentials $\hat{U}$ and $\hat{V}$ that are related to U and V by

$$\hat{U}(X) = U(X) - 2\sigma'(X) \qquad \hat{V}(X) = e^{2\sigma(X)} V(X) . \tag{B.3}$$

The functions $\hat{Q}(X)$ and $\hat{w}(X)$ transform according to

$$\hat{Q}(X) = Q(X) - 2\sigma(X) \qquad \hat{w}(X) = w(X) . \tag{B.4}$$

A change of variables with $2\sigma(X) = Q(X)$ leads to vanishing $\hat{U}$. This transformation is common in the dilaton gravity literature because it simplifies many calculations. Notice that $w(X)$ is invariant under (B.1). Referring to (2.15), which depends only on $w(X)$, X, and the parameter M, we confirm that the on-shell action is invariant. This is as it should be: the value of the action cannot depend on the variables we use to evaluate it.

Now we want to examine the variation of the action using the new metric variable. In terms of $\hat{g}_{\mu\nu}$ and σ a general variation of $g_{\mu\nu}$ takes the form

$$\delta g_{\mu\nu} = e^{2\sigma} \delta\hat{g}_{\mu\nu} + 2 e^{2\sigma} \hat{g}_{\mu\nu}\sigma'\delta X . \tag{B.5}$$

The change in the action (2.16) due to a small variation of the fields can be written as

$$\delta I[X, g_{\mu\nu}] = \int_{\mathcal{M}} d^2x \sqrt{\hat{g}} \left[\hat{\mathcal{E}}^{\mu\nu}\delta\hat{g}_{\mu\nu} + \hat{\mathcal{E}}_X \, \delta X \right] + \int_{\partial\mathcal{M}} dx \sqrt{\hat{\gamma}} \left[\hat{\pi}^{ab}\delta\hat{\gamma}_{ab} + \hat{\pi}_X\delta X \right] . \tag{B.6}$$

Both the equations of motion $\hat{\mathcal{E}}_{\mu\nu}$ and $\hat{\mathcal{E}}_X$, and the momenta $\hat{\pi}^{ab}$ and $\hat{\pi}_X$ have the same functional form as in the main text, with all un-hatted quantities being replaced by their hatted counterparts. The solution of the equations of motion is

$$X = X(\hat{r}) \qquad\qquad d\hat{s}^2 = \hat{\xi}(\hat{r}) \, d\tau^2 + \frac{1}{\hat{\xi}(\hat{r})} \, d\hat{r}^2 \tag{B.7}$$

$$\partial_{\hat{r}}X = e^{-\hat{Q}(X)} \qquad\qquad \hat{\xi}(X) = e^{\hat{Q}(X)} \, (w(X) - 2M) . \tag{B.8}$$

This is just the solution from section 2.1, expressed in terms of the new metric and written in a coordinate system $(\tau, \hat{r})$ with $\partial_{\hat{r}} = e^{2\sigma(X)} \partial_r$. Evaluating (B.6) for this solution gives

$$\delta I[X, g_{\mu\nu}]\Big|_{\hat{\mathcal{E}}=0} = \int_{\partial\mathcal{M}} dx \left[-\frac{1}{2} e^{-\hat{Q}} \delta\hat{\xi} + e^{-\hat{Q}} \left(\hat{U}\hat{\xi} - \frac{1}{2} \hat{\xi}' \right) \delta X \right] . \tag{B.9}$$

We can now choose the function $\sigma(X)$ in (B.1) to simplify the analysis of this expression. If we define

$$e^{2\sigma(X)} = w(X) \, e^{Q(X)} , \tag{B.10}$$

then the (square of the) Killing norm for the new metric is given by

$$\hat{\xi}(X) = 1 - \frac{2M}{w(X)} . \tag{B.11}$$

This makes the meaning of Dirichlet boundary conditions unambiguous: $\hat{\xi}$ takes the constant value $\hat{\xi} = 1$ at $\partial\mathcal{M}$, and any variation $\delta\hat{\xi}$ that preserves this condition must vanish at $\partial\mathcal{M}$. The functions $\hat{Q}$ and $\hat{U}$ are determined as

$$e^{\hat{Q}(X)} = \frac{1}{w(X)} \qquad \hat{U} = -\frac{w'(X)}{w(X)} \tag{B.12}$$

so that (B.9) yields

$$\delta I_{\text{reg}} = \int d\tau \left[-\frac{1}{2} w \, \delta\hat{\xi} - \left(1 - \frac{M}{w} \right) w' \, \delta X \right] . \tag{B.13}$$

The subscript indicates that this quantity is evaluated using a regulator, as in section 2.3. Following the arguments in the main text we ignore the δX term and concentrate on a variation of the metric $\delta\hat{\xi} = \delta M e^{\hat{Q}} = \delta M / w$. The corresponding change in the action is

$$\lim_{X_{\text{reg}} \to \infty} \delta I_{\text{reg}} = \int d\tau \, \delta M \neq 0 \tag{B.14}$$

which of course agrees with (2.20). Changing variables to a metric that obeys standard Dirichlet conditions at $\partial\mathcal{M}$ simplifies the analysis of the improved action, as well, and is useful in establishing the result (2.37).

In subsection 3.9 we treat (B.1) as a Weyl transformation, as opposed to a change of variable. Then the action (B.2) is interpreted as a functional of X and $\hat{g}_{\mu\nu}$. This gives a new model of type (1.1) with potentials given by (B.3).

C. Proof of gauge-independence

In this appendix we show that the main results of section 2 are independent of the choice of gauge (2.3). We begin by establishing two important properties of solutions of (2.1) and (2.2). Consider the vector k^μ defined by

$$k^\mu = \epsilon^{\mu\nu} e^{Q(X)} \nabla_\nu X \tag{C.1}$$

where $e^{Q(X)}$ was given in (2.6). The Lie derivative of the dilaton along k^μ vanishes by construction

$$\mathcal{L}_k X = e^{Q(X)} \epsilon^{\mu\nu} \, \partial_\mu X \partial_\nu X = 0 \ . \tag{C.2}$$

The same result also applies to the metric. It follows from the equation of motion (2.1) that the vector k^μ satisfies Killing's equation

$$\mathcal{L}_k g_{\mu\nu} = \nabla_\mu k_\nu + \nabla_\nu k_\mu = 0 \ . \tag{C.3}$$

Therefore, solutions of (2.1) and (2.2) always possess at least one Killing vector whose orbits are isosurfaces of the dilaton.

The properties of k^μ allow us to construct a coordinate system (τ, r) where the Killing vector is $k^\mu \partial_\mu = \partial_\tau$, and the metric and dilaton take the form

$$ds^2 = N(r)^2 \, dr^2 + \xi(r) \, (d\tau + N^\tau(r) \, dr)^2 \qquad X = X(r) \ . \tag{C.4}$$

The functions N and N^τ that appear in the metric are the lapse function and shift 'vector', respectively, associated with evolution in the direction orthogonal to k^μ. Normally, the lapse and shift both appear in the action as Lagrange multipliers. In this case, the fact that all fields are constant along the direction singled out by the Killing vector implies that the action is independent of $N^\tau(r)$. Furthermore, the coefficient of δN^τ in the variation of the action vanishes identically. This means that we can freely set $N^\tau = 0$ in (C.4) with no loss of generality. The metric then simplifies to

$$ds^2 = N(r)^2 \, dr^2 + \xi(r) \, d\tau^2 \ . \tag{C.5}$$

One can substitute this form of the metric into the action and recover all of the results of a fully covariant analysis. From the point of view of the action, the functions $\xi(r)$ and $N(r)$ must be treated as independent fields. However, we are always free to present a solution in a coordinate system where $N(r)^2 = \xi(r)^{-1}$, as in (2.3).

Now we want to check whether or not the conclusions of subsection 2.3 are sensitive to the choice of gauge. Integrating the equations of motion for an arbitrary lapse function, we find

$$\partial_r X = N(r) \sqrt{\xi(X)} \, e^{-Q(X)}$$
$$\xi(X) = w(X) \, e^{Q(X)} \left(1 - \frac{2M}{w(X)} \right) \ . \tag{C.6}$$

The only difference between these expressions and the solution (2.4) and (2.5) is an overall factor that appears in $\partial_r X$. Notably, we obtain the same expression for $\xi(X)$, concurrent with its coordinate independent interpretation as square of the Killing norm. The change in the action due to a variation in the metric, with the bulk terms set to zero, is now given by

$$\delta I_{\text{reg}} = -\frac{1}{2} \int d\tau \, \frac{\partial_r X}{N(r) \sqrt{\xi(r)}} \, \delta \xi(r) \ . \tag{C.7}$$

The subscript is a reminder that the integrand is evaluated using a regulator. This result can be obtained either by evaluating the covariant expression (2.16) for the solution (C.6),

or by expressing the action (1.1) in terms of $\xi(r)$ and $N(r)$ and then varying the fields. It is clear from (C.6) that the explicit factor of $N(r)$ in (C.7) cancels against the factor in $\partial_r X$, and we recover the result (2.20). Therefore, the conclusions of subsection 2.3 are independent of the choice of gauge. This exercise can be repeated for the variation of the action Γ from (2.30). Regardless of the choice of gauge, $\delta\Gamma = 0$ for all variations of the fields that preserve the boundary conditions.

D. Formulas

The variations of X_h and T can be expressed entirely in terms of dM, using the definitions of the horizon (2.11) and T (2.13)

$$dX_h = \frac{1}{2\pi T}\, dM \qquad dT = \frac{1}{2\pi T}\, \frac{\partial T}{\partial X_h}\, dM \ . \tag{D.1}$$

Similarly, the definitions of ξ (2.5), Q (2.6), and w (2.7) give

$$d\xi_c = -2\, e^{Q_c}\, dM + \left(U_c\, \xi_c + e^{Q_c}\, w'_c \right) dD_c \ . \tag{D.2}$$

When D_c is held fixed the variation of ξ_c reduces to

$$d\xi_c\Big|_{D_c} = -2\, e^{Q_c}\, dM \ . \tag{D.3}$$

A useful consequence is

$$\frac{\partial\xi_c}{\partial T}\Big|_{D_c} = -4\pi e^{Q_c}\frac{w'_h}{w''_h} \ . \tag{D.4}$$

Holding the temperature T_c fixed implies a relationship between dM and dD_c, which can be obtained using the definition (2.14), and the relations (D.1) and (D.2):

$$\left(\frac{1}{2\pi}\, \frac{\partial T}{\partial X_h} + e^{Q_c}\, T_c^2 \right) dM\Big|_{T_c} = \frac{1}{2}\, T_c^2\, \left(U_c\, \xi_c + e^{Q_c}\, w'_c \right) dD_c\Big|_{T_c} \ . \tag{D.5}$$

Using (3.5) and the formulas above we obtain

$$\frac{\partial\psi_c}{\partial D_c}\Big|_{T_c} = -\left(f_c + g_c\sqrt{\xi_c} + h_c\frac{1}{\sqrt{\xi_c}} \right) \tag{D.6}$$

with

$$\begin{aligned}
f &= \frac{e^{-Q/2}}{2\sqrt{w}}\left[w'' - \frac{w'^2}{2w} - (wU)' + \frac{1}{2}wU^2 \right]\\
g &= \frac{1}{2}e^{-Q}\left(U' - U^2 + UN \right)\\
h &= -\frac{1}{2}w'' + \frac{1}{2}w'N
\end{aligned} \tag{D.7}$$

where

$$N = -w''_h\, \frac{(2M - w)U - w'}{w'^2_h - 2w''_h(2M - w)} \ . \tag{D.8}$$

– 57 –

The relations (D.4) and

$$\frac{\partial \xi_c}{\partial T_c}\bigg|_{D_c} = \frac{\partial \xi_c}{\partial T}\bigg|_{D_c} \sqrt{\xi_c} \left(1 - \frac{T}{2\xi_c} \frac{\partial \xi_c}{\partial T}\bigg|_{D_c}\right)^{-1} \tag{D.9}$$

allow to determine

$$\frac{\partial \psi_c}{\partial T_c}\bigg|_{D_c} = -\frac{1}{4} \xi_c^{-3/2} \left(w_c' + \xi_c U_c e^{-Q_c}\right) \frac{\partial \xi_c}{\partial T_c}\bigg|_{D_c} . \tag{D.10}$$

Another useful formula is

$$T_c \frac{\partial S}{\partial T_c}\bigg|_{\psi_c} = -T_c \frac{\frac{\partial \psi_c}{\partial T_c}\big|_S}{\frac{\partial \psi_c}{\partial S}\big|_{T_c}} . \tag{D.11}$$

References

[1] E. Witten, *On string theory and black holes*, *Phys. Rev.* **D 44** (1991) 314.

[2] R. Dijkgraaf, H.L. Verlinde and E.P. Verlinde, *String propagation in a black hole geometry*, *Nucl. Phys.* **B 371** (1992) 269.

[3] D. Grumiller, W. Kummer and D.V. Vassilevich, *Dilaton gravity in two dimensions*, *Phys. Rept.* **369** (2002) 327 [hep-th/0204253].

[4] J. York, James W., *Role of conformal three geometry in the dynamics of gravitation*, *Phys. Rev. Lett.* **28** (1972) 1082.

[5] G.W. Gibbons and S.W. Hawking, *Action integrals and partition functions in quantum gravity*, *Phys. Rev.* **D 15** (1977) 2752.

[6] H. Liebl, D.V. Vassilevich and S. Alexandrov, *Hawking radiation and masses in generalized dilaton theories*, *Class. and Quant. Grav.* **14** (1997) 889 [gr-qc/9605044].

[7] T. Regge and C. Teitelboim, *Role of surface integrals in the hamiltonian formulation of general relativity*, *Ann. Phys. (NY)* **88** (1974) 286.

[8] R.B. Mann and D. Marolf, *Holographic renormalization of asymptotically flat spacetimes*, *Class. and Quant. Grav.* **23** (2006) 2927 [hep-th/0511096].

[9] D.J. Gross, M.J. Perry and L.G. Yaffe, *Instability of flat space at finite temperature*, *Phys. Rev.* **D 25** (1982) 330.

[10] J.L. Davis and R. McNees, *Boundary counterterms and the thermodynamics of 2D black holes*, *JHEP* **09** (2005) 072 [hep-th/0411121].

[11] J. York, James W., *Black hole thermodynamics and the euclidean Einstein action*, *Phys. Rev.* **D 33** (1986) 2092.

[12] G.W. Gibbons and M.J. Perry, *The physics of 2-D stringy space-times*, *Int. J. Mod. Phys.* **D 1** (1992) 335 [hep-th/9204090].

[13] V.P. Frolov, *Two-dimensional black hole physics*, *Phys. Rev.* **D 46** (1992) 5383.

[14] T. Klosch and T. Strobl, *Classical and quantum gravity in (1 + 1)-dimensions. Part 1: a unifying approach*, *Class. and Quant. Grav.* **13** (1996) 965 [gr-qc/9508020].

[15] M. Bañados, C. Teitelboim and J. Zanelli, *The black hole in three-dimensional space-time*, *Phys. Rev. Lett.* **69** (1992) 1849 [hep-th/9204099]; *Geometry of the* $(2 + 1)$ *black hole*, *Phys. Rev.* **D 48** (1993) 1506 [gr-qc/9302012].

[16] H.-J. Schmidt, *Scale invariant gravity in two-dimensions*, *J. Math. Phys.* **32** (1991) 1562.

[17] T. Banks and M. O'Loughlin, *Two-dimensional quantum gravity in minkowski space*, *Nucl. Phys.* **B 362** (1991) 649.

[18] R.C. Tolman, *Relativity, thermodynamics, and cosmology*, Oxford University Press, Oxford (1934).

[19] I. Papadimitriou and K. Skenderis, *Thermodynamics of asymptotically locally AdS spacetimes*, *JHEP* **08** (2005) 004 [hep-th/0505190].

[20] R. McNees, *A new boundary counterterm for asymptotically AdS spacetimes*, hep-th/0512297.

[21] M. Henningson and K. Skenderis, *The holographic Weyl anomaly*, *JHEP* **07** (1998) 023 [hep-th/9806087].

[22] V. Balasubramanian and P. Kraus, *A stress tensor for anti-de Sitter gravity*, *Commun. Math. Phys.* **208** (1999) 413 [hep-th/9902121].

[23] R. Emparan, C.V. Johnson and R.C. Myers, *Surface terms as counterterms in the AdS/CFT correspondence*, *Phys. Rev.* **D 60** (1999) 104001 [hep-th/9903238].

[24] S. de Haro, S.N. Solodukhin and K. Skenderis, *Holographic reconstruction of spacetime and renormalization in the AdS/CFT correspondence*, *Commun. Math. Phys.* **217** (2001) 595 [hep-th/0002230].

[25] R.B. Mann, D. Marolf and A. Virmani, *Covariant counterterms and conserved charges in asymptotically flat spacetimes*, *Class. and Quant. Grav.* **23** (2006) 6357 [gr-qc/0607041].

[26] D. Marolf and L. Patino, *The non-zero energy of* $2 + 1$ *Minkowski space*, *Phys. Rev.* **D 74** (2006) 024009 [hep-th/0604127].

[27] D. Astefanesei and E. Radu, *Quasilocal formalism and black ring thermodynamics*, *Phys. Rev.* **D 73** (2006) 044014 [hep-th/0509144].

[28] D. Astefanesei, R.B. Mann and C. Stelea, *Note on counterterms in asymptotically flat spacetimes*, *Phys. Rev.* **D 75** (2007) 024007 [hep-th/0608037].

[29] R. Olea, *Regularization of odd-dimensional AdS gravity: Kounterterms*, hep-th/0610230.

[30] D. Marolf and A. Virmani, *Holographic renormalization of gravity in little string theory duals*, hep-th/0703251.

[31] T.H. Buscher, *A symmetry of the string background field equations*, *Phys. Lett.* **B 194** (1987) 59; *Path integral derivation of quantum duality in nonlinear sigma models*, *Phys. Lett.* **B 201** (1988) 466.

[32] C.R. Nappi and A. Pasquinucci, *Thermodynamics of two-dimensional black holes*, *Mod. Phys. Lett.* **A 7** (1992) 3337 [gr-qc/9208002].

[33] J. Gegenberg, G. Kunstatter and D. Louis-Martinez, *Observables for two-dimensional black holes*, *Phys. Rev.* **D 51** (1995) 1781 [gr-qc/9408015].

[34] D. Grumiller and R. Meyer, *Ramifications of lineland*, *Turk. J. Phys.* **30** (2006) 349 [hep-th/0604049].

[35] J.D. Brown and J. York, James W., *Quasilocal energy in general relativity*, gr-qc/9209012; *Quasilocal energy and conserved charges derived from the gravitational action*, *Phys. Rev.* **D 47** (1993) 1407.

[36] G.W. Gibbons, M.J. Perry and C.N. Pope, *The first law of thermodynamics for Kerr-anti-de Sitter black holes*, *Class. and Quant. Grav.* **22** (2005) 1503 [hep-th/0408217].

[37] G. Ruppeiner, *Riemannian geometry in thermodynamic fluctuation theory*, *Rev. Mod. Phys.* **67** (1995) 605 [Erratum *ibid.* **68** (1996) 313].

[38] G. Arcioni and E. Lozano-Tellechea, *Stability and critical phenomena of black holes and black rings*, *Phys. Rev.* **D 72** (2005) 104021 [hep-th/0412118]; J.-y. Shen, R.-G. Cai, B. Wang and R.-K. Su, *Thermodynamic geometry and critical behavior of black holes*, *Int. J. Mod. Phys.* **A 22** (2007) 11 [gr-qc/0512035]; J.E. Aman, I. Bengtsson and N. Pidokrajt, *Flat information geometries in black hole thermodynamics*, *Gen. Rel. Grav.* **38** (2006) 1305 [gr-qc/0601119]; T. Sarkar, G. Sengupta and B. Nath Tiwari, *On the thermodynamic geometry of btz black holes*, *JHEP* **11** (2006) 015 [hep-th/0606084].

[39] J.E. Aman, J. Bedford, D. Grumiller, N. Pidokrajt and J. Ward, *Ruppeiner theory of black hole thermodynamics*, gr-qc/0611119.

[40] D. Birmingham, M. Blau, M. Rakowski and G. Thompson, *Topological field theory*, *Phys. Rept.* **209** (1991) 129.

[41] S.W. Hawking and D.N. Page, *Thermodynamics of black holes in anti-de sitter space*, *Commun. Math. Phys.* **87** (1983) 577.

[42] L. Bergamin, D. Grumiller and W. Kummer, *Supersymmetric black holes are extremal and bald in 2-D dilaton supergravity*, *J. Phys.* **A 37** (2004) 3881 [hep-th/0310006].

[43] M.O. Katanaev, W. Kummer and H. Liebl, *On the completeness of the black hole singularity in 2D dilaton theories*, *Nucl. Phys.* **B 486** (1997) 353 [gr-qc/9602040].

[44] R. Jackiw and C. Teitelboim in *Quantum Theory Of Gravity*, S. Christensen ed., Adam Hilger, Bristol (1984).

[45] S. Elitzur, A. Forge and E. Rabinovici, *Some global aspects of string compactifications*, *Nucl. Phys.* **B 359** (1991) 581; G. Mandal, A.M. Sengupta and S.R. Wadia, *Classical solutions of two-dimensional string theory*, *Mod. Phys. Lett.* **A 6** (1991) 1685.

[46] D. Grumiller, *An action for the exact string black hole*, *JHEP* **05** (2005) 028 [hep-th/0501208].

[47] V.A. Kazakov and A.A. Tseytlin, *On free energy of 2-D black hole in bosonic string theory*, *JHEP* **06** (2001) 021 [hep-th/0104138].

[48] D. Grumiller, *Logarithmic corrections to the entropy of the exact string black hole*, hep-th/0506175.

[49] C. Ford, G. Jorjadze and G. Weigt, *Integration of the* SL(2, ℝ)/u(1) *gauged WZNW theory by reduction and quantum parafermions*, *Theor. Math. Phys.* **128** (2001) 1046 [hep-th/0003246]; C. Kruger, *Exact operator quantization of the euclidean black hole CFT*, hep-th/0411275.

[50] D. Grumiller, W. Kummer and D.V. Vassilevich, *Positive specific heat of the quantum corrected dilaton black hole*, *JHEP* **07** (2003) 009 [hep-th/0305036].

[51] P. Kraus, F. Larsen and R. Siebelink, *The gravitational action in asymptotically AdS and flat spacetimes*, *Nucl. Phys.* **B 563** (1999) 259 [hep-th/9906127].

[52] J.D. Brown et al., *Thermodynamic ensembles and gravitation*, *Class. and Quant. Grav.* **7** (1990) 1433;
J.D. Brown and J. York, James W., *The path integral formulation of gravitational thermodynamics*, gr-qc/9405024;
J.D. Brown, S.R. Lau and J. York, James W., *Action and energy of the gravitational field*, gr-qc/0010024.

[53] J.D. Brown, J. Creighton and R.B. Mann, *Temperature, energy and heat capacity of asymptotically anti-de Sitter black holes*, *Phys. Rev.* **D 50** (1994) 6394 [gr-qc/9405007].

[54] M.M. Akbar, *Schwarzschild-anti de Sitter within an isothermal cavity: thermodynamics, phase transitions and the Dirichlet problem*, hep-th/0401228.

[55] S. Hollands, A. Ishibashi and D. Marolf, *Comparison between various notions of conserved charges in asymptotically AdS-spacetimes*, *Class. and Quant. Grav.* **22** (2005) 2881 [hep-th/0503045].

[56] A. Achucarro and M.E. Ortiz, *Relating black holes in two-dimensions and three-dimensions*, *Phys. Rev.* **D 48** (1993) 3600 [hep-th/9304068].

[57] S. Carlip, *What we don't know about BTZ black hole entropy*, *Class. and Quant. Grav.* **15** (1998) 3609 [hep-th/9806026].

[58] O.B. Zaslavsky, *Geometry of nonextreme black holes near the extreme state*, *Phys. Rev.* **D 56** (1997) 2188 [gr-qc/9707015];
A.J.M. Medved and G. Kunstatter, *One-loop corrected thermodynamics of the extremal and non-extremal spinning BTZ black hole*, *Phys. Rev.* **D 63** (2001) 104005 [hep-th/0009050];
J.M. Izquierdo, J. Navarro-Salas and P. Navarro, *Kaluza-Klein theory, AdS/CFT correspondence and black hole entropy*, *Class. and Quant. Grav.* **19** (2002) 563 [hep-th/0107132].

[59] S. Deser, R. Jackiw and S. Templeton, *Three-dimensional massive gauge theories*, *Phys. Rev. Lett.* **48** (1982) 975; *Topologically massive gauge theories*, *Ann. Phys. (NY)* **140** (1982) 372 [Erratum *ibid.* **185** (1988) 406].

[60] G. Guralnik, A. Iorio, R. Jackiw and S.Y. Pi, *Dimensionally reduced gravitational Chern-Simons term and its kink*, *Ann. Phys. (NY)* **308** (2003) 222 [hep-th/0305117].

[61] D. Grumiller and W. Kummer, *The classical solutions of the dimensionally reduced gravitational Chern-Simons theory*, *Ann. Phys. (NY)* **308** (2003) 211 [hep-th/0306036].

[62] B. Sahoo and A. Sen, *BTZ black hole with Chern-Simons and higher derivative terms*, *JHEP* **07** (2006) 008 [hep-th/0601228].

[63] R.M. Wald, *Black hole entropy in the Nöther charge*, *Phys. Rev.* **D 48** (1993) 3427 [gr-qc/9307038].

[64] D. Grumiller and R. Jackiw, *Duality in 2-dimensional dilaton gravity*, *Phys. Lett.* **B 642** (2006) 530 [hep-th/0609197].

[65] W. Kummer, H. Liebl and D.V. Vassilevich, *Exact path integral quantization of generic 2-D dilaton gravity*, *Nucl. Phys.* **B 493** (1997) 491 [gr-qc/9612012];
L. Bergamin, D. Grumiller, W. Kummer and D.V. Vassilevich, *Classical and quantum integrability of 2D dilaton gravities in euclidean space*, *Class. and Quant. Grav.* **22** (2005) 1361 [hep-th/0412007].

[66] J. de Boer, E.P. Verlinde and H.L. Verlinde, *On the holographic renormalization group*, *JHEP* **08** (2000) 003 [hep-th/9912012];
J. de Boer, *The holographic renormalization group*, *Fortschr. Phys.* **49** (2001) 339 [hep-th/0101026].

[67] D. Martelli and W. Muck, *Holographic renormalization and Ward identities with the Hamilton-Jacobi method*, *Nucl. Phys.* **B 654** (2003) 248 [hep-th/0205061].

[68] F. Larsen and R. McNees, *Holography, diffeomorphisms and scaling violations in the CMB*, *JHEP* **07** (2004) 062 [hep-th/0402050].

[69] A. Batrachenko, J.T. Liu, R. McNees, W.A. Sabra and W.Y. Wen, *Black hole mass and Hamilton-Jacobi counterterms*, *JHEP* **05** (2005) 034 [hep-th/0408205].

[70] K. Schwarzschild, *On the gravitational field of a mass point according to Einstein's theory*, *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **189** (1916) [physics/9905030].

[71] C.G. Callan Jr., S.B. Giddings, J.A. Harvey and A. Strominger, *Evanescent black holes*, *Phys. Rev.* **D 45** (1992) 1005 [hep-th/9111056].

[72] J.P.S. Lemos and P.M. Sa, *The black holes of a general two-dimensional dilaton gravity theory*, *Phys. Rev.* **D 49** (1994) 2897 [gr-qc/9311008].

[73] A. Fabbri and J.G. Russo, *Soluble models in 2D dilaton gravity*, *Phys. Rev.* **D 53** (1996) 6995 [hep-th/9510109].

[74] D. Grumiller, *Long time black hole evaporation with bounded Hawking flux*, *JCAP* **05** (2004) 005 [gr-qc/0307005].

[75] Y. Nakayama, *Liouville field theory: a decade after the revolution*, *Int. J. Mod. Phys.* **A 19** (2004) 2771 [hep-th/0402009].

[76] M.O. Katanaev and I.V. Volovich, *String model with dynamical geometry and torsion*, *Phys. Lett.* **B 175** (1986) 413 [hep-th/0209014].

[77] H. Reissner, *Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie*, *Ann. Phys. (NY)* **50** (1916) 106;
G. Nordström, *On the energy of the gravitation field in Einstein's theory*, *Proc. Kon. Ned. Akad. Wet.* **20** (1916) 1238.

[78] N. Berkovits, S. Gukov and B.C. Vallilo, *Superstrings in 2D backgrounds with RR flux and new extremal black holes*, *Nucl. Phys.* **B 614** (2001) 195 [hep-th/0107140].