University of California
Santa Barbara

Topics in Quantum Gravity and Field Theory

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Benjamin Lewis Michel

Committee in charge:

Professor Joseph Polchinski, Chair
Professor Mark Srednicki
Professor David Stuart

September 2017

The Dissertation of Benjamin Lewis Michel is approved.

Professor Mark Srednicki

Professor David Stuart

Professor Joseph Polchinski, Committee Chair

September 2017

Topics in Quantum Gravity and Field Theory

by

Benjamin Lewis Michel

To my grandparents:

Bob and Joan,
Vivi and Herb.

Acknowledgements

If I attempted a proper acknowledgement of everyone who has helped me to this point it would be very long, mostly boring and woefully incomplete. Instead, I will just say that I owe many deep debts of gratitude: to the hep-th community at UCSB and beyond, for wonderful interactions; to my collaborators and co-conspirators, especially my academic siblings; to my cohort, and to old friends; to my family, and to Sarah; to Joe, for some of the most thrilling years of my intellectual life, and to Mark, for his steady wisdom.

Curriculum Vitæ

Benjamin Lewis Michel

Education

2017	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2014	M.A. in Physics, University of California, Santa Barbara.
2010	A.B. in Physics, Harvard University.

Publications

- F. Chen, B. Michel, J. Polchinski and A. Puhm, *Journey to the Center of the Fuzzball*, JHEP **02** (2015) 081 doi:10.1007/JHEP02(2015)081 [arXiv:1408.4798 [hep-th]]
- B. Michel, E. Mintun, J. Polchinski, A. Puhm and P. Saad, *Remarks on brane and antibrane dynamics*, JHEP **1509**, 021 (2015) doi:10.1007/JHEP09(2015)021 [arXiv:1412.5702 [hep-th]].
- B. Michel, J. Polchinski, V. Rosenhaus and S. J. Suh, *Four-point function in the IOP matrix model*, JHEP **1605**, 048 (2016) doi:10.1007/JHEP05(2016)048 [arXiv:1602.06422 [hep-th]].
- A. Belin, J. de Boer, J. Kruthoff, B. Michel, E. Shaghoulian and M. Shyani, *Universality of sparse $d > 2$ conformal field theory at large N*, JHEP **1703**, 067 (2017) doi:10.1007/JHEP03(2017)067 [arXiv:1610.06186 [hep-th]].
- W. Donnelly, B. Michel, D. Marolf and J. Wien, *Living on the Edge: A Toy Model for Holographic Reconstruction of Algebras with Centers*, JHEP **1704**, 093 (2017) doi:10.1007/JHEP04(2017)093 [arXiv:1611.05841 [hep-th]].
- W. Donnelly, B. Michel, and A. Wall, *Electromagnetic Duality and Entanglement Anomalies*, *Phys. Rev.* **D96** no. 4 (2017) 045008 10.1103/PhysRevD.96.045008 [arXiv:1611.0592 [hep-th]].
- B. Michel, *Remarks on Rindler Quantization*, arXiv:1612.03158 [hep-th].
- D. Marolf, B. Michel and A. Puhm, *A rough end for smooth microstate geometries*, JHEP **1705**, 021 (2017) doi:10.1007/JHEP05(2017)021 [arXiv:1612.05235 [hep-th]].
- B. Michel and M. Srednicki, *Entanglement Entropy and Boundary Conditions in 1+1 Dimensions*, arXiv:1612.08682 [hep-th].

Abstract

Topics in Quantum Gravity and Field Theory

by

Benjamin Lewis Michel

This dissertation addresses a variety of open questions in quantum field theory and quantum gravity. The work fits broadly into two categories: attempts to study black holes and brane dynamics in models of quantum gravity, and attempts to study the entangling surface in quantum field theory.

Contents

Chapter 1

Introduction

This dissertation studies a number of open problems in quantum gravity and quantum field theory. In quantum gravity the questions largely center around black holes and brane dynamics while in field theory the main focus is entanglement entropy, with particular attention to the entangling surface. More details are given in the relevant sections.

1.1 Permissions and Attributions

1. The content of § 2.2 is the result of two collaborations: one with Fang Chen, Joseph Polchinski and Andrea Puhm; the other with Donald Marolf and Andrea Puhm. Both have previously appeared in the Journal of High Energy Physics (JHEP) [57, 319]. They are reproduced here under the terms of the Creative Commons Attribution Noncommercial License: `https://creativecommons.org/licenses/by-nc/3.0/us/`.

2. The content of § 2.3 is the result of a collaboration with Joseph Polchinski, Vladimir Rosenhaus and S. Josephine Suh, and has previously appeared in the Journal of High Energy Physics (JHEP) [315]. It is reproduced here under the terms of the Creative Commons Attribution Noncommercial License: `https://creativecommons.org/licenses/by-nc/3.0/us/`.

3. The content of § 2.4 is the result of a collaboration with Eric Mintun, Joseph Polchinski, Andrea Puhm and Philip Saad, and has previously appeared in the Journal of High Energy Physics (JHEP) [314]. It is reproduced here under the terms of the Creative Commons Attribution Noncommercial License: `https://creativecommons.org/licenses/by-nc/3.0/us/`.

4. The content of § 2.5 is the result of a collaboration with Alexandre Belin, Jan de Boer, Jorrit Kruthoff, Edgar Shaghoulian and Milind Shyani, and has previously appeared in the Journal of High Energy Physics (JHEP) [316]. It is reproduced here under the terms of the Creative Commons Attribution Noncommercial License: `https://creativecommons.org/licenses/by-nc/3.0/us/`.

5. The content of § 3.2 is the result of a collaboration with Mark Srednicki and is available as a preprint on arXiv.org [320]. It is reproduced here under the terms of their non-exclusive distribution license: `https://arxiv.org/licenses/nonexclusive-distrib/1.0/license.html`.

6. The content of § 3.3 is the result of a collaboration with William Donnelly and Aron C. Wall, and has previously appeared in Phys. Rev. D [222]. It is reproduced here with permission from the publisher, the American Phys-

ical Society: http://publish.aps.org/copyrightFAQ.html#thesis. See http://forms.aps.org/author/copytrnsfr.pdf for the official copyright transfer agreement.

7. The content of § 3.4 is available as a preprint on arXiv.org [318]. It is reproduced here under the terms of their non-exclusive distribution license: https://arxiv.org/licenses/nonexclusive-distrib/1.0/license.html.

8. The content of § 3.5 is the result of a collaboration with William Donnelly, Donald Marolf and Jason Wien, and has previously appeared in the Journal of High Energy Physics (JHEP) [317]. It is reproduced here under the terms of the Creative Commons Attribution Noncommercial License: https://creativecommons.org/licenses/by-nc/3.0/us/.

Chapter 2

Black holes and branes

2.1 Introduction

Black holes are objects of central interest in modern physics for reasons numerous, varied and deep. Beyond their ubiquity in general relativity and their astrophysical importance (as galactic epicenters [321, 322], sources of detectable gravitational waves [323] and much more [324, 325]) they geometrize a fascinating array of non-gravitational phenomena via holography [327] and, further, manifest the core tension between quantum mechanics and general relativity: the information paradox [40]. This paradox has recently been the subject of great controversy as its study has lead to new fundamental questions about the of nature black holes – especially, whether their interiors even exist, or if instead spacetime ends violently in a "firewall" near the event horizon [41]. The firewall is a dramatic departure from the conventional expectations of general relativity, but its existence is supported by a wide variety of arguments [42, 328]; nonetheless a dynamical

mechanism for its formation is lacking.

Faced with such a conflict between the predictions of quantum mechanics and general relativity, it is natural to look for resolution in string theory, where the two are united. Of course, this approach is nothing new: string theorists have been studying black holes for decades (see [329] for references), prompted by the original information paradox and mysteries about the origins of black hole entropy [45]. Black holes in string theory correspond to bound states of strings and branes [44], providing (in principle) a concrete setting for the study of black hole physics in the full quantum gravity arena. Studies of particularly supersymmetric black holes have even yielded proposals [5] for the realization of the black hole microstates as individual spacetimes, dubbed "fuzzballs", whose aggregate dynamics reproduces the naive black hole geometry of general relativity [30].

Branes are also of general interest as fundamental objects in string theory: they are the extended charges coupling to the p-forms of the massless sector and the objects on which open strings can end [238]; they arise as the duals of ordinary string configurations, and manifest the web of dualities that unifies the different string theories [330, 331]. Apart from black holes they describe a surprising variety of physics: they geometrize field-theoretic phenomena such as confinement [36, 38], the moduli space of vacua in supersymmetric theories [332] and Seiberg-Witten duality [333], and can be used to construct new field theories even in the absence of a Lagrangian description [334]; they naturally give rise both to lower-dimensional subspaces and gauge groups in string compactifications, and their phenomenological applications are broad.

This section presents several publications that aim to understand the nature

and implications of brane physics in the black hole context and beyond. In the two papers in § 2.2 my collaborators and I studied the fuzzball proposal, with particular attention to the nature of the horizon. We found evidence, in a particularly tractable circumstance, that one very rarely encounters smooth spacetime at the horizon, despite arguments in the literature to the contrary. In § 2.3 we studied the fitness of a brane-inspired solvable model of black hole physics known to capture a key feature of the information paradox: apparent thermalization in the field theory limit that corresponds to the classical limit on the gravity side. We found that this simplified model does not, however, display another key feature of black hole dynamics: classical chaos [95].

In § 2.4 we sought to address the question of metastability in the context of supersymmetry breaking induced by branes. Such supersymmetry-breaking "antibranes" were crucial in the construction of vacua with a positive cosmological constant [131], but it has been argued that these vacua, unprotected by supersymmetry, decay too quickly to be phenomenologically viable (see [336] for references). The presence of divergent fluxes near the antibranes in the corresponding supergravity solutions was taken as an indicator of violent instability. We argued, to the contrary, that these divergences (and brane actions more generally) should be understood in the context of effective field theory, and that the divergences must be resolved by matching onto the corresponding finite string diagram. We estimate the diagrams that capture the antibrane coupling to the background and show that the antibranes induce little backreaction at weak coupling, even at string-scale distances from the antibrane, where the supergravity description breaks down.

Finally, in § 2.5 we studied the question of when a field theory is holographic. Any holographic dual to ordinary (super)gravity must have black holes in its dual description, and general relativity in AdS further dictates a particularly simple phase structure [335] in which a thermal gas of gravitons collapses to form a black hole as the temperature rises. On either side of the transition there is universality: in the free energy in the graviton phase, and in the black hole entropy after the transition. The same must be true of the dual field theory, leading to strong constraints, especially when combined with the modular symmetry exchanging the spatial and temporal circles of the field theory at finite temperature. Following up on work in which these constraints were explored in CFT_2 [156] my collaborators and I studied higher dimensions, where modular symmetry is not as powerful, finding stronger constraints that must be satisfied by any holographic dual to Einstein gravity.

2.2 Journey to the Center of the Fuzzball, and A Rough End for Smooth Microstate Geometries

The papers presented in this section study the geometric realization of black hole microstates according to the fuzzball program. Our work focuses on the D1-D5 system, the simplest fuzzballs, and broadly aims to address implications for the nature of the horizon. Starting in § 2.2.1[57], we re-examined arguments [9] that typical microstates of the D1-D5 black hole are smooth spacetime geometries. We found that these arguments were in the wrong duality frame: as one approaches the fuzzball a circle becomes small, triggering a duality cascade towards the fuzzball core. In the appropriate duality frame, typical states have structure in place of the horizon that is singular, rather than smooth.

Starting in § 2.2.7 [319] we studied atypical black hole microstates, which have large angular momenta on the S^3. These microstates correspond to spacetimes that end before any duality transitions can occur and so are typically smooth. However, it was recently argued by [52] that the corresponding spacetimes are unstable. This instability takes a particularly simple form: any particle in the spacetime near the fuzzball core tends to counter-rotate with the geometry, reducing the angular momentum. We showed explicitly that the backreaction of a particle at the core (in the BPS limit) produces a fuzzball with smaller (hence more typical) angular momentum and interpreted the instability in terms of typicalization of the microstate geometry, arguing that the instability is shut off in typical microstates by stringy effects.

2.2.1 Journey to the Center of the Fuzzball: Introduction

The conflict between quantum theory and general relativity exposed by the black hole information paradox has swung back and forth for nearly four decades, recently inflamed by the firewall paradox. There have been a variety of previous proposals that the black hole horizon is not as general relativity describes. In particular, the fuzzball program argues that the structure of the horizon is necessarily modified by the extended objects of string theory. Indeed, key features of the firewall argument were first put forward as evidence for fuzzballs [1].

In this paper we focus primarily on the simplest version of fuzzballs, the two-charge system of D1-D5 branes compactified on a circle. In § 2.2.2 we reexamine the argument that the naive two-charge geometry is unphysical, and that fuzzball solutions are the correct description. We begin by noting that as one approaches the singularity of the naive geometry, the first sign of a breakdown is that the radius of a circle drops below the string scale. This suggests a T-duality from the original IIB picture to IIA, and indeed this provides a description valid down to smaller radii. Eventually the coupling grows large, and an S-duality takes us to M theory. In this regime the four-torus shrinks toward zero size, and a further STS duality brings us to a new Type II description, in which the charges are carried by fundamental strings and momentum. Finally this breaks down due to the spacetime curvature becoming large, and no further stringy duality can save us. Rather, the final picture is a weakly coupled CFT.

This onion-like layered structure has already been described in detail by Martinec and Sahakian [2], building on the classic analysis of non-conformal branes in Ref. [3]. However, its significance for the fuzzball program does not seem to have

been discussed.

Fuzzball solutions approximate the naive geometry outside some crossover radius, which depends inversely on the average harmonic excited, $\overline{m}$. For different values of $\overline{m}$, the crossover radius may lie within any of the IIB/IIA/M/II′/CFT regimes, and the parametrically valid description is a fuzzball solution in the given duality frame. For *typical* states, the crossover occurs right at the transition between the final geometric picture, II′, and the free CFT. In particular, this changes the standard picture of two-charge D1-D5 fuzzballs. The smooth geometries [4] are not an accurate description for typical states. Rather, the best (though still marginal) supergravity description is one with explicit stringy sources.

Indeed, it is well-known that typical two-charge fuzzballs lie right at the breakdown of supergravity. In fact, there are three important radii that are known to coincide: the typical *fuzzball radius* $r_{\rm f}$; the *entropy radius* $r_{\rm S}$, where the area in Planck units just matches the density of states of the system; and the *breakdown radius* $r_{\rm b}$, beyond which supergravity cannot be continued. Historically the D1-D5 fuzzballs were derived by a duality chain from F1-p solutions. These are the same as the charges of our II′ description. We trace the relation between these descriptions, and we emphasize the distinction between two free orbifold CFTs that arise in the D1-D5 system.

Much of the discussion of two-charge fuzzballs focuses on this final transition radius, and compares fuzzballs with a black hole solution including α' corrections. Our focus is rather on descriptions that are parametrically valid. In search of a more interesting situation, we consider in § 2.2.3 states with large angular momentum J, for which the naive geometry is a black ring. This geometry breaks

down due to large curvature as we approach the ring. We find that, as measured from the ring, the fuzzball and entropy radii again coincide, but the breakdown radius can be larger or smaller, depending on parameters. Thus we identify a regime where the fuzzball description is parametrically valid and physically correct, even though the naive geometry still has small curvature. We suggest that the breakdown of the naive geometry is instead signaled by the entropy radius, beyond which the naive geometry would describe more states than holography allows. In § 2.2.4 we discuss further directions.

2.2.2 The $J = 0$ system

Naive geometry: small black hole

Consider the background

$$\begin{aligned} ds^2_{\rm IIB} &= \frac{1}{\sqrt{H_1 H_5}}(-dt^2 + R^2 dy^2) + \sqrt{H_1 H_5}\, dx_4^2 + \sqrt{\frac{H_1}{H_5}}\sqrt{V} dz_4^2\,, \\ e^{\Phi_{\rm IIB}} &= g\sqrt{\frac{H_1}{H_5}}\,, \\ C_2 &= g^{-1}\left[H_1^{-1} dt \wedge R dy + Q_5 R\cos^2\tilde\theta d\psi \wedge d\varphi\right]\,, \end{aligned} \tag{2.1}$$

where

$$\begin{aligned} H_1 &= 1 + \frac{g N_1}{V r^2} \equiv 1 + \frac{Q_1}{r^2}\,, \\ H_5 &= 1 + \frac{g N_5}{r^2} \equiv 1 + \frac{Q_5}{r^2}\,. \end{aligned} \tag{2.2}$$

We work in units such that $\alpha' = 1$. The four flat transverse directions x are non-compact, and can be coordinatized as $dx_4^2 = dr^2 + r^2(d\tilde\theta^2 + \sin^2\tilde\theta d\varphi^2 + \cos^2\tilde\theta d\psi^2)$,

where the tildes are included to conform with standard notation [5]. The T^4 coordinates z have period 2π. We consider the case where the T^4 is replaced by K3 in Appendix 2.2.5.

For non-compact y, the infrared geometry is $AdS_3 \times S^3 \times T^4$ in Poincaré coordinates. If we then identify $y \cong y + 2\pi$, the horizon $r = 0$ is a fixed point and becomes a cusp singularity. For the compact theory there are three moduli: the coupling g, the circle radius R, and the torus volume V. In the attractor limit where we ignore the 1's in the harmonic functions, only the modulus g remains. The torus volume flows to the attractor value $V = N_1/N_5$, while R appears only in the combinations Rr and y/R. For simplicity we fix V to its attractor value, so that $Q_1 = Q_5 \equiv Q$ and $H_1 = H_5 \equiv H$. We are most interested in the attractor region, but it is useful to keep the harmonic function H general. The background is then given by (2.2) with $e^{\Phi_{IIB}} = g$ and $H = 1 + Q/r^2$ where the 1 drops out in the attractor.

In order for this D1-D5 description to be the correct duality frame asymptotically, we need the coupling and curvature to be small, and the circle and torus to be larger than the string scale. Thus,

$$g < 1\,, \quad Q > 1\,, \quad R > 1\,, \quad N_1 > N_5\,. \tag{2.3}$$

Discussions of this system often begin with a dual F1-p description. In § 2.2.2 we will discuss connections with this frame.

Into the black onion

In the fuzzball program, it is argued that for y compact the geometry (2.2) breaks down even before the singularity, and must be replaced by fuzzball solutions. We wish to ask, is there some signal of this breakdown as we approach the singularity?

While this work was in progress, we learned that this question had already been addressed by Martinec and Sasakian [2]. Since this result does not seem to be widely known, we review their analysis.

Note that the y circle is shrinking, and at a radius $r \sim r_{\rm IIA} = Q^{1/2}/R$ it reaches the string scale. In the D1-D5 regime (2.3) this is always inside the crossover to the near-horizon region, $r \sim r_{\rm nh} = Q^{1/2}$.[1] This breakdown suggests a T-duality along the y circle to a IIA solution, and indeed this extends the range of validity to smaller r. The solution is

$$\begin{aligned} ds^2_{\rm IIA} &= -H^{-1}dt^2 + H(d\tilde{y}^2/R^2 + dx_4^2) + \sqrt{V}dz_4^2\,, \\ e^{\Phi_{\rm IIA}} &= \frac{g\sqrt{H}}{R}\,, \\ C_1 &= \frac{R}{gH}dt\,, \\ C_3 &= g^{-1}QR\cos^2\tilde{\theta}d\psi\wedge d\varphi\wedge d\tilde{y}\,. \end{aligned} \tag{2.4}$$

In the IIA frame, the charges are carried by D0- and D4-branes localized in the $\tilde{y}$ direction. We are interested in single-particle states, so the branes should

[1] We will encounter a long list of significant radii as we move along. Figure 2.1 gives an overview. Because of the scaling in the attractor region noted above, most radii are proportional to $1/R$.

be coincident in the $\tilde{y}$ direction. Unsmearing the sources gives

$$H = \frac{Q}{r^2} \to \frac{\pi Q}{R} \sum_n \frac{1}{[r^2 + (\tilde{y} - 2\pi n)^2/R^2]^{3/2}} \sim \frac{\pi Q}{R\rho^3}\,, \tag{2.5}$$

where the normalization is fixed by the large-r behavior. The crossover to the unsmeared solution is at $r \sim r_{\rm u} = 1/R$. In the last line we have given the form as we approach the $\tilde{y} = 0$ image, where $\rho^2 = r^2 + \tilde{y}^2/R^2$.

As we continue toward the singularity, the IIA coupling becomes large, suggesting a lift to M theory. If we work with the smeared metric, this occurs at $r \sim r_{\rm M} = gQ^{1/2}/R$. Thus $r_{\rm M}/r_{\rm u} = gQ^{1/2}$. In the D1-D5 regime (2.3), g is small and Q is large, but the product $gQ^{1/2}$ is not restricted. If $gQ^{1/2} > 1$, the transition to the M theory picture occurs in the smeared regime, at $r \sim r_{\rm M}$. If $gQ^{1/2} < 1$ it occurs at in the unsmeared regime, at $\rho = \rho_{\rm M} = g^{2/3}Q^{1/3}/R$.

Either way, we end up with the M theory solution

$$\begin{aligned} ds^2_{\rm M} &= e^{-2\Phi_{\rm IIA}/3} ds^2_{\rm IIA} + e^{4\Phi_{\rm IIA}/3}(dx_{11} + C_1)^2 \\ &= \left(\frac{R^2}{g^2 H}\right)^{1/3} \left[-H^{-1}dt^2 + H(d\tilde{y}^2/R^2 + dx_4^2) + \sqrt{V}dz_4^2\right] \\ &\quad + \left(\frac{g^2 H}{R^2}\right)^{2/3} \left(dx_{11} + \frac{R}{gH}dt\right)^2\,, \\ A_3 &= C_3\,, \end{aligned} \tag{2.6}$$

(here x_{11} denotes the M direction, and the units are such that the M theory Planck scale is 1) which has p_{11} and wrapped M5 charges.

As we proceed to smaller r, both the transverse S^3 and the T^4 may shrink. The S^3 metric is proportional to $r^{2/3}$ in the smeared regime but constant ρ^0 in the

unsmeared regime, the latter property following from the conformal behavior of the M5 solution. One can check that the S^3 radius never falls below the coincident M5-brane value $N_5^{1/3}$, so this never leads to a breakdown of the solution. For the T^4, the radii become Planckian when $H = R^2V^{3/2}/g^2$. In the smeared solution this is at $r_{\rm II'} = gQ^{1/2}/V^{3/4}R = r_{\rm M}/V^{3/4}$. In the unsmeared solution it is at $\rho_{\rm II'} = g^{2/3}Q^{1/3}/V^{1/2}R$. If $r_{\rm II'} > r_{\rm u}$ the M theory solution breaks down in the unsmeared regime at $r_{\rm II'}$, otherwise it breaks down at $\rho_{\rm II'}$.

In order to extend the solution further, we must first reduce to IIA along one of the T^4 directions. The other three torus radii remain small, so a T-duality along these is needed next. This leaves the IIB coupling large, so a further S duality is needed. The net result of this STS transformation is a parametrically valid type II description

$$\begin{aligned} ds^2_{\rm II'} &= V\left[{dx_{11}}^2 + \frac{2R}{gH}dtdx_{11} + \frac{R^2}{g}\left(d\tilde{y}^2/R^2 + dx_4^2\right)\right] + d\tilde{z}_3^2\,, \\ e^{\Phi_{\rm II'}} &= \frac{RV^{3/4}}{gH^{1/2}}\,, \\ B_2^{\rm II'} &= \frac{R^2V}{gH}dt\wedge dx_{11}\,. \end{aligned} \tag{2.7}$$

In this solution one of the original torus directions has become the M direction, while x_{11} has emerged as a new periodic direction. The three $\tilde{z}$-circles remaining from the original T^4 are now string-sized. We therefore label this solution simply as II′, since it is midway between the IIA and IIB descriptions. The charges are F-string winding in the 11-direction and p_{11}.

In this final form, the curvature becomes large at $r_{\rm b} = \rho_{\rm b} = g/V^{1/2}R$. This is inside the unsmearing radius $r_{\rm u}$, so it is $\rho_{\rm b}$ that matters. When curvature

becomes large, no further string duality can save us. However, we note that the II′ description and its breakdown are very similar to those of the supergravity description of the D1-brane in Ref. [3]. There, the final supergravity description is in terms of F-strings, and it is argued that dynamics at smaller r (lower energy) is given by the long string CFT identified in Refs. [6, 7, 8]. We expect the same to hold here as well, although the additional momentum charge means that we are looking at excited states in this theory.

This conjecture is in keeping with the general expectation that when the curvature becomes large while the string coupling goes to zero, as it does in the solution (2.7), one should look for a weakly coupled CFT description. The leading twist interaction in the CFT is irrelevant [8], so that the coupling continues to go to zero in this regime.

The full picture is summarized in Figure 2.1. Martinec and Sahakian do not restrict to the asymptotic D1-D5 regime (2.3) and so cover a wider range of phases (Ref. [2], Fig. 4). Note also that they use different variables for the axes. For such non-conformal branes [3], the physics at a given scale or temperature is governed by the weakly coupled description at the corresponding holographic radius. For example, at the lowest energies the weakly coupled field theory is the appropriate description, as it is for Dp-branes with $p > 3$.

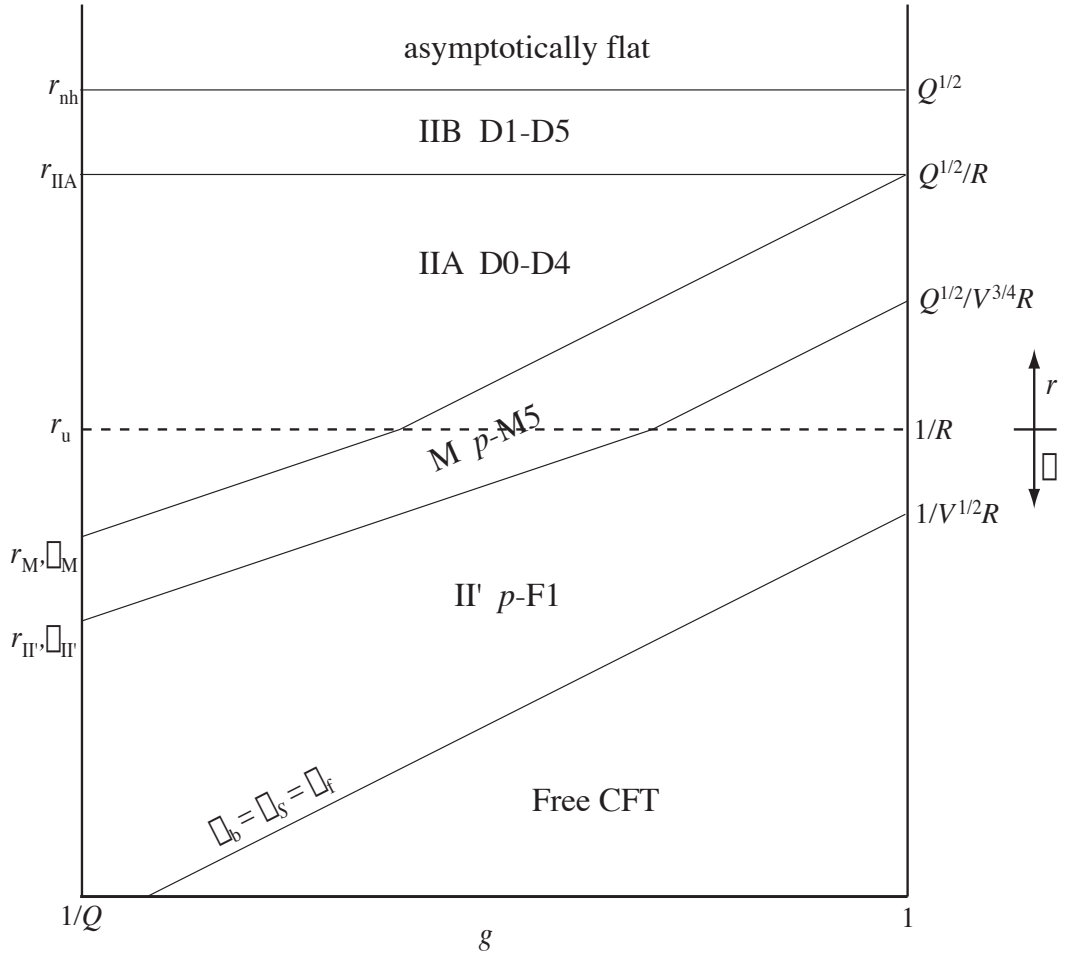


Figure 2.1: Domains of duality frames, on a log-log plot of radius and coupling. The dashed line divides smeared geometries (above) from unsmeared (below).

Fuzzball geometries

Fuzz and the onion

A more general class of two-charge geometries is characterized by a curve $\vec{F}(v)$ in the non-compact $\mathbb{R}^4$ [4]:

$$ds^2 = \frac{1}{\sqrt{H_1 H_5}}\left[-(dt+A)^2+(Rdy+B)^2\right]+\sqrt{H_1 H_5}dx_4^2+\sqrt{\frac{H_1}{H_5}}\sqrt{V}dz_4^2\,,$$

$$
\begin{aligned}
e^{\Phi} &= g\sqrt{\frac{H_1}{H_5}}\,, \\
C_2 &= g^{-1}\left[H_1^{-1}(dt+A)\wedge(Rdy+B)+\zeta\right]\,,
\end{aligned}
\tag{2.8}
$$

where the harmonic functions are

$$
\begin{aligned}
H_5 &= 1+\frac{Q_5}{L}\int_0^L \frac{dv}{|\vec{x}-\vec{F}(v)|^2}\,, \\
H_1 &= 1+\frac{Q_5}{L}\int_0^L \frac{|\dot{\vec{F}}|^2 dv}{|\vec{x}-\vec{F}(v)|^2}\,, \\
A^i &= \frac{Q_5}{L}\int_0^L \frac{\dot{F}^i dv}{|\vec{x}-\vec{F}(v)|^2}\,,
\end{aligned}
\tag{2.9}
$$

with $L=\frac{2\pi Q_5}{R}$.[2] The remaining quantities are defined via $dB=\star_4 dA$, $d\zeta=-\star_4\, dH_5$.

To be precise, this solution describes only oscillations in the transverse directions. The complete solution with oscillations in the torus directions is given in Refs. [9, 10]. It is slightly more complicated in form, but qualitatively similar, and the same estimates of radii apply.

At $r>|\vec{F}|$ these solutions go over to the naive geometry (2.2), with

$$
Q_1=\frac{Q_5}{L}\int_0^L |\dot{\vec{F}}|^2 dv\,. \tag{2.10}
$$

Expanding $\vec{F}$ in harmonics,

$$
\vec{F}=\sum_{m=1}^{\infty}\vec{F}_m e^{2\pi i m v/L}+\text{c.c.}\,, \tag{2.11}
$$

[2]The range L is a vestige of the original derivation of these solutions and does not have particular significance.

this becomes

$$2\sum_{m=1}^{\infty} m^2|\vec{F}_m|^2 = \frac{Q_1Q_5}{R^2} \tag{2.12}$$

or

$$\frac{2\ VR^2}{g^2}\sum_{m=1}^{\infty} m^2|\vec{F}_m|^2 = N_1N_5. \tag{2.13}$$

This last form is compatible with the quantization condition

$$|\vec{F}_m|^2 = \frac{g^2n_m}{2mVR^2} = \frac{r_{\rm b}^2n_m}{2m} \quad \Rightarrow \quad \sum_{m=1}^{\infty} mn_m = N_1N_5\,, \tag{2.14}$$

which can be derived either by duality from the F1-p system [4] or by quantization of the D1-D5 solution [11]. Note that the breakdown radius $r_{\rm b}$ is the same as the parameter μ in the literature, meaning that $r_{\rm b}$ maps to the string length in the F1-p frame.

For a solution with average harmonic $\overline{m}$, the sum (2.12) implies that

$$|\vec{F}| \sim \frac{\sqrt{Q_1Q_5}}{\overline{m}R} = \frac{\sqrt{N_1N_5}}{\overline{m}}r_{\rm b} \equiv r_{\overline{m}}\,. \tag{2.15}$$

As long as $r_{\overline{m}} > r_{\rm b}$ this should be a valid supergravity solution. This translates to $\overline{m} < \sqrt{N_1N_5}$. Note that $r_1 > r_{\rm IIA}$, so the largest solutions are described in the original IIB D1-D5 frame. The ratio $r_1/r_{\rm nh}$ is of order $Q^{1/2}/R$. In the asymptotic regime (2.3) this can be either large or small, but we are usually interested in scaling up the charges with other parameters held fixed. In this case the $\overline{m} \sim 1$ solutions extend into the flat Minkowski region.

As $\overline{m}$ decreases, the parametrically valid description of the state moves among the IIA, M, and II$'$ frames. Since the y and z directions remain flat in the fuzzball

solutions, it is straightforward to dualize them in the same way as for the naive solution, including the unsmearing; we do so in Appendix 2.2.6. For $\overline{m} > \sqrt{N_1 N_5}$ the states are described by the low energy CFT rather than supergravity. As $\overline{m} \to \infty$, the fuzzball solution approaches the naive solution, although the quantization condition puts the limit $m \leq N_1 N_5$ on the highest Fourier mode of $\vec{F}$.

For *typical* states, $\overline{m} \sim \sqrt{N_1 N_5}$, which defines the fuzzball radius $r_{\rm f} = r_{\sqrt{N_1 N_5}} = r_{\rm b}$. That is, these states live at the boundary of validity between the last supergravity solution and the free CFT. The fact that these fuzzballs live at the boundary of validity of supergravity is well-known in the F1-p frame [12], and remains true here. The duality cascade that we have found means that the D1-D5 geometries are never good descriptions of these typical fuzzball states. The best supergravity description would be the F1-p solutions [13, 14].

Note that for both the fuzzball and naive D1-D5 geometries, the IIB curvature is always small in terms of the tension of a probe F-string, seemingly in contradiction with what we have found. The point of the duality cascade is that there is a lighter string-like object: a probe KK monopole (charged on the y-circle, wrapped on the torus and extended in one transverse direction) which maps to a probe F-string in the II' picture. It has a tension $\tau_{\rm KK} \sim R_y^2(r)V(r)/g^2(r) = R^2 V/g^2 H(r)$, which goes to zero as it approaches the singularity and matches the IIB curvature Q^{-1} at $\rho_{\rm b}$, signaling a breakdown.

Before we go on, there is one additional radius of interest. The two-charge system has a known microscopic entropy of order

$$S \sim \sqrt{N_1 N_5}\,. \tag{2.16}$$

Let us compare this to the Bekenstein-Hawking entropy that we would ascribe to a spherical shell surrounding the singularity in the naive geometry. In the smeared regime $rR > 1$ this would be

$$\frac{\text{8d area}}{l_p^8} \sim R_y \times V_{S^3} \times V_{T^4} \times e^{-2\Phi} = \frac{r}{r_{\rm b}} \sqrt{N_1 N_5}\,. \tag{2.17}$$

The area in Planck units is the same in any duality frame; the decomposition (2.17) corresponds to the IIB picture. In the smeared regime $rR < 1$ it is

$$\frac{\text{8d area}}{l_p^8} \sim V_{S^4} \times V_{T^3} \times L_{11} \times e^{-2\Phi} = \frac{\rho}{\rho_{\rm b}} \sqrt{N_1 N_5}\,, \tag{2.18}$$

where we have used the II′ description. It is now interesting to ask, at what radius is the holographic value equal to the actual entropy? We see that this is true at $\rho = \rho_{\rm b} \equiv \rho_S$. Again this reproduces a result known from the F1-p frame [17, 12], that the horizon radius corresponding to the microscopic entropy is comparable to the breakdown radius and the typical fuzzball radius.

It is not clear then whether the fuzzball solutions are any better as a description than the naive geometry.

From F1-p to D1-D5 and back again

The D1-D5 fuzzball geometries were originally obtained [4] via U-duality from F1-p geometries describing a string with left-moving excitations:

$$\begin{pmatrix} \text{F1} \\ p \end{pmatrix} \stackrel{S}{\to} \begin{pmatrix} \text{D1} \\ p \end{pmatrix} \stackrel{T_{T^4}}{\longrightarrow} \begin{pmatrix} \text{D5} \\ p \end{pmatrix} \stackrel{S}{\to} \begin{pmatrix} \text{NS5} \\ p \end{pmatrix} \stackrel{T_y T_6}{\longrightarrow} \begin{pmatrix} \text{NS5} \\ F1 \end{pmatrix} \stackrel{S}{\to} \begin{pmatrix} \text{D5} \\ D1 \end{pmatrix} . \tag{2.19}$$

This relates the F1-p and D1-D5 moduli as

$$g_{\text{F1-}p} = \left(\frac{V^{3/4}R}{g}\right)_{\text{D1-D5}} , \quad R_{\text{F1-}p} = \left(\sqrt{V}\right)_{\text{D1-D5}} , \quad V^{1/4}_{\text{F1-}p} = \left(\frac{\sqrt{V}}{g}\right)_{\text{D1-D5}} . \tag{2.20}$$

The F1-p solutions describe the physics in a corner of the moduli space where, in terms of the asymptotic D1-D5 moduli, $V^{3/4}R/g < 1$, $V > 1$, and $\sqrt{V}/g > 1$. In this regime the D1-D5 description at infinity breaks down.

It is amusing that the descent into the fuzzball core leads us back to the F1-p duality frame in which the solutions were originally obtained, a sort of "ontogeny recapitulates phylogeny." Unlike the horizontal duality chain (2.19), the asymptotics are held fixed as we descend. The II$'$ frame in the deep IR is related to the asymptotic IIB frame by

$$\begin{pmatrix} D5 \\ D1 \end{pmatrix} \stackrel{T_y}{\to} \begin{pmatrix} D4 \\ D0 \end{pmatrix} \stackrel{S_{11}}{\longrightarrow} \begin{pmatrix} M5 \\ p \end{pmatrix} \stackrel{S_6}{\to} \begin{pmatrix} D4 \\ p \end{pmatrix} \stackrel{T_{789}}{\longrightarrow} \begin{pmatrix} D1 \\ p \end{pmatrix} \stackrel{S}{\to} \begin{pmatrix} F1 \\ p \end{pmatrix} , \tag{2.21}$$

which inverts the horizontal chain: $ST_{T^4}ST_{y6}ST_yS_{11}S_6T_{789}S = 1$.

Examining the II$'$ metric, one finds $R^{11}_{\text{II}'} = \sqrt{N_1/N_5} = R^y_{\text{F1-P}}$, while $R^y_{\text{II}'} =$

$g^{-1}\sqrt{N_1/N_5} = V_{\text{F1-P}}^{1/4}$. The long chain from F1-p to D1-D5 and back again just switches the (y, x^6) circles of the original F1-p picture with the (x_{11}, y) circles of II$'$: the emergent II$'$ description of D1-D5 at low energies matches the F1-p description obtained by moving on the asymptotic moduli space.

Orbifolds and orbifolds

The target space of the free CFT is the orbifold $(R^4 \times T^4)^{N_5}/S_{N_5}$. This should not be confused with the orbifold $(T^4)^{N_1 N_5}/S_{N_1 N_5}$ which also appears in the D1-D5 system. The latter is relevant in an entirely different duality frame where $N_5' = 1$, reached by turning on form fields on the T^4. We also note some other differences between these:

- For $(R^4 \times T^4)^{N_5}/S_{N_5}$ we are interested in states with N_1 left-moving excitations. For $(T^4)^{N_1 N_5}/S_{N_1 N_5}$ we are interested in ground states.
- For $(R^4 \times T^4)^{N_5}/S_{N_5}$ we are only interested in the sector with a single long string, because only this corresponds to a single-particle state. For $(T^4)^{N_1 N_5}/S_{N_1 N_5}$ the fractionalized strings are all bound to the D5-branes, so all winding sectors correspond to single-particle states.
- For $(R^4 \times T^4)^{N_5}/S_{N_5}$ the twist interaction is irrelevant as noted above. For $(T^4)^{N_1 N_5}/S_{N_1 N_5}$ it is marginal.

Lessons

Our conclusion is that the typical fuzzball is at the transition between two descriptions, a supergravity description with stringy sources and a weakly coupled

CFT description. There is yet a third description that has been given for this system: the black hole solution with a horizon, which exists when higher derivative terms are included [15, 16]. This is usually discussed in systems with half as much supersymmetry, where the T^4 is replaced by K3, but as shown in Appendix 2.2.5 the onion structure is the same in this case.[3] This solution allows a precise counting of supersymmetric states, but like the naive and fuzzball geometries it is on the boundary of its range of validity.

We are primarily interested in regimes where the fuzzball geometries are parametrically valid, and we will find one in § 2.2.3, but here we make a few remarks about the marginal case found above. Ref. [18] argues that two-charge systems fall into two classes, those whose description is given by smooth horizonless solutions, and those where it is a black hole from a higher derivative action. The D1-D5 system was argued to be of the first type, but the onion structure shows that, if this classification is correct, then it is of the second type.

The fuzzball description might seem to retain more information by distinguishing individual microstates, but this information may not be meaningful. As argued in [18], interactions mix the BPS states of interest into a larger space of non-BPS states, so that the resulting BPS states may bear little resemblance to their naive form. This phenomenon can be seen for example in the low-energy CFT frame. There is a twist interaction, which mixes the BPS single-long-string sector with non-BPS multi-string states (these are somewhat localized in the transverse directions and so have supersymmetry-breaking $p_\perp$).

However, there is an interesting counterargument. The one-point functions

[3]We thank Nori Iizuka for discussions of the K3 case and the relation between different pictures.

of chiral operators distinguish microstates [19, 20], and these one-point functions are not renormalized [21].[4] It is puzzling to reconcile this with the point of view above. Note that in a Haar-random state the one-point functions will be of order $e^{-S/2}$ [23]. Curiously, the same is true for Schwarzschild black holes. In thermal systems, variations of the one-point functions from their thermal averages are of order $e^{-S/2}$ [24]. However, this implies that the eigenvalues are $O(1)$, and one can find a basis in which the one-point functions are of this size in *any* thermal system.

Indeed, a similar basis has been used to argue for the genericity of firewalls, namely the basis in which the Hawking occupation numbers are diagonal [25, 26]. These would be analogous to number eigenstates for the $\vec{F}_m$. So the 'firewall' basis in these papers seems to be the Schwarzschild equivalent of the two-charge fuzzball states. This parallel is somewhat unexpected, since extremal and non-extremal horizons are in many respects quite different. Clearly it is interesting to contemplate this further.

2.2.3 The $J > 0$ system

Naive geometry: small black ring

We now focus on fuzzball states having angular momentum J in the 1-2 plane of the transverse space. The maximum value $J_{\max} = N_1 N_5$ corresponds to the

[4] In Ref. [22], it has been shown that these same one-point functions imply that the entanglement entropy distinguishes microstates.

classical solution [27, 28]

$$\vec{F}_{\max} = (a\cos\omega v, a\sin\omega v)\,, \tag{2.22}$$

where only the $m = 1$ harmonic is excited. Here

$$a = r_1 = \sqrt{Q_1 Q_5}/R\,, \quad \omega = 2\pi/L = R/Q_5\,. \tag{2.23}$$

For near maximal J, i.e.

$$\epsilon \equiv \frac{J_{\max} - J}{J_{\max}} \ll 1\,, \tag{2.24}$$

most of the excitation goes into the first harmonic. Such a solution can be described by the profile

$$\begin{aligned} \vec{F} &= \vec{F}^{(0)} + \delta\vec{F}\,, \\ \vec{F}^{(0)} &= (a_0\cos\omega v, a_0\sin\omega v)\,, \end{aligned} \tag{2.25}$$

with $a_0 = a\sqrt{J/J_{\max}}$. The sum rule (2.13) gives

$$\frac{2\ VR^2}{g^2}\sum_{m=1}^{\infty} m^2 |\delta\vec{F}_m|^2 = \epsilon N_1 N_5\,. \tag{2.26}$$

For typical states, the dominant harmonic is then $\overline{m} \sim \sqrt{\epsilon N_1 N_5}$. We have $|\delta\vec{F}|/|\vec{F}^{(0)}| \sim \sqrt{\epsilon}/\overline{m} \sim 1/\sqrt{N_1 N_5}$, so the geometry is a fuzzy ring, with thickness much less than its radius.

As in the $J = 0$ case, we can think of the naive geometry as obtained by taking

the $\overline{m} \to \infty$ limit, or equivalently by interpolating the geometry outside the fuzz down to the core of the ring. This gives [12, 30]

$$\begin{aligned} H_5 &\approx 1 + \frac{Q_5}{L}\int_0^L \frac{dv}{|\vec{x} - \vec{F}^{(0)}(v)|^2}\,, \\ H_1 &\approx 1 + \frac{J_{\max}}{J}\frac{Q_5}{L}\int_0^L \frac{|\dot{\vec{F}}^{(0)}|^2 dv}{|\vec{x} - \vec{F}^{(0)}(v)|^2}\,, \\ A^i &\approx \frac{Q_5}{L}\int_0^L \frac{\dot{F}^{(0)i} dv}{|\vec{x} - \vec{F}^{(0)}(v)|^2}\,, \end{aligned} \tag{2.27}$$

which is shown in [29, 30] to be a special case of the black ring [31, 32, 33].

Because of the factor of $J_{\max}/J$, the cancellation of singular behaviors that gives rise to a smooth geometry [27, 9] no longer occurs, and there is a singularity in the core of the ring. Using "ring coordinates" as in [31, 32] the flat metric dx_4^2 on $\mathbb{R}^4$ is

$$dx_4^2 = \frac{a_0^2}{(X-Y)^2}\left[\frac{dY^2}{Y^2-1} + (Y^2-1)d\psi^2 + \frac{dX^2}{1-X^2} + (1-X^2)d\varphi^2\right], \tag{2.28}$$

and $\mathbb{R}^4$ is foliated by surfaces of constant Y with topology $S^1 \times S^2$. The coordinates X, Y take values in the range $-1 \le X \le 1$ and $-\infty < Y \le -1$ and ψ, φ are polar angles in two orthogonal planes in $\mathbb{R}^4$ with period 2π. The angle ψ is along the ring and the ring singularity is located at $Y = -\infty$. In terms of the ring coordinates we have

$$H_1 = 1 + \frac{Q_1}{\Sigma}\,, \quad H_5 = 1 + \frac{Q_5}{\Sigma}\,, \quad \text{where} \quad \Sigma = \frac{2a_0^2}{X-Y}\,, \tag{2.29}$$

and

$$A_\psi = \frac{R}{2}(1+Y)\,, \quad B_\varphi = \frac{R}{2}(1+X)\,, \quad \zeta_{\psi\varphi} = \frac{Q_5}{2}\left[Y - \frac{1-Y^2}{X-Y}\right]\,. \tag{2.30}$$

In the near-ring limit it is useful to switch from the ring coordinates X, Y to $\theta, x_\perp$:

$$X \approx -\cos\theta\,, \qquad 1+Y \approx -\frac{a_0}{x_\perp}\,, \tag{2.31}$$

where the angle coordinate θ combines with φ to form an S^2 and $x_\perp$ is the radial coordinate transverse to the ring. The ring singularity is now located at $x_\perp = 0$. The leading behaviors (simplified again to $Q_1 = Q_5 = Q$) are

$$H_5 = H_1 \approx \frac{R}{2cx_\perp}\,, \quad A_\psi \approx -\frac{Qc}{2x_\perp}\,, \quad B_\varphi \approx \frac{R}{2}(1-\cos\theta)\,, \quad \zeta_{\psi\varphi} \approx -\frac{Q}{2}(1-\cos\theta)\,, \tag{2.32}$$

where we have introduced $c = \sqrt{J/J_{\max}} = \sqrt{1-\epsilon}$. The naive near-ring metric becomes

$$\begin{aligned} ds^2_{\rm near} \;\approx\; & \frac{2cx_\perp}{R}\Big[-\left(dt - \frac{Qc}{2x_\perp}d\psi\right)^2 + R^2\left(dy + \frac{1-\cos\theta}{2}d\varphi\right)^2\Big] \\ & + \frac{Rc}{2x_\perp}\Big[dx_\perp^2 + x_\perp^2(d\theta^2 + \sin^2\theta d\varphi^2)\Big] + \frac{cQ^2}{2Rx_\perp}d\psi^2 + \sqrt{V}dz_4^2\,. \end{aligned} \tag{2.33}$$

For $c = 1$ this is smooth at $x_\perp = 0$, but for $c < 1$ it becomes singular there. The near-ring dilaton is simply $e^\Phi = g$ and the RR potential is given by

$$C_2 \approx 2cx_\perp dt \wedge \left[dy + \frac{1-\cos\theta}{2}d\varphi\right] + Qc^2\left[dy + \left(1+\frac{1}{c^2}\right)\frac{1-\cos\theta}{2}d\varphi\right] \wedge d\psi\,. \tag{2.34}$$

In the near-ring limit there are four local charges corresponding to D1 and D5 branes wrapped on the y circle and the torus, KK monopoles wrapping the $y\psi$ directions and the torus and momentum charge along the ψ direction.[5]

No black onion rings

As we proceed toward smaller $x_\perp$, the y-circle again shrinks. However, this is merely a coordinate effect: the metric in the $x_\perp$-y plane is just $\mathbb{R}^2$, with y an angular coordinate. A T-duality provides a useful description only if the shrinking circle does not cap off smoothly, as in the $J = 0$ metric (2.2). Hence there is no repetition of the layered structure found before: there is no black onion ring.

The first breakdown of the naive geometry (2.33) is due to the divergence of the curvature, because of the uncanceled $1/x_\perp$ in $g_{\psi\psi}$ and the squashing of the Hopf fibration. The curvature invariant is calculated to be

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = \frac{22}{R^2 x_\perp^2}\epsilon^2 \,. \tag{2.35}$$

This defines the breakdown radius $x_{\perp \mathrm{b}} = \epsilon/R$.

As for the $J = 0$ case there are two other radii to compare. From the discussion below Eq. (2.26) it follows that the fuzzball radius is

$$x_{\perp \mathrm{f}} \sim r_1/\sqrt{N_1 N_5} = g/R\sqrt{V} \,. \tag{2.36}$$

To obtain the entropy radius, the area in Planck units of a torus surrounding the

[5]Note that in the near-ring geometry (2.33) the circumference of the ψ-circle seems to go to zero at large $x_\perp$. However, this occurs outside of the range of validity of (2.33). In the full solution (2.28) the 1's in the harmonic functions prevent the ψ-circle from shrinking.

ring is

$$\frac{\text{8d area}}{l_p^8} \sim L_\psi \times L_y \times L_{S^2} \times L_{T^4} \times e^{-2\Phi} \sim Q\sqrt{V} x_\perp R\sqrt{\epsilon}/g^2 \,. \tag{2.37}$$

Equating this to the entropy $\sqrt{\epsilon N_1 N_5}$, we obtain $x_{\perp S} = x_{\perp \rm f} = g/R\sqrt{V}$.

The matching of the fuzzball and entropy radii for the ring has been noted previously [12]. But unlike the $J = 0$ case considered above, the breakdown radius differs from these:

$$\frac{x_{\perp \rm b}}{x_{\perp \rm f}} = \frac{x_{\perp \rm b}}{x_{\perp S}} = \frac{\epsilon\sqrt{V}}{g} \,. \tag{2.38}$$

This ratio can be either large or small.

The interesting case is when $x_{\perp \rm f,S} \gg x_{\perp \rm b}$: the fuzzballs appear at a radius where the curvature is still small.[6] Thus they are good supergravity solutions, and give a parametrically valid description of the states in this regime. It is interesting to ask whether the naive geometry shows any signs of this premature breakdown.

For comparison, in the enhançon [35] and the $\mathcal{N} = 1^*$ geometries [36], singularities are resolved by branes expanding out to radii where the naive curvature is small. In these cases, brane probes give an indication of this: if one tries to add branes to the singularity, they feel a repulsive potential at radii where the curvature is still small. This does not seem to be the case for the black ring: one can consider atypical solutions with larger harmonics, and these can approach the

[6] The curvature is smaller than the $1/\mu^2$ that might have been expected from the curvature in the original F1-p frame (μ is defined below Eq. (2.14)). This happens because terms arising originally from $B_{\mu\nu}$ combine with the metric to produce a smoother Hopf-fibered metric. In the parameter regime where the F1-p duality frame applies, the curvature becomes stringy and there is a higher-derivative black hole solution [34].

ring much more closely. In the Klebanov-Tseytlin geometry [37], resolved in supergravity [38], a flux takes an unphysical negative value at finite radius; nothing analogous happens here.

The signal of the breakdown of the naive geometry for the black ring seems to be the entropy radius. If the naive geometry were valid, we could consider a torus thinner than $x_{\perp S}$, and the number of quantum states contained within would be larger than the exponential of the Bekenstein-Hawking entropy for the torus. It is natural to conjecture that this cannot happen: that if a system has a Hilbert space of dimension $\mathcal{D}$, then the states must be distinguishable at a radius where a surrounding surface has area $\log \mathcal{D}$, in Planck units.

For $x_{\perp \mathrm{f},S} \ll x_{\perp \mathrm{b}}$, we have not yet found a good description.

2.2.4 Discussion

Our study of two-charge fuzzballs has led to some surprises.

For $J = 0$, we find that the appropriate duality frame depends on the size of the fuzzball state, which is determined by the average harmonic $\overline{m}$. For typical states, the best supergravity description is not in terms of smooth D1-D5 solutions but rather has stringy sources. We emphasize the importance of three radii: the radius of the typical fuzzball, the radius where the transverse area is equal to the microscopic entropy, and the radius where the curvature approaches the string scale. For the two-charge system, these three radii agree, meaning in particular that the supergravity description is beginning to break down for typical states. This triple agreement is well-known in the original F1-p duality frame; it is therefore unsurprising to find it here since the II$'$ frame with F1-p charges is actually

the correct duality frame for the typical fuzzball.

Fuzzballs with other values of $\overline{m}$ are parametrically valid in one of the supergravity pictures, or in the free CFT. These descriptions accurately capture dynamical behavior and excited states, not just BPS properties.

For three-charge black holes the entropy is $S_{\text{3-charge}} \sim \sqrt{N_p N_1 N_5}$. When $N_p \ll N_1, N_5$, the geometry resembles the two-charge geometry at large radius. It begins to differ at the entropy radius (2.17, 2.18) that would correspond to $S_{\text{3-charge}}$. This is

$$r_{\text{3-charge}}(N_p) \sim \sqrt{N_p}\, r_{\text{b}}\,. \tag{2.39}$$

We see that the correct description of these solutions can be any of IIB, IIA, M, or II′, depending on N_p.

For $J \neq 0$, we have found a regime near J_{max} where the fuzzball solutions are of low curvature. It is interesting that the naive solution gives no direct indication of breakdown at the corresponding radius. The curvature is small, and probe branes see no breakdown. The key indicator seems to be the entropy radius: if the naive geometry were the correct description down to smaller radii, there would not be room for all the microstates. This leads us to conjecture that if some sets of microstates give rise to a common geometry, then this geometry must break down when the transverse area is of order the entropy in Planck units.

If we apply this to the Schwarzschild geometry in a naive way, the entropy radius r_{S} is the Schwarzschild radius r_{s}. If we pass through this radius into the interior where $r < r_{\text{s}}$, there are then too many microstates unless we begin to see deviations from the Schwarzschild geometry: this is the fuzzball proposal. Of course it is a speculation to extend such a principle from the two-charge geometry

to Schwarzschild, but we have noted other parallels in § 2.2.2.

Acknowledgments

We would like to thank J. Maldacena, E. Martinec, and S. Mathur for helpful discussions. We have also benefited from discussions with N. Iizuka, K. Skenderis, M. Taylor and other participants at the Aspen Center for Physics, supported in part by National Science Foundation Grant No. PHYS-1066293. F.C. is supported by National Science Foundation Grant No. PHY11-25915. B.M. is supported by the NSF Graduate Research Fellowship Grant No. DGE-1144085. J.P. is supported by National Science Foundation Grant Nos. PHY11-25915 and PHY13-16748. A.P. is supported by National Science Foundation Grant No. PHY12-05500.

2.2.5 Appendix A: Black onions on K3

Taking the D1-D5 system to live on K3 instead of a T^4, we find a heterotic theory at the core of the onion.[7] This is as expected, since the duality chain of § 2.2.2 sending type II F1-p to D1-D5 on T^4 maps heterotic F1-p to D1-D5 on K3.

Starting from the naive metric (2.2) with K3 replacing the torus, one is led along the same duality chain until the K3 becomes small in the M theory description. Past this point, string-string duality suggests that the appropriate picture is the heterotic theory on T^3. This follows from the same STS series that we used before, but now the duals go through a IIA orientifold, type I, and then heterotic $SO(32)$ [39]. The transformations on the metric, B-field, and dilaton are the same

[7] We thank Nori Iizuka for asking about this case.

as before, so we obtain

$$\begin{aligned} ds^2_{\rm het} &= V\left[{dx_{11}}^2 + \frac{2R}{gH}dtdx_{11} + \frac{R^2}{g}\left(d\tilde{y}^2/R^2 + dx_4^2\right)\right] + d\tilde{z}_3^2\,, \\ e^{\Phi_{\rm het}} &= \frac{RV^{3/4}}{gH^{1/2}}\,, \\ B_2^{\rm het} &= \frac{R^2V}{gH}dt\wedge dx_{11}\,. \end{aligned} \tag{2.40}$$

This matches the II′ solution (2.7) exactly; the only difference from the T^4 case is that we have ended up in a heterotic theory. As before, this description is parametrically valid until $\rho_{\rm b}$, where the curvature becomes large.

2.2.6 Appendix B: Fuzzy onions

We repeat the analysis of § 2.2.2 for the fuzzball geometries, obtaining descriptions valid for fuzzballs with various values of m.

Starting from the IIB frame with fuzz (2.8),

$$\begin{aligned} ds^2_{\rm IIB} &= H^{-1}\left[-(dt+A)^2 + (Rdy+B)^2\right] + Hdx_4^2 + \sqrt{V}dz_4^2\,, \\ e^{\Phi_{\rm IIB}} &= g\,, \\ C_2 &= g^{-1}\left[H^{-1}(dt+A)\wedge(Rdy+B) + \zeta\right]\,, \end{aligned} \tag{2.41}$$

the IIA fuzzball geometry is

$$ds^2_{\rm IIA} = -H^{-1}(dt+A)^2 + H\left[d\tilde{y}^2/R^2 + dx_4^2\right] + \sqrt{V}dz_4^2\,,$$

$$\begin{aligned} e^{\Phi_{\rm IIA}} &= g\sqrt{H}/R\,, \\ B_2^{\rm IIA} &= R^{-1}B\wedge d\tilde{y}\,, \\ C_1 &= \frac{R}{gH}(dt+A)\,, \\ C_3 &= g^{-1}\zeta\wedge d\tilde{y}\,. \end{aligned} \tag{2.42}$$

The B-field corresponds to NS5 dipole charge along $\vec{F}$, T-dual to the KK dipole in IIB. The branes unsmear for $r < r_u$ just as in the naive geometry.

The fuzzy IIA becomes strongly coupled beyond r_M/ρ_M, suggesting an M theory description:

$$\begin{aligned} ds^2_{\rm M} &= e^{-2\Phi_{\rm IIA}/3}ds^2_{IIA} + e^{4\Phi_{\rm IIA}/3}(dx_{11}+C_1)^2 \\ &= \left(\frac{R^2}{g^2H}\right)^{1/3}\left\{-H^{-1}(dt+A)^2 + H\left[d\tilde{y}^2/R^2 + dx_4^2\right] + \sqrt{V}dz_4^2\right\} \\ &\quad + \left(\frac{g^2H}{R^2}\right)^{2/3}\left[dx_{11}+\frac{R}{gH}(dt+A)\right]^2\,, \\ A_3 &= C_3 + B_2^{\rm IIA}\wedge dx_{11}\,, \end{aligned} \tag{2.43}$$

with the NS5 lifting to M5 dipole.

Once again the torus becomes small past $r_{II'}/\rho_{II'}$, and performing an STS transformation as for the naive geometry yields fuzzy II′:

$$\begin{aligned} ds^2_{\rm II'} &= V\left[{dx_{11}}^2 + \frac{2R}{gH}(dt+A)dx_{11} + \frac{R^2}{g}\left(d\tilde{y}^2/R^2 + dx_4^2\right)\right] + d\tilde{z}_3^2\,, \\ e^{\Phi_{\rm II'}} &= \frac{RV^{3/4}}{gH^{1/2}}\,, \end{aligned} \tag{2.44}$$

and $B_2^{\rm II'}$ whose field strength satisfies $H_3^{\rm II'} = \star d\,(A_3\wedge d\tilde{z}_3)$. The M5 dipole de-

scends to F1 dipole in the final frame, localized along $\vec{F}$.

2.2.7 A Rough End for Smooth Microstate Geometries: Introduction

The black hole information paradox [40, 1] and more recently the firewall argument [41, 42] have reignited the search for the correct microscopic description of black holes. The study of supersymmetric black holes in string theory has been a useful arena for this study, providing many insights. For example, such black holes may be described as bound states of strings and branes [43], which can then be explored using either the low-energy perturbative worldvolume gauge theory on the branes or supergravity at finite coupling [44]. One of the great triumphs of this approach is the explicit stringy counting [45] by Strominger and Vafa of the number of microstates of the D1-D5-P system, which famously agrees precisely with the Bekenstein-Hawking entropy of the naive black hole solution.

The fuzzball program [5, 46, 47, 48, 49, 50, 51] is an attempt to describe these microstates at finite coupling. It argues that the extended objects of string theory modify the structure of the black hole horizon and solves the information paradox by construction: there is no horizon, only an end to spacetime. Some of the major goals of the program are to explain the Bekenstein-Hawing entropy, construct representative microstates and, especially in light of the firewall paradox, to understand the consequences of the stringy/braney physics at the horizon.

Within this program one may distinguish 3 types of microstates [51]: (i) microstate *geometries*, smooth horizonless solutions of supergravity; (ii) microstate *solutions*, horizonless solutions of supergravity with singularities corresponding to D-brane sources or which can be dualized patch-wise into smooth geometries; and (iii) general *fuzzballs*, horizonless configurations which may be arbitrarily quan-

tum and/or strongly curved. Since the horizon is a classical notion, it may well be that this definition of general *fuzzball* includes all black hole microstates in any approach to the information problem. In any case, it remains an open question what fraction of black hole microstates fall into each category. In particular, while in several examples of supersymmetric black holes it has been argued [9, 51] that many microstates do in fact have a consistent description entirely within supergravity, it is far from clear that they are typical.

New questions about this program were recently raised by Eperon, Reall and Santos (ERS) [52]. Focusing on supersymmetric microstate *geometries*, they identified a non-linear classical instability due to the growth of excitations at an "evanescent ergosurface" [53] of infinite redshift. On such a surface, there are null geodesics with zero energy relative to infinity which are *stably trapped* in the potential well near the ergosurface. They find that perturbing the microstate by adding a massive particle or general wavepacket near the evanescent ergosurface eventually leads to large backreaction, even if the particle has negligible energy at infinity. In particular, the coupling of the particle to supergravity fields will allow it to gradually radiate energy and angular momentum and its trajectory will approach a geodesic that minimizes the energy. Since the particle is now following an almost-null trajectory, the local energy and hence backreaction will be very large. The instability is non-linear in the sense that it involves interactions between the particle and the radiation field. A corresponding effect arises in perturbative field theory due to the coupling of modes near the evanescent ergosurface (playing the role of the massive particle above) to radiative degrees of freedom at infinity. [1]

[1] It has long been known that a class of non-supersymmetric fuzzball solutions [54] exhibits a linear ergoregion instability [55]. However, such a stability analysis of supersymmetric microstate

The emission of angular momentum reduces the size of a fuzzball. However, at least in well-understood cases, typical fuzzballs have structure on microscopic scales and thus are not described by smooth solutions [57]. The ERS instability implies that smooth solutions can only describe the system for a short time when it is coupled to the environment. In a dual CFT description of the near-horizon region, the instability corresponds to motion among the ground states towards larger (and more generic) twist numbers [58, 59]. As a result, and as we emphasize below, such an instability might have been deduced on entropic grounds even before the identification of a dynamical mechanism by ERS.

The implications of the ERS instability for the fuzzball program depend on its endpoint. ERS proposed that it could lead to a collapse of the evanescent ergosurface and thus drive the initially smooth horizonless microstate geometry to an almost-supersymmetric black hole with the same brane charges as the microstate geometry but with different angular momenta. In particular, they suggested that the endstate of the instability (for supersymmetric D1-D5 microstates with additional momentum charge) might be a near-extremal black hole [60] or a black ring[31]. To support this argument one may note that as the solution shrinks it is described by the duality cascade of [57], but since the evanescent ergosurface is a consequence of supersymmetry it persists in every duality frame and so the ERS instability argument continues to apply.

However, entropic reasoning leads to the expectation that the endpoint is instead a typical microstate with angular momentum j_{typical} which maximizes the

geometries had not been performed until the recent work by ERS. Another recent study of dynamics focuses on the quantum tunneling of branes into microstate geometries [56]; the result suggests that a collapsing shell of matter might tunnel into a fuzzball configuration before a horizon can form.

microstate density of states $S(j)$. In particular, we suggest that the string-scale structure of a typical microstate leads to corrections that remove the instability for $j \sim j_{\text{typical}}$ and prevents the collapse to a black hole. Within the supergravity approximation the stabilized geometry is indistinguishable from a supergravity black hole but has structure at the horizon that differentiates the two in the full string theory. This structure is located at the bottom of the duality cascade described in [57], and supergravity will not capture the full physics at the fuzzball core.

To obtain a measure of analytic control over the ERS instability, we take an adiabatic limit in which the particle is well-described by an Aichelburg-Sexl-like shockwave on the evanescent ergosurface. We focus on 2-charge microstates, for which the general microstate geometries are known. Solutions with such shockwaves preserve the same supersymmetries as the microstate geometries and are thus independent of time, but a small departure from this limit will lead to slow evolution. In particular, growth of the instability leads to growth of the shockwave and thus to motion along this family of solutions. The geometries accounting for the backreaction of the shock are known explicitly [61] and in fact correspond to special cases of the more general family of microstate geometries. The CFT states dual to their near-horizon limits were described in [62]. These facts can be used to justify the entropic reasoning used above.

Analysis of any potential instability in typical microstates would require a better understanding of black hole microstates beyond supergravity. In the absence of such knowlege, we describe a simple toy model displaying what we believe to be key features of their stringy physics. In particular, the model includes both

a low-energy region near the evanescent ergosurface, a parameter that we also call j controlling the microstate size, and an analog of the internal structure that would be associated with stringy excitations used to perturb the microstates. We then study the model as one decreases j in analogy with the adiabatic evolution described above. At small enough j the low-energy region displays features on scales smaller than those set by the internal structure of the probe. The probe can then no longer take full advantage of the low-energy region, raising the ground state energy and shutting off the instability. Thus we argue that the net effect of the ERS instability is to drive smooth solutions through the duality cascade of [57] towards typicality, and the instability is stabilized by stringy corrections just as supergravity breaks down: a rough end for smooth microstate geometries.

The organization of this paper is as follows. In §2.2.8 we review some of the salient features of the supergravity and CFT descriptions of the 2-charge system. We then address the ERS instability in §2.2.9. After reviewing the main argument of [52], we study Aichelburg-Sexl-like shockwaves described above and discuss their identification in terms of known microstate geometries. This allows us to give a concrete description of adiabatic evolution along this family. §2.2.10 then describes and analyzes our toy model illustrating our proposed mechanism for stabilizing the system once the microstates become typical. We conclude with a discussion of our results in §2.2.11. Appendix §2.2.12 describes the analogous physics for a special class of 3-charge microstate solutions.

2.2.8 2-charge microstates

Our analysis will focus on 2-charge supersymmetric microstate geometries; discussion of the 3-charge case is relegated to appendix 2.2.12. There is now considerable evidence [4, 12, 5] supporting the identification of particular states $|\Psi\rangle$ in the D1-D5 CFT at small string coupling g_s and large brane charges Q_1, Q_5 with (the near-horizon limit of) a class of horizonless supergravity solutions characterized by a profile $\vec{F}$ in the four non-compact transverse spatial dimensions. The map between these descriptions takes the form

$$|\Psi\rangle = \prod_{k=1}^{N} (\sigma_k^{ss'})^{N_k} |0\rangle \quad \longleftrightarrow \quad \vec{F}(v) = \sum_{k=1}^{N} \vec{F}_k e^{ik\omega v} , \tag{2.45}$$

where the N_k are related to the Fourier amplitudes $\vec{F}_k$. We will discuss the details of the CFT and supergravity descriptions, and thus the two sides of (2.45), in §2.2.8 and §2.2.8.

CFT Review

Let us consider IIB string theory compactified to $M^{1,4} \times S^1 \times T^4$, with n_1 D1 branes wrapping the S^1 and n_5 D5 branes wrapping $S^1 \times T^4$. At parametrically large S^1 the low-energy dynamics of the bound state of these branes is described by a $(1+1)$ dimensional sigma model whose target space is the moduli space of n_1 instantons in the D5-brane gauge theory [58, 59], a resolution of the orbifold $(T^4)^N/S_N$ (the symmetric product of $N = n_1 n_5$ copies of T^4). The CFT has $\mathcal{N} = (4,4)$ supersymmetry and a moduli space of supersymmetric deformations. It is conjectured that this moduli space contains the "orbifold point" where the

target space is just the orbifold $(T^4)^N/S_N$. This is the symmetric product of a seed with 4 real bosons X_i (4 torus directions), 4 real left moving fermions ψ_i, 4 real right-moving fermions ψ_i' and central charge $c = 6$.

The complete theory with target space $(T^4)^N/S_N$ has N copies of the $c = 6$ CFT with states symmetrized between the N copies. Many details of this theory are given in [30], here we just review some relevant aspects. Modular invariance requires that we introduce twisted sectors, created by bosonic and fermionic twist operators permuting the N copies. These operators are labeled by conjugacy classes of cycles of S_N, which can be decomposed into irreps σ_k labeled by a single cycle of length k (the particular elements are irrelevant because of the symmetrization, which will be implicit). For simplicity, in our discussion below we place all oscillators ossciated with the T^4 in their ground state. A general such twisted sector state corresponds to

$$|\{N_k\}\rangle = \prod_{k=1}^{N} (\sigma_k^{ss'})^{N_k} |0\rangle \tag{2.46}$$

where $s, s' = \pm$ and

$$\sum_{k=1}^{N} kN_k = N\,, \tag{2.47}$$

since each copy must be involved in the permutation. We take the field theory on the D1-D5 system to be in the Ramond sector [9]. The σ_k in (2.46) have $(h, \tilde{h}) = (\frac{c}{24}, \frac{c}{24})$ and so any set of $\{N_k\}$ satisfying (2.47) is a Ramond ground state. This fact underlies the argument matching the ground state degeneracy with the black hole entropy.

In the D1-D5 CFT the R-symmetry is geometrized as the rotational symmetry

of the non-compact directions $SO(4) \approx SU(2)_L \times SU(2)_R$. Maximal R-charge hence corresponds to maximal angular momentum. The left-moving fermions ψ_i carry spin $\frac{s}{2}$ under $SU(2)_L$ while the right-moving fermions ψ'_i carry spin $\frac{s'}{2}$ under $SU(2)_R$; the R-charge of the state is given by $(j, j') = (j^3_L, j^3_R)$. The σ_k form bi-doublets of the $SU(2) \times SU(2)$ R-symmetry and in (2.46) and below we take the $(s, s') = (-, -)$ component.

We can now explain the main features of the density of states $S(j)$ as a function of angular momentum. The state $|N_1\rangle = \sigma_1^{N_1}|0\rangle$ with $N_1 = N$ and all other modes zero is the unique completely untwisted state, corresponding to the Ramond ground state with maximal R-charge, and thus to a state of maximal angular momentum $j_{\max} = n_1 n_5$. Less finely-tuned states have smaller angular momentum, so $S(j)$ is a decreasing function of j near $j_{\max}$. Indeed, for $j \gg \sqrt{n_1 n_5}$ (and once the oscillators associated with the internal T^4 are included as well) one finds [30]

$$S(j) = 2\pi\sqrt{2}\sqrt{n_1 n_5 - |j|} \tag{2.48}$$

to leading order in N. On the other hand, since the twist operators can contribute angular momentum with any sign, charge conjugation symmetry implies that the ensemble of all ground states has vanishing expectation value for the angular momentum. Fluctuations about the average imply typical states to have non-zero angular momentum of order $\sqrt{N} = \sqrt{n_1 n_5}$, so $S(j)$ is maximized in this regime and decreases when j is decreased further.

Before proceeding to discuss geometries, we remind the reader that states in the Ramond sector can be mapped to states in the Neveu-Schwarz sector via a symmetry of theories with $\mathcal{N} \geq 2$ in 2 dimensions known as spectral flow. The

dimensions h and R-charges j of operators change along the flow according to[63]:

$$h_\alpha = h - \alpha j + \alpha^2 \frac{c}{24}\,, \qquad j_\alpha = j - \alpha \frac{c}{12}\,. \tag{2.49}$$

In particular, a Ramond ground state with maximal R-charge $(h, j) = (\frac{c}{24}, \frac{c}{12})$ can be mapped via (2.49) with $\alpha = 1$ to the Neveu-Schwarz vacuum $(h, j) = (0, 0)$. Ramond ground states of non-maximal R-charge map to chiral primaries in the NS sector. As a result, the completely untwisted state $|N_1\rangle$ becomes the NS vacuum dual to global AdS. In particular, on the gravity side spectral flow of the near-horizon limit for the corresponding solution will give simply $AdS_3 \times S^3$ in global coordinates.

Geometries

The two-charge D1-D5 geometries are type IIB compactifications on $S^1 \times T^4$ (or K3) characterized by a curve $\vec{F}(v)$ in $\mathbb{R}^4 \times T^4$. Due to the fact that these solutions were originally constructed in a duality frame where the charges are P-F1, the curve $\vec{F}(v)$ is known as the string profile.

We will focus on solutions describing only oscillations in the four non-compact transverse directions x. The complete solution with oscillations in the T^4 directions z is given in Refs. [9, 10]. Since the T^4 factor plays no further role in our discussion of the ERS instability we will usually omit it henceforth. The argument $v = t - y$ of the string profile is a lightcone coordinate involving the spatial coordinate y along the S^1. The metric, dilaton and RR 2-form for such solutions

are given by [9]

$$\begin{aligned} ds^2 &= \frac{1}{\sqrt{H_1 H_5}}\left[-(dt-A)^2+(dy+B)^2\right]+\sqrt{H_1 H_5}dx_4^2+\sqrt{\frac{H_1}{H_5}}\sqrt{V}dz_4^2\,, \\ e^{\Phi} &= g\sqrt{\frac{H_1}{H_5}}\,, \\ C_2 &= g^{-1}\left[H_1^{-1}(dt-A)\wedge(dy+B)+\zeta\right]\,, \end{aligned} \tag{2.50}$$

where the harmonic functions are

$$H_5 = 1+\frac{Q_5}{L}\int_0^L \frac{dv}{|\vec{x}-\vec{F}(v)|^2}\,,\; H_1 = 1+\frac{Q_5}{L}\int_0^L \frac{|\dot{\vec{F}}|^2 dv}{|\vec{x}-\vec{F}(v)|^2}\,,\; A^i = -\frac{Q_5}{L}\int_0^L \frac{\dot{F}^i dv}{|\vec{x}-\vec{F}(v)|^2}\,. \tag{2.51}$$

The remaining quantities are defined via $dB = \star_4 dA$, $d\zeta = -\star_4 dH_5$.[2] $L = \frac{2\pi Q_5}{R}$, and its presence in (2.51) is a vestige of the original derivation of these solutions. The profile $\vec{F}$ relates the D5 charge Q_5 to the D1 charge:

$$Q_1 = \frac{Q_5}{L}\int_0^L |\dot{\vec{F}}|^2 dv\,. \tag{2.52}$$

These supergravity charges Q_1, Q_5 are related to the dimensionless quantized charges n_1, n_5 by

$$Q_1 = \frac{g\alpha'^3}{V}n_1\,, \quad Q_5 = g\alpha' n_5\,. \tag{2.53}$$

The y coordinate is identified under $y \to y+2\pi R$ and V is the asymptotic volume of the T^4 whose coordinates z have period 2π. The four flat transverese directions x are non-compact and can be coordinatized as $dx_4^2 = d\tilde{r}^2 + \tilde{r}^2(d\tilde{\theta}^2 + \sin^2\tilde{\theta} d\tilde{\varphi}^2 + \cos^2\tilde{\theta} d\tilde{\psi}^2)$. The relation between the Cartesian coordinates (x_1, x_2, x_3, x_4) and the

[2] Our functions H_5, H_1, A correspond, respectively, to $H^{-1}, K+1, A$ in *e.g.* [62].

spherical coordinates $(\tilde{r}, \tilde{\theta}, \tilde{\varphi}, \tilde{\psi})$ is given by $x_1 = \tilde{r}\sin\tilde{\theta}\cos\tilde{\varphi}$, $x_2 = \tilde{r}\sin\tilde{\theta}\sin\tilde{\varphi}$, $x_3 = \tilde{r}\cos\tilde{\theta}\cos\tilde{\psi}$, $x_4 = \tilde{r}\cos\tilde{\theta}\sin\tilde{\psi}$.

Supersymmetry fixes the energy to be

$$\mathcal{E} = Q_1 + Q_5 \tag{2.54}$$

while the angular momentum depends on $\vec{F}$ through [9]

$$J_{ij} = \frac{Q_5 R}{L}\int_0^L (F_i\dot{F}_j - F_j\dot{F}_i)dv\,. \tag{2.55}$$

This quantity has dimensions [length]^4 and is related to the quantized angular momentum j by

$$J_{12} = \frac{g^2}{V}j, \qquad J_{34} = \frac{g^2}{V}j' \tag{2.56}$$

in units where $\alpha' = 1$. For details relevant to computing energy and angular momentum in the above 6d geometries, see [64, 65].

It will be useful to estimate the size of a given curve $\vec{F}$ as this determines the validity of the supergravity description at the string profile [57]. As argued in [30], the size of the curve is roughly proportional to its angular momentum $J = \sqrt{J_{ij}J^{ij}}$. The $\vec{F}$ that carries maximal angular momentum $J_{\max} = Q_1Q_5$ extends to a distance $\sqrt{Q_1Q_5}/R$ from the center while strings carrying a fraction $J_{\max}/m$ of the maximum angular momentum are smaller by a factor $1/m$. As noted in § 2.2.8, most CFT states have $j_{\text{typical}}/j_{\max} \approx 1/\sqrt{n_1n_5}$ and so have size of order 1 in string units. The supergravity description is valid (i.e. weakly curved) in the large $N = n_1n_5$ limit, so from this perspective both J_{typical} and the typical

size are indistinguishable from zero. Indeed, in the strict supergravity limit one may compute the density of states in direct analogy with [66, 67, 68, 11] to obtain (2.48) (which is maximized only at $j = 0$).

In a different duality frame the solution (2.50) describes a singular string source along $\vec{F}(v)$ carrying momentum, but in the corner of moduli space where the asymptotic charges are D1-D5 it has long been argued [9] that the geometry is completely smooth. This feature is particularly intriguing, as the ensemble of 2-charge solutions approximates the $M = 0$ BTZ black hole [30], so one could argue that the actual black hole microstates were horizon-free geometries that cap off smoothly at the string profile. However, for typical states it turns out [57] that maintaining the validity of the supergravity description while descending toward the fuzzball requires a duality cascade. Furthermore, the cascade terminates in a frame where the D1-D5 charges have become P-F1 and curvature of the S^3 becomes string-scale, so that even this final supergravity description breaks down near the location of the typical string profile. Typical 2-charge states are thus not well-described by smooth geometries. However, states with atypically large angular momenta have string profiles that vary slowly enough for supergravity to remain valid even at the locus defined by $\vec{F}(v)$, in some cases using only a single duality frame. Such states are indeed described by smooth geometries.

It is therefore of particular interest that ERS [52] found an instability for the geometry with maximal angular momentum which is the prime example of such a solution. Since we will also begin our discussion of shockwaves in §2.2.9 with this special case, we now pause to describe it in some detail.

The maximally-rotating microstate

The angular momentum (2.55) obtains its maximum value for the profile function

$$\vec{F}(v) = (a\cos(\omega v)\,, a\sin(\omega v), 0, 0)\,, \quad 0 \le v \le L\,, \tag{2.57}$$

where

$$a = \frac{\sqrt{Q_1 Q_5}}{R}\,, \quad \omega = \frac{2\pi}{L}\,. \tag{2.58}$$

The D1 charge (2.52) for this profile is

$$Q_1 = Q_5 a^2 \omega^2\,, \tag{2.59}$$

and the angular momentum (2.55) in the $x_1 - x_2$ plane, or equivalently, along the $\tilde{\varphi}$ direction, takes the value

$$J_{\tilde{\varphi}} = J_{12} = Q_1 Q_5 = J_{\rm max}\,. \tag{2.60}$$

With the profile (2.57) the harmonic functions become (in the notation of [5])[3]

$$H_5 = 1 + \frac{Q_5}{r^2 + a^2\cos^2\theta}\,, \quad H_1 = 1 + \frac{Q_1}{r^2 + a^2\cos^2\theta}\,, \tag{2.62}$$

[3]We use coordinates $(r, \theta, \tilde{\varphi}, \tilde{\psi})$ in which the flat metric takes the form

$$dx_4^2 = (r^2 + a^2\cos^2\theta)\left(\frac{dr^2}{r^2 + a^2} + d\theta^2\right) + (r^2 + a^2)\sin^2\theta d\tilde{\varphi}^2 + r^2\cos^2\theta d\tilde{\psi}^2\,, \tag{2.61}$$

and $\epsilon_{r\theta\tilde{\varphi}\tilde{\psi}} = \sqrt{g} = (r^2 + a^2\cos^2\theta) r\sin\theta\cos\theta$. These coordinates are related to $(\tilde{r}, \tilde{\theta}, \tilde{\varphi}, \tilde{\psi})$ in which the S^3 takes its standard form $d\Omega_3^2 = (Q_1 Q_5)^{1/4}(d\tilde{\theta}^2 + \sin^2\tilde{\theta} d\tilde{\varphi}^2 + \cos^2\tilde{\theta} d\tilde{\psi}^2)$ by $\tilde{r} = \sqrt{r^2 + a^2\sin^2\theta}$ and $\cos\tilde{\theta} = \frac{r\cos\theta}{\sqrt{r^2 + a^2\sin^2\theta}}$.

$$A_{\tilde{\varphi}} = -Q_5 a^2 \omega \frac{\sin^2\theta}{(r^2 + a^2\cos^2\theta)}\,, \quad B_{\tilde{\psi}} = -Q_5 a^2 \omega \frac{\cos^2\theta}{(r^2 + a^2\cos^2\theta)}\,. \tag{2.63}$$

The full solution is given by

$$\begin{aligned} ds_R^2 &= -\frac{1}{h}(dt^2 - dy^2) + hf\left(d\theta^2 + \frac{dr^2}{r^2+a^2}\right) - \frac{2a\sqrt{Q_1 Q_{\tilde{5}}}}{hf}\left(\cos^2\theta dy d\tilde{\psi} + \sin^2\theta dt d\tilde{\varphi}\right) \\ &\quad + h\left[\left(r^2 + \frac{a^2 Q_1 Q_5 \cos^2\theta}{h^2 f^2}\right)\cos^2\theta d\tilde{\psi}^2 + \left(r^2 + a^2 - \frac{a^2 Q_1 Q_5 \sin^2\theta}{h^2 f^2}\right)\sin^2\theta d\tilde{\varphi}^2\right], \end{aligned} \tag{2.64}$$

with

$$f = r^2 + a^2\cos^2\theta\,, \quad h = \sqrt{H_1 H_5} = \left[\left(1 + \frac{Q_1}{f}\right)\left(1 + \frac{Q_5}{f}\right)\right]^{1/2}\,. \tag{2.65}$$

In the near-horizon limit, $r \ll (Q_1 Q_5)^{1/4}$, $a \ll (Q_1 Q_5)^{1/4} \ll R$, this solution is dual to a Ramond ground state with maximal R charge. To see this, we remind the reader that spectral flow maps the Ramond ground state to the Neveu-Schwarz ground state and that this flow is implemented by the large coordinate transformation

$$\psi = \tilde{\psi} - \frac{y}{R}\,, \quad \varphi = \tilde{\varphi} - \frac{t}{R}\,. \tag{2.66}$$

Applying (2.66) to the above metric yields

$$ds_{NS}^2 = \sqrt{Q_1 Q_5}\left[-(r'^2+1)\frac{dt^2}{R^2} + r'^2\frac{dy^2}{R^2} + \frac{dr'^2}{r'^2+1} + d\theta^2 + \cos^2\theta d\psi^2 + \sin^2\theta d\varphi^2\right]\,. \tag{2.67}$$

This is just global $AdS_3 \times S^3$ and is indeed dual to the NS vacuum state as desired.

Evanescent ergosurface

The ERS instability relies on a key feature of supersymmetric microstate geometries dubbed the evanescent ergorsurface in [53]. To describe this surface, recall [69] that supersymmetry implies the existence of a globally null Killing vector field which, when there exists a Kaluza-Klein Killing field ∂_y, may be written

$$V = \partial_t + \partial_y \,. \tag{2.68}$$

Here ∂_t and ∂_y are commuting Killing vector fields. The Killing field ∂_y is spacelike and is associated with the Kaluza-Klein direction of the 6d geometry, while ∂_t becomes timelike and canonically normalized near infinity. As a result, V can also be related to a non-spacelike Killing vector of the 5d geometry obtained from dimensional reduction along the y circle. Since V is globally null it is everywhere tangent to affinely parametrized null geodesics. It will be convenient to refer to V as the SUSY Killing field below.

The evanescent ergosurface $\mathcal{S}$ is then defined by $V \cdot \partial_y = 0$. It is thus located at $f = 0$ in the geometry (2.64), where $r = 0$ and $\theta = \pi/2$. Hence $\mathcal{S}$ is a 2d timelike submanifold of the 6d geometry. At this locus the Kaluza-Klein circle y pinches off smoothly, as does ψ. At constant t the topology of $\mathcal{S}$ is S^1 where the coordinate around this circle is φ. The Killing vector field ∂_t is timelike everywhere except on $\mathcal{S}$ where it is null (V is null everywhere and ∂_y vanishes on $\mathcal{S}$). There are zero-energy null geodesics with tangent vector V which are stably trapped on $\mathcal{S}$ and thus stay at constant $(r, \theta) = (0, \pi/2)$; more on this in §2.2.9. This evanescent ergorsurface will be the location of our Aichelburg-Sexl pp-wave.

2.2.9 Adiabatic instability of 2-charge microstate geometries

We are now ready to add null particles moving in the φ direction of the S^3 at $\theta = \pi/2$ and at the center of AdS_3 ($r = 0$). This is the location of the evanescent ergosurface after spectral flow. Our focus will be on studying the backreaction induced by such particles.

From the CFT perspective, the addition of a particle corresponds to exciting higher harmonics N_k. Starting with the NS vacuum or, after spectral flow, the Ramond ground state with maximal R-charge, we will see in §2.2.9 that the instability found in [52] will take us towards more complex and typical states $|\{N_k\}\rangle$. Our main focus, however, is on explaining the physical implications of the instability found in [52] for the gravity solutions (2.50). We therefore begin with a brief review of this instability.

The ERS instability

The instability identified in [52] is a consequence of a property called stable trapping, which is exhibited by the microstate geometries near the evanescent ergosurface $\mathcal{S}$ where the SUSY Killing field V is tangent to affinely parameterized null geodesics with zero energy. These geodesics are at rest relative to infinity, in contrast to the microstate geometries which have a non-zero angular momentum. This implies that particles following orbits of V resist the frame-dragging effect caused by the rotation of the background geometry. In this sense, the zero-energy null geodesics can be seen as possessing angular momentum opposite to that of the microstate geometry. These geodesics remain within the bounded region of

the evanescent ergosurface and are thus trapped. Because they sit at the bottom of a gravitational potential well they minimize the energy and so the trapping phenomenon is stable.

Now imagine perturbing the spacetime by adding an uncharged massive particle near to the evanescent ergosurface. If we neglect backreation, the particle moves on a geodesic. When coupled to supergravity fields it will gradually radiate energy and angular momentum and its trajectory will approach a geodesic that minimizes the energy. Hence the trajectory of the particle will approach one of the zero-energy trapped null geodesics tangent to V on the evanescent ergosurface. The particle will have very small energy as measured at infinity but, since the massive particle is now following an almost null trajectory, the energy measured by a local observer will be very large. It will thus give rise to strong backreaction. As argued in [52], this suggests an instability that triggers a large change in the spacetime geometry.

While the above reasoning used particles, one should obtain the same conclusions using a field-theoretic analysis in the WKB limit, and analogous physics follows from studying quasi-normal modes [52]. In the particle context, the fact that interactions played an important role (by allowing the massive particle to radiate) means that the instability is a non-linear effect. Note that the instability is fundamentally a consequence of the existence of stably trapped null geodesics and that an evanescent ergosurface per se is not required. In particular, one expects this instability to arise even in supersymmetric microstate geometries that do not possess a Kaluza-Klein Killing vector field and thus no concept of an evanescent ergosurface. In this sense, the ERS instability appears to be a rather robust

feature of supersymmetric microstate geometries.

What could be the endpoint of this instability? Its overall effect is to remove angular momentum from the microstate geometry via radiation. This will cause the evanescent ergosurface to shrink. It was suggested in [52] that a natural endpoint is a non-supersymmetric black hole with the same conserved charges as the microstate geometry but different angular momenta.

We will now argue for a different conclusion. To do so, we recall [52] that orbits of the SUSY Killing field V on the evanescent ergosurface are null geodesics. We then return to the above discussion of adding a particle and consider the limit where the particle becomes massless and travels precisely along such a geodesic. Such particles preserve the supersymmetry of the background geometry, so in this limit one expects there to be a stationary supergravity solution that incorporates the full backreaction from the particle even when the local energy and momentum of the null particle are large. This is not to say that the ERS instability has been completely removed, as even tiny deformations away from this limit will still trigger its effects. However, continuity implies that the ERS instability proceeds very slowly when the system is close to this SUSY null particle limit. Furthermore, we recall that the ERS instability tends only to make the particle more null and to move it even closer to the above null geodesics while increasing the locally-measured energy. As a result, close to our SUSY null limit, one may approximate the evolution induced by the ERS instability as adiabatic evolution along a one-parameter family of fully-backreacted supersymmetric supergravity solutions describing null particles on the above SUSY geodesics. The natural parameter labeling the solutions is just the locally-measured energy of the null

particle, and dynamical evolution drives this energy to slowly increase.

Our first task is thus to identify the relevant supergravity solutions. As is well known, the backreaction of a null particle in flat space is described by the Aichelburg-Sexl solution [70], which preserves the desired supersymmeries [71]. We therefore seek supersymmetric solutions of the D1-D5 system which locally take the Aichelburg-Sexl form near the null geodesic on which the particle travels. To simplify the analysis, we will in fact consider a more symmetric situation describing an ensemble of such particles that preserves both translation invariance on the internal T^4 and rotational invariance under ∂_φ: in the language commonly used to describe such solutions, we smear the particles over these directions. It will be convenient to begin with the maximally rotating microstate and in fact to start our discussion in the near-horizon limit which, under the spectral flow transformation discussed in § 2.2.8 becomes just $AdS_3 \times S^3$.

Aichelburg-Sexl solutions

We therefore consider the addition to $AdS_3 \times S^3$ of an Aichelburg-Sexl shock wave associated with a ring of particles moving at the speed of light around a circle on the S^3 at the center of AdS. As shown in [61, 62], the resulting geometry is

$$\begin{aligned} d\bar{s}^2_{NS} &= \sqrt{Q_1 Q_5}\left[-(r'^2+1)\frac{dt^2}{R^2} + r'^2\frac{dy^2}{R^2} + \frac{dr'^2}{r'^2+1} + d\theta^2 + \cos^2\theta d\psi^2 + \sin^2\theta d\varphi^2\right] \\ &\quad + \frac{q\sqrt{Q_1 Q_5}}{r'^2+\cos^2\theta}\left[\left((r'^2+1)\frac{dt}{R} + \sin^2\theta d\varphi\right)^2 - \left(r'^2\frac{dy}{R} - \cos^2\theta d\psi\right)^2\right], \end{aligned} \tag{2.69}$$

where we have corrected some typos in the expressions of [61, 62]. In (2.69), q parametrizes the locally-measured energy of the null particle; i.e., it describes the strength of the shock. For $q = 0$ (2.69) is empty $AdS_3 \times S^3$ as desired.

The geometry (2.69) has a curvature singularity at the locus of the shockwave. Near the evanescent ergosurface $(r, \theta) = (0, \pi/2)$, the leading terms in (2.69) yield

$$d\bar{s}^2_{NS} = \sqrt{Q_1 Q_5}\left[-\frac{dt^2}{R^2} + dr'^2 + d\theta^2 + d\varphi^2 + \frac{q}{f}\left(\frac{dt}{R} + d\varphi\right)^2\right], \tag{2.70}$$

which is precisely an Aichelburg-Sexl shock in otherwise-flat space propagating along $\tilde{\varphi} = \varphi + \frac{t}{R}$. Note that, as for the 2-charge geometry without the shockwave (2.64), the y and ψ circles pinch off at $f = 0$.

It is now straightforward to invert the spectral flow (2.66) and obtain the R sector solution. We further restore the asymptotically flat region by judiciously adding back the appropriate constants inside the harmonic functions. Defining the parameter $\xi = 1 - q$, this construction suggests that taking the maximally-rotating geometry (2.64), adding a ring of particles to the evanescent ergosurface and incorporating their backreaction, one obtains the geometry

$$\begin{aligned} d\bar{s}^2_R &= -\frac{1}{\bar{h}}(dt^2 - dy^2) + \bar{h}\bar{f}\left(d\theta^2 + \frac{d\bar{r}^2}{\bar{r}^2 + \bar{a}^2}\right) - \xi\frac{2a\sqrt{Q_1 Q_5}}{\bar{h}\bar{f}}\left(\cos^2\theta dy d\tilde{\psi} + \sin^2\theta dt d\tilde{\varphi}\right) \\ &\quad + \bar{h}\left[\left(\bar{r}^2 + \xi\frac{\bar{a}^2 Q_1 Q_5 \cos^2\theta}{\bar{h}^2\bar{f}^2}\right)\cos^2\theta d\tilde{\psi}^2 + \left(\bar{r}^2 + \bar{a}^2 - \xi\frac{\bar{a}^2 Q_1 Q_5 \sin^2\theta}{\bar{h}^2\bar{f}^2}\right)\sin^2\theta d\tilde{\varphi}^2\right], \end{aligned} \tag{2.71}$$

where

$$\bar{h} = \sqrt{\bar{H}_1 \bar{H}_5} = \left[\left(1 + \frac{Q_1}{\bar{f}}\right)\left(1 + \frac{Q_5}{\bar{f}}\right)\right]^{1/2}, \quad \bar{f} = \bar{r}^2 + \bar{a}^2\cos^2\theta = \xi f\,. \tag{2.72}$$

One can show that (2.71) is generated by the string profile

$$\begin{aligned}\vec{\bar{F}}(v) &= (\bar{a}\cos(\omega v/\xi+\varphi_0), \bar{a}\sin(\omega v/\xi+\varphi_0), 0, 0)\,, \quad 0\le v\le L\xi \\ \vec{\bar{F}}(v) &= (\bar{a}\cos\varphi_0, \bar{a}\sin\varphi_0, 0, 0)\,, \quad L\xi\le v<L \end{aligned} \tag{2.73}$$

after smearing over φ_0 [62]. The smearing operation should be understood as generalizing (2.50) by adding further terms to the harmonic functions sourced by a set of independent string profiles $\vec{F}_i$ with independent values of φ_0 and then taking a limit where the profiles in fact coincide and the ensemble of φ_0 values forms the uniform distribution on $[0, 2\pi]$. This construction makes it clear that the result (2.73) is indeed an appropriately supersymmetric solution, once augmented by the appropriate dilaton and form fields generated by (2.73).[4] Readers concerned about the breakdown of the supergravity description near the shock may think of (2.73) as an approximation to a smooth profile whose Fourier decomposition has no excitations higher than the $N^{\rm th}$ harmonic.

Returning to the string profile (2.73), before smearing one sees that the profile describes a string that winds once around the φ-circle on the interval $v\in[0, L\xi]$ and then remains at the same x-location for the remaining v-length $(1-\xi)L$. The last straight segment corresponds to the added particle: just a bump on a

[4]While the profile function (2.73) is very similar to the profile function that generates the solutions dual to spectral flows of the conical deficits [72, 27, 30], it has a different range of integration which destroys the Hopf structure that leads to the conical singularity. With (2.73) one finds a curvature singularity instead.

fuzzball. [5] From this profile one obtains the harmonic functions [62]

$$\bar{H}_5 = 1 + \frac{Q_5\xi}{\bar{r}^2 + \bar{a}^2\cos^2\theta} + \frac{Q_5(1-\xi)}{(x_1 - \bar{a}\cos\varphi_0)^2 + (x_2 - \bar{a}\sin\varphi_0)^2 + x_3^2 + x_4^2} \quad (2.74)$$

$$\bar{H}_1 = 1 + \frac{Q_5\bar{a}^2\omega^2/\xi}{\bar{r}^2 + \bar{a}^2\cos^2\theta}\,, \quad (2.75)$$

$$\bar{A}_{\tilde{\varphi}} = -Q_5\bar{a}^2\omega\frac{\sin^2\theta}{\bar{r}^2 + \bar{a}^2\cos^2\theta}\,, \quad (2.76)$$

where the radial coordinate at infinity $\bar{r}$ is related to r by

$$\bar{r} = \sqrt{\xi}\, r\,, \quad (2.77)$$

so that $\epsilon_{\bar{r}\theta\tilde{\varphi}\tilde{\psi}} = \sqrt{g} = (\bar{r}^2 + \bar{a}^2\cos^2)\bar{r}\sin\theta\cos\theta$ and the flat metric takes the form

$$dx_4^2 = (\bar{r}^2 + \bar{a}^2\cos^2\theta)\left(\frac{d\bar{r}^2}{\bar{r}^2 + \bar{a}^2} + d\theta^2\right) + (\bar{r}^2 + \bar{a}^2)\sin^2\theta d\tilde{\varphi}^2 + \bar{r}^2\cos^2\theta d\tilde{\psi}^2\,. \quad (2.78)$$

Averaging over φ_0 gives

$$\bar{H}_5 = 1 + \frac{Q_5}{\bar{r}^2 + \bar{a}^2\cos^2\theta}\,, \quad \bar{H}_1 = 1 + \frac{Q_1}{\bar{r}^2 + \bar{a}^2\cos^2\theta}\,, \quad \bar{A}_{\tilde{\varphi}} = -Q_5\bar{a}^2\omega\frac{\sin^2\theta}{\bar{r}^2 + \bar{a}^2\cos^2\theta}\,, \quad (2.79)$$

which leads to the geometry (2.71). Note that the relation (2.52) yields

$$\bar{a} = \sqrt{\xi}\, a\,. \quad (2.80)$$

Though it seems innocent enough, this equation is actually key to our analysis. It implies the backreacted solution to be scaled down by a factor $\sqrt{\xi}$.

[5] We are grateful to Iosif Bena for emphasizing this viewpoint.

The shrinking shockwave

We argued above that the ERS instability admits an adiabatic limit described by the family of solutions (2.71) with increasing strength q of the Aichelburg-Sexl shock, and thus with decreasing ξ. From (2.55) and the asymptotics of the metric (2.71) one finds that the angular momentum of any such solution is smaller than in the maximally rotating case by a factor $\xi = 1 - q \leq 1$ while the total energy is unchanged. We find

$$\bar{\mathcal{E}} = Q_1 + Q_5 \,, \qquad \bar{J}_{\tilde{\varphi}} = \xi Q_1 Q_5 \tag{2.81}$$

which corresponds to $j = \xi j_{\max} = \xi n_1 n_5$. Since (2.71) still possesses an evanescent ergosurface, the solution will continue to shrink and radiate angular momentum to infinity so long as the ERS analysis remains valid. Indeed, while a consistent supergravity description will require a series of duality frames as we decrease j [57], the existence of a (perhaps singular) evanescent ergosurface is guaranteed in all frames by the supersymmetry of the solution.

The solution will continue to shrink at least until we can no longer trust the ERS analysis at $\xi \sim 1/\sqrt{n_1 n_5} = 1/\sqrt{N}$. In the large N limit this corresponds to taking $\xi \to 0$, which gives

$$\bar{a} = \sqrt{\xi} a \to 0 \,, \quad \bar{f} = \bar{r}^2 + \xi a^2 \cos^2\theta \to \bar{r}^2 \,, \quad \bar{h} = \left[\left(1 + \frac{Q_1}{\bar{f}} \right) \left(1 + \frac{Q_5}{\bar{f}} \right) \right]^{1/2} \to \frac{\sqrt{Q_1 Q_5}}{\bar{r}^2} \,. \tag{2.82}$$

In this limit we recover the near-horizon metric of the $M = 0$ extremal BTZ black

hole with transverse S^3 [73, 74]:

$$d\bar{s}_R^2 = \frac{\bar{r}^2}{\sqrt{Q_1 Q_5}}(-dt^2 + dy^2) + \sqrt{Q_1 Q_5}\left(\frac{d\bar{r}^2}{\bar{r}^2} + d\theta^2 + \cos^2\theta d\tilde{\psi}^2 + \sin^2\theta d\tilde{\varphi}^2\right). \tag{2.83}$$

This is consistent with the ERS suggestion that the system evolves to become a black hole.

One effect not taken into account by ERS is the possibility that the particle seeding the instability will decay. So long as the decay products continue to be treated as classical particles, one presumes this to give rise to a set of ERS-like instabilities all acting in concert. But since we consider a limit where the instability is adiabatically slow, this system of particles will reach some sort of equilibrium at each j. Indeed, in the absence of other constraints, a coarse-grained description of this equilibrium should resemble the microcanonical ensemble of all appropriately supersymmetric states with the given value of j; after including backreaction, this is just the microcanonical ensemble of microstate geometries.

The ERS analysis thus suggests that there is a general tendency for asymptotically flat microstate geometries to evolve towards smaller j. This is no surprise for $j \sim j_{\max}$, as the microcanoncal entropy $S(j)$ decreases with increasing j in this regime according to (2.48). In fact, $S(j)$ behaves this way for all $j > j_{\text{typical}}$, and so *any* interaction should lead to this behavior when the microstate is well-described by supergravity.

On the other hand, we recall from § 2.2.8 that $S(j)$ is maximized at j_{typical} of order $\sqrt{n_1 n_5}$. As a result, so long as our microcanonical ensemble approximation

remains valid and the entropy in radiation at infinity can be neglected[6], unitarity prohibits *any* interaction from causing j to decrease below j_{typical}. This strongly suggests that – at least for generic microstates – the ERS mechanism shuts down for j near j_{typical}. The effect of the ERS instability is thus to drive smooth solutions towards stringy typicality - a rough end for these supposedly smooth spacetimes.

There is indeed ample room for corrections to the ERS analysis in this regime. As noted in § 2.2.8, microstate geometries with $j \sim j_{\text{typical}}$ have string-scale structure and could well require large corrections to the classical supergravity description used by ERS. While a full analysis is beyond the scope of this work, we describe a particular stringy effect in § 2.2.10 below that could plausibly provide such corrections and illustrate the resulting stabilization in a simple toy model.

2.2.10 A model for stabilization at typicality

The ERS analysis considered test particles and fields propagating on microstate geometries. At large j the geometries are quite smooth, so stringy corrections can be incorporated via an asymptotic expansion in α'. However, due to the presence of string-scale structure when $j \sim j_{\text{typical}}$, an accurate analysis in this regime requires any probes to be treated as quantum strings. In particular, the zero-point oscillations of probe strings mean that they will not sit sharply at the minimum of any background potential. One may thus expect this effect to raise the energy of the probe above what would be expected by naively extrapolating results from the smoother geometries at larger j. As a result, this mechanism has

[6]This is a subtle point. The entropy of radiation at infinity is divergent. We may regulate the model by placing the system in a finite-sized box. Then near j_{typical}, in the limit of large charges n_1, n_5 with fixed box size, the entropy in the radiation is negligible when compared with the microstate density of states $S(j)$.

the potential to deactivate the ERS instability at $j \sim j_{\rm typical}$. While a complete stringy analysis is beyond the scope of our work, we provide a simple toy model below exhibiting what we believe to be key features of the physics.

To set the context for our model, let us briefly return to the ERS discussion of massive particles. As discussed in § 2.4 of ERS, the energy of such particles is minimized at $j_{particle} = -\infty$ in the geometry (2.64) with $j = j_{\rm max}$. In particular, the minimum of the energy $E_{\rm min}(j_{particle})$ decreases as $j_{particle} \to -\infty$ and so the particle tends to roll down this effective-potential hill by radiating into the asymptotically flat region.

Of course, once $j_{particle}$ becomes large one must take backreaction into account. One would then like to compute the minimum energy $E_{\rm backreacted,\ min}(j)$ consistent with a given total angular momentum j (including $j_{particle}$) and the existence of the particle. Doing so will be complicated away from the adiabatic limit of § 2.2.9, but one expects the result to give an effective potential $E_{\rm backreacted,\ min}(j)$ whose qualitative features are similar to the above $E_{\rm min}(j_{particle})$, and in particular which again decreases as we make j more negative.

A toy model for such an effective potential computation is given by a family of one-dimensional models in non-relativistic quantum mechanics defined by potentials $V_j(x)$ for which we wish to compute the energy $E_{\rm model,\ min}(j)$ of the ground state. We consider the Hamiltonian

$$H = \frac{p^2}{2m} + V_j(x) \tag{2.84}$$

for each value of a parameter that we will also call j. Here there is no explicit

notion of backreaction, though it has been incorporated implicitly through our comparison of ground state energies for different values of the external parameter j.

One would like this potential to model the effective potential for timelike particles in a microstate geometry, which is minimized at the evanescent ergosurface and which becomes constant far away. It thus takes the general shape of the potential in figure 2.2. For simplicity, we model this shape by choosing

$$V_j(x) = \begin{cases} \frac{1}{2}m\omega(j)^2x^2 - V_0(j) & |x| < L \\ V_1(j) & |x| > L \end{cases} . \tag{2.85}$$

L characterizes the scale over which the potential differs from its asymptotic value, and continuity of the potential requires

$$\frac{1}{2}m\omega^2L^2 - V_0 = V_1. \tag{2.86}$$

To model the ERS instability, all the parameters should depend on j except the particle mass m. We will often leave this functional dependence implicit.

Near $x = 0$ the eigenfunctions match the harmonic oscillator, but the effects of the flat potential in the $|x| > L$ region begin to affect the n^{th} and higher states when the position fluctuations

$$\langle x^2 \rangle \approx \frac{(2n+1)}{2m\omega} \tag{2.87}$$

become $O(L^2)$. In particular, in states with $\langle x^2 \rangle \gg L^2$ the particle will not be bound to the harmonic trap. It will be useful to define the dimensionless quantity

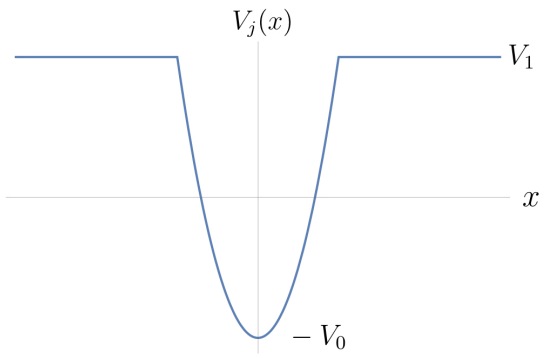


Figure 2.2: The potential in our toy model. The real shape of the effective potential for timelike particles in a microstate geometry is a smoothed version.

$$C := m\omega L^2 \tag{2.88}$$

which essentially counts the number of bound states in the potential $V_j(x)$: states with excitation number $n \ll C$ are well-approximated by harmonic oscillator eigenstates, while those with $n \sim C$ are still bound but receive significant corrections from the turning point in the potential (2.85). For $n \gg C$ the particle is effectively free.

As j decreases, the length scales of structures in our potential should decrease in analogy with the decreasing size of structures in the microstate geometries. Thus we take ω to increase, and in order to keep the number of bound states constant we hold ωL^2 fixed (this is the natural scaling in non-relativistic quantum mechanics). We take V_0 to slowly decrease with j in order to make the ground state energy $E_{\text{model, min}}(j)$ behave like $E_{\text{min}}(j_{particle})$, and in particular to slowly drive the solution towards smaller j.

So far we have merely constructed a simple quantum-mechanical toy model of the original ERS particle analysis. However, we wish to consider effects associated with the zero-point oscillations of stringy probes of the microstate geometries. In our model this can be accomodated by letting the test particle have internal structure. For the present purposes, it will be enough to regard the particle as a bound state of K partons (say, each of mass m/K) coupled by an additional internal potential $\epsilon_{\rm int}$ that depends only on the relative separations of the partons and not on j. If one likes, one may take these K particles to be connected by springs in a ring in order to give a discrete model of a quantum string.

Since the potential V_j largely models gravitational redshift effects in each microstate background, we will take each parton to experience the same potential V^{parton} whose parameters $\omega_{parton}, L_{parton}, V_0^{parton}$ we fix below in terms of the parameters of the particle model (2.85). The full Hamiltonian is

$$H = \sum_{i=1}^{K} \left(\frac{K p_i^2}{2m} + V^{parton}(x_i) \right) + \epsilon_{\rm int}. \tag{2.89}$$

We begin in the regime where the external potential ω_{parton} is small compared to all scales in the internal potential $\epsilon_{\rm int}$. This models microstates, like the maximally-rotating solution, whose structures are large compared to the string scale. In this regime the internal degrees of freedom are effectively in their ground state and we obtain a "tight binding" limit in which any differences between the x_i are small compared to any scales in the external potential. The result is an effective description of the parton composite as a single particle of mass m moving in a 1-particle potential KV^{parton} evaluated at the center of mass coordinate $\bar{x}$. The

effective physics exactly matches the single-particle model above if we identify

$$V^{parton} = \frac{V_j}{K}. \tag{2.90}$$

This implies $L_{parton} = L$, $\omega_{parton} = \omega$ and $KV_0^{parton} = V_0$.

So long as $C > 1$, there is a ground state bound to the well in which

$$\langle \bar{x}^2 \rangle \approx \frac{1}{2m\omega}. \tag{2.91}$$

The harmonic oscillator approximation to $V(\bar{x})$ implies that the ground state energy of the composite system is

$$E_0 \approx E_{\text{tight binding}} := \frac{\omega}{2} - V_0 + \epsilon_0, \tag{2.92}$$

where ϵ_0 is the ground state energy of Hamiltonian describing the intra-parton couplings.

However, the properties of the model become very different at $\omega \gg \epsilon_{\text{int}}$, i.e. as j decreases towards typicality. Any bound partons are much more strongly coupled to the external potential than to each other; if the partons remained bound, the ground state of the composite system would have each parton separately in the ground state of the potential V^{parton}. However, defining C_{parton} in analogy with (2.88) yields

$$C_{parton} := \frac{m}{K}\omega_{parton} L_{parton}^2 = \frac{C}{K}. \tag{2.93}$$

This is the quantity that counts states bound to the external potential when interactions between partons can be ignored. Taking $K \gtrsim C$ partons, the number

of such bound states will become less than one in this regime and it will be inconsistent to continue to treat all partons as bound in the external potential.

Instead, the partons pop out of the external potential well and experience only the flat potential $V_1^{parton} = V_1/K$ to good approximation when $K \gg 1$.[7] As a result, the actual ground state energy in this regime will be

$$E_{\text{model, min}} \approx V_1 + \epsilon_{\text{int}} = \frac{1}{2}m\omega^2 L^2 - V_0 + \epsilon_{\text{int}} = E_{\text{tight binding}} + (\epsilon_{\text{int}} - \epsilon_0) + \frac{\omega}{2}(C-1). \tag{2.94}$$

Taking $C > 1$ so that there is at least initially a bound state, the corrections to the tight binding energy are positive. They scale with ω at large ω and so counteract any tendency of $E_{\text{tight binding}}$ to slowly decrease due to the j-dependence of V_0. The behavior at smaller K is similar.

Note that the analogue of the ERS effective potential is $E_{\text{tight binding}}$, and that this generally differs from the actual ground state energy that would arise from putting all the particles inside the external potential well. The latter knows about the internal structure of the composite particle, while the ERS potential does not. Writing $E_{\text{model, min}}$ in terms of $E_{\text{tight binding}}$ clearly displays the extra positive term that exhibits stabilization.

To summarize, in our toy model decreasing j causes the ground state energy to decreases for a while as the instability proceeds. However, it then begins to increase again when the zero-point oscillations of the probe string no longer fit

[7] At any given time, some of the partons will in fact lie within their potential well. This effect can be estimated by studying the effective potential $K\langle V^{parton}\rangle_{\bar{x}}$, where the notation indicates the expectation value of V^{parton} for some one parton in the approximation that $\bar{x}$ is held fixed but that the system is otherwise in its ground state. One finds it to be of order $1/\sqrt{K}$, so we neglect it.

into the external potential well. Analogous behavior for the ERS phenomenon would mean that the instability stabilizes when the evanescent ergosurface develops string-scale structure, which occurs as the CFT state approaches typicality.

2.2.11 Discussion

We have argued that an adiabatic limit of the ERS instability of the 2-charge D1-D5 system is described by motion along a family of microstate geometries associated with the D1-D5 CFT. In particular, due to the emission of radiation to infinity, the angular momentum labelling the relevant microstate geometries should be thought of as a slowly-evolving function of time $j(t)$. When the instability is very weak and this evolution is especially slow, there is time for any perturbation to induce transitions between microstates and the geometry at any time t should admit an approximate description as the ensemble of all supersymmetric geometries with angular momentum $j(t)$, described in [47]. At large j the ERS instability is consistent with entropic reasoning in the CFT and indeed could have been anticipated on such grounds. From the field theory point of view, the instability simply causes evolution from states described by rare collections of twist operators to those described by more generic such collections.

On the other hand, entropic reasoning suggests that the instability terminates when j approaches $j_{\text{typical}} \sim \sqrt{n_1 n_5}$. Since this is also the regime where stringy corrections to [52] naturally become large, we suggested that the system is indeed stabilized at such j. A plausible scenario is that the zero-point oscillations of any perturbing string then prohibit it from taking full advantage of the strong redshift near the evanescent ergosurface as this surface also exhibits string-scale structure.

A full analysis is beyond our scope, but the toy model of § 2.2.10 illustrates how this effect might tame the instability.

It is important to emphasize that we have argued for stabilization only in our adiabatic limit. Since the ERS instability is non-linear, it will evolve quickly under large perturbations that take the system far away from the supersymmetric moduli space. It appears difficult to analyze this regime, and one could well imagine the endpoint in the case being either a horizon-free (but not smooth) solution with string-scale structure (a.k.a. a *rough* microstate), or a traditional black hole. As usual in this field, the question remains open for future investigation.

It would be interesting to consider a similar analysis for the 3-charge system. While in that setting it is unclear that there is any geometric analogue of typical microstates, one may in any case choose to study known classes of geometric solutions. Some initial steps involving the addition of Aichelburg-Sexl shockwaves to one such family are taken in appendix 2.2.12, but it remains to check that the conjectured fields do in fact satisfy the supergravity equations of motion, or to study more typical 3-charge microstates [75].

Even with our presumed stabilization at $j \sim j_{\text{typical}}$, the fact that it modifies the ERS instability only when the supergravity description breaks down means that much of the physical interpretation of ERS remains intact: the slightest perturbation will cause microstates with large angular momentum to collapse, with the likely endpoint being (geometrically) indistinguishable from the $M = 0$ BTZ black hole. This does not prevent one from preparing the black hole in such a microstate but, depending on parameters, it could well cause the microstate to collapse and absorb the observer into its structure before she can sail through any

smooth region where the spacetime caps off.

Acknowledgements

We would like to thank Iosif Bena, Roberto Emparan, Henry Maxfield, Harvey Reall, Jorge Santos, and David Turton for useful discussions. DM was supported in part by the U.S. National Science Foundation under grant number PHY15-04541 and by funds from the University of California. BM was supported by NSF Grant PHY13-16748. AP acknowledges support from the Black Hole Initiative (BHI) at Harvard University, which is funded by a grant from the John Templeton Foundation and support from the Simons Investigator Award 291811 and the DOE grant DE-SC0007870.

2.2.12 Appendix: Instability of 3-charge microstate geometries

We now discuss the ERS instability for the special class of 3-charge geometries constructed in [76, 77, 78, 79] and studied also by ERS [52]. Building on the 2-charge solution of §2.2.8 (but now with rotation along both angles φ and ψ of the S^3 turned on) dual to Ramond ground states, the action of spectral flow (2.49) with $\alpha \neq \pm 1$ yields excited states. In addition to D1 and D5 brane charge, these solutions have momentum excitations along the common D1-D5 direction. We review this special class of 3-charge solutions from the CFT and geometry descriptions and then briefly discuss the ERS instability along the same lines as § 2.2.9.

CFT

States in the D1-D5 CFT with momentum excitations along the common y direction correspond to excited Ramond sector states. Starting with the Neveu-Schwarz vacuum we can generate excited states in the Ramond sector through the action of spectral flow (2.49). The 3-charge states of interest are obtained by acting on the Neveu-Schwarz vacuum in the left-moving sector with

$$\alpha = 2n + 1 \quad \text{with } n \text{ integer}\,, \tag{2.95}$$

and in the right-moving sector with $\alpha = 1$ (so that the right movers are in their Ramond ground state and the CFT is supersymmetric). After spectral flow (2.49) with (2.95) the states in the symmetric product theory have dimensions $(h, \tilde{h})$ and charges (j, j'):

$$h = \tfrac{1}{4}(2n+1)^2 n_1 n_5\,, \quad \tilde{h} = \tfrac{1}{4} n_1 n_5\,, \tag{2.96}$$

$$j = -\tfrac{1}{2}(2n+1) n_1 n_5\,, \quad j' = -\tfrac{1}{2} n_1 n_5\,. \tag{2.97}$$

We get D1-D5-p states carrying momentum charge

$$n_p = h - \tilde{h} = n(n+1) n_1 n_5\,, \tag{2.98}$$

along the S^1 and angular momenta

$$j_\psi = -j' + j = -n\, n_1 n_5\,, \quad j_\varphi = -j' - j = (n+1) n_1 n_5\,, \tag{2.99}$$

on the angles of the S^3.

Geometry

The special class of 3-charge solutions obtained from the spectral flow (2.108) of the maximally rotating 2-charge solution (2.64) are given by [77, 79]

$$\begin{aligned} ds_R^2 &= -\frac{1}{h}(dt^2-dy^2)+\frac{Q_p}{hf}(dt-dy)^2+hf\left(\frac{dr^2}{r^2+(\tilde\gamma_1+\tilde\gamma_2)^2\eta}+d\theta^2\right)\\ &+h\left(r^2+\tilde\gamma_1(\tilde\gamma_1+\tilde\gamma_2)\eta-\frac{(\tilde\gamma_1^2-\tilde\gamma_2^2)\eta Q_1Q_5\cos^2\theta}{h^2f^2}\right)\cos^2\theta d\tilde\psi^2\\ &+h\left(r^2+\tilde\gamma_2(\tilde\gamma_1+\tilde\gamma_2)\eta+\frac{(\tilde\gamma_1^2-\tilde\gamma_2^2)\eta Q_1Q_5\sin^2\theta}{h^2f^2}\right)\sin^2\theta d\tilde\varphi^2\\ &+\frac{Q_p(\tilde\gamma_1+\tilde\gamma_2)^2\eta^2}{hf}(\cos^2\theta d\tilde\psi+\sin^2\theta d\tilde\varphi)^2\\ &-2\frac{\sqrt{Q_1Q_5}}{hf}\left(\tilde\gamma_1\cos^2\theta d\tilde\psi+\tilde\gamma_2\sin^2\theta d\tilde\varphi\right)(dt-dy)\\ &-2\frac{(\tilde\gamma_1+\tilde\gamma_2)\eta\sqrt{Q_1Q_5}}{hf}\left(\cos^2\theta d\tilde\psi+\sin^2\theta d\tilde\varphi\right)dy\,, \end{aligned} \tag{2.100}$$

where

$$\eta\equiv\frac{Q_1Q_5}{Q_1Q_5+Q_1Q_p+Q_5Q_p}\,, \tag{2.101}$$

$$f=r^2+(\tilde\gamma_1+\tilde\gamma_2)\eta(\tilde\gamma_1\sin^2\theta+\tilde\gamma_2\cos^2\theta)\,, \tag{2.102}$$

$$\tilde\gamma_1=a\frac{j_\psi}{n_1n_5}=-an\,,\quad \tilde\gamma_2=a\frac{j_\varphi}{n_1n_5}=a(n+1)\,, \tag{2.103}$$

while the functions h, H_1, H_5 are as in §2.2.8. The dilaton and gauge fields are

$$
\begin{aligned}
e^{\Phi} &= g\sqrt{\frac{H_1}{H_5}}, && (2.104)\\
C_2 &= -\frac{\sqrt{Q_1Q_5}\cos^2\theta}{H_1 f}\left(\tilde{\gamma}_2 dt + \tilde{\gamma}_1 dy\right)\wedge d\psi + \frac{\sqrt{Q_1Q_5}\cos^2\theta}{H_1 f}\left(\tilde{\gamma}_1 dt + \tilde{\gamma}_2 dy\right)\wedge d\varphi \\
&\quad + \frac{(\tilde{\gamma}_1+\tilde{\gamma}_2)\,\eta Q_p}{\sqrt{Q_1Q_5}H_1 f}\left(Q_1 dt + Q_5 dy\right)\wedge\left(\cos^2\theta d\psi + \sin^2\theta d\varphi\right) \\
&\quad - \frac{Q_1}{H_1 f}dt\wedge dy - \frac{Q_5\cos^2\theta}{H_1 f}\left(r^2 + \tilde{\gamma}_2(\tilde{\gamma}_1+\tilde{\gamma}_2)\eta + Q_1\right)d\psi\wedge d\varphi. && (2.105)
\end{aligned}
$$

This solution has n_1 units of D1 branes and n_5 units of D5 branes wrapping the S^1, n_p units of momentum along the S^1 and j_ψ, j_φ units of angular momenta on the S^3. The dimensionful quantities in (2.100) are related to these quantized values by (using (2.98))

$$
Q_1 = \frac{g\alpha'^3}{V}n_1\,, \quad Q_5 = g\alpha' n_5\,, \quad Q_p = \frac{g^2\alpha'^4}{VR^2}n_p = -\tilde{\gamma}_1\tilde{\gamma}_2\,. \tag{2.106}
$$

For $n = 0$, *i.e.* in the absence of momentum $Q_p = 0$, we have $\eta = 1\,, \tilde{\gamma}_1 = 0\,, \tilde{\gamma}_2 = a$ thus recovering (2.64).

The energy and angular momenta are

$$
\mathcal{E} = Q_1 + Q_5 + 2Q_p\,, \quad J_{\tilde{\psi}} = \tilde{\gamma}_1 R\sqrt{Q_1Q_5}\,, \quad J_{\tilde{\varphi}} = \tilde{\gamma}_2 R\sqrt{Q_1Q_5} \tag{2.107}
$$

and the coordinate transformation correponding to spectral flow (2.49) with (2.95)

is given by

$$\psi = \tilde{\psi} - \tilde{\alpha}\frac{a}{\sqrt{Q_1Q_5}}y + (\tilde{\alpha}-1)\frac{a}{\sqrt{Q_1Q_5}}t\,, \quad \varphi = \tilde{\varphi} - \tilde{\alpha}\frac{a}{\sqrt{Q_1Q_5}}t + (\tilde{\alpha}-1)\frac{a}{\sqrt{Q_1Q_5}}y\,. \tag{2.108}$$

For $\tilde{\alpha} = 1$ this reduces to the coordinate transformation (2.66) for which the metric in the near-horizon limit $r \ll \sqrt{Q}$ and $a \ll \sqrt{Q} \ll R$ (implying $Q_p \ll Q$ and $\eta \to 1$) becomes $AdS_3 \times S^3$ dual to the NS vacuum. The exicted Ramond states obtained from spectrally flowing the NS vaccuum with (2.95) are dual to geometries obtained from $AdS_3 \times S^3$ *via* the coordinate transformation (2.108) with $\tilde{\alpha} = n$.

Aichelburg-Sexl in excited $AdS_3 \times S^3$

The same procedure as in §2.2.9 suggests that the addition of massless particles to the class of 3-charge solutions (2.100) is described by the geometries [8]

$$\begin{aligned} ds_R^2 \;=\; & -\frac{1}{\bar{\bar{h}}}(dt^2-dy^2) + \frac{Q_p}{\bar{\bar{h}}\bar{\bar{f}}}(dt-dy)^2 + \bar{h}\bar{f}\left(\frac{d\bar{r}^2}{\bar{r}^2+(\bar{\bar{\gamma}}_1+\bar{\bar{\gamma}}_2)^2\eta} + d\theta^2\right) \\ & +\bar{h}\left(\bar{r}^2 + \bar{\bar{\gamma}}_1(\bar{\bar{\gamma}}_1+\bar{\bar{\gamma}}_2)\eta - \xi\frac{(\bar{\bar{\gamma}}_1^2-\bar{\bar{\gamma}}_2^2)\eta Q_1Q_5\cos^2\theta}{\bar{h}^2\bar{f}^2}\right)\cos^2\theta d\tilde{\psi}^2 \\ & +\bar{h}\left(\bar{r}^2 + \bar{\bar{\gamma}}_2(\bar{\bar{\gamma}}_1+\bar{\bar{\gamma}}_2)\eta + \xi\frac{(\bar{\bar{\gamma}}_1^2-\bar{\bar{\gamma}}_2^2)\eta Q_1Q_5\sin^2\theta}{\bar{h}^2\bar{f}^2}\right)\sin^2\theta d\tilde{\varphi}^2 \\ & +\frac{Q_p(\bar{\bar{\gamma}}_1+\bar{\bar{\gamma}}_2)^2\eta^2}{\bar{h}\bar{f}}(\cos^2\theta d\tilde{\psi} + \sin^2\theta d\tilde{\varphi})^2 \\ & -2\xi\frac{\sqrt{Q_1Q_5}}{\bar{h}\bar{f}}\left(\tilde{\gamma}_1\cos^2\theta d\tilde{\psi} + \tilde{\gamma}_2\sin^2\theta d\tilde{\varphi}\right)(dt-dy) \\ & -2\xi\frac{(\tilde{\gamma}_1+\tilde{\gamma}_2)\eta\sqrt{Q_1Q_5}}{\bar{h}\bar{f}}\left(\cos^2\theta d\tilde{\psi} + \sin^2\theta d\tilde{\varphi}\right)dy \end{aligned} \tag{2.109}$$

[8] The dilaton is as in (2.104), while the RR 2-form picks up an extra piece proportionl to $(1-\xi)$ relative to (2.104) as in [62].

where

$$\bar{\tilde{\gamma}}_i = \sqrt{\xi}\tilde{\gamma}_i. \tag{2.110}$$

In particular, in the near-horizon limit this yields an Aichelburg-Sexl shockwave propagating along both angles of the S^3:

$$\begin{aligned} ds^2_{NS} &= -\left(r^2+a^2\right)\frac{dt^2}{\sqrt{Q_1Q_5}} + r^2\frac{dy^2}{\sqrt{Q_1Q_5}} + \sqrt{Q_1Q_5}\frac{dr^2}{r^2+a^2} \\ &\quad + \sqrt{Q_1Q_5}\left(d\theta^2 + \cos^2\theta d\psi^2 + \sin^2\theta d\varphi^2\right) \\ &\quad + \frac{q\sqrt{Q_1Q_5}}{f}\Bigg\{\left[\left(r^2+a^2\right)\frac{dt}{\sqrt{Q_1Q_5}} + \tilde{\gamma}_1\cos^2\theta d\psi + \tilde{\gamma}_2\sin^2\theta d\varphi\right]^2 \\ &\qquad - \left[r^2\frac{dy}{Q} - \tilde{\gamma}_2\cos^2\theta d\psi - \tilde{\gamma}_1\sin^2\theta d\varphi\right]^2\Bigg\} \end{aligned} \tag{2.111}$$

We have not checked that this is a solution other than for the trivial cases $q = 0$ and $q = 1$. Assuming that it is, we may then again describe an adibatic limit of the ERS instability as the growth of q with time. Again, this causes the backreacted solution to shrink as a function of time, decreasing the angular momentum by a factor $\xi = 1 - q$ while leaving the total energy unchanged:

$$\bar{\mathcal{E}} = Q_1 + Q_5 + 2Q_p\,, \quad \bar{J}_{\tilde{\psi}} = \tilde{\gamma}_1 R\sqrt{Q_1Q_5}\xi\,, \quad \bar{J}_{\tilde{\varphi}} = \tilde{\gamma}_2 R\sqrt{Q_1Q_5}\xi\,. \tag{2.112}$$

As in the 2-charge case, the solution will continue to shrink at least until we can no longer trust the ERS analysis at $\xi \sim 1/\sqrt{n_1n_5} = 1/\sqrt{N}$. In the large N limit this corresponds to taking $\xi \to 0$. Making this replacement in (2.109) yields the near-horizon metric of extremal BTZ black hole with a transverse S^3:

$$ds^2_R = \frac{\bar{r}^2}{\sqrt{Q_1Q_5}}(-dt^2 + dy^2) + \frac{Q_p}{\sqrt{Q_1Q_5}}(dt - dy)^2$$

$$+\sqrt{Q_1 Q_5}\left(\frac{d\bar{r}^2}{\bar{r}^2}+d\theta^2+\cos^2\theta d\tilde{\psi}^2+\sin^2\theta d\tilde{\varphi}^2\right). \qquad (2.113)$$

This is the near-horizon limit of the 5d non-rotating D1-D5-p (Strominger-Vafa) black hole [80]. Hence this preliminary analysis suggests that, as in the 2-charge case, the ERS instability proceeds until the 3-charge microstate is geometrically indistinguishable from the extremal BTZ black hole outside its putative horizon.

2.3 Four-point function in the IOP matrix model

String theory captures the microscopic physics of black holes via the brane description, but the corresponding spacetime description of the microstates is not always obvious; most systems do not seem to have a straightforward geometric description of the microstates along the lines of the D1-D5 model above. However, one can still study the field theories that describe the dynamics of these brane systems at low energies and hope to learn something about the underlying black hole mechanics. Perhaps the simplest such system consists entirely of branes of one type – we can consider N D0-branes, by T-duality. While the spacetime description of the D0-brane black hole is complex, its field theory description is very simple: maximally supersymmetric matrix quantum mechanics [129, 8, 6].

Much can be learned by studying probes of the D0 black hole, and a particularly simple probe is another D0 brane. The field theory describing the interactions of the D0 black hole and probe is not solvable but slight modification simplifies the system significantly [81, 83], in fact leading to exact solubility in the limit where the number of D0 branes in the black hole becomes large. This enables one to study features of the black hole in the brane field theory, provided the simplified models accurately capture those aspects of the black hole physics.

In [81, 83] it was shown that these simplified models capture an important aspect of black hole physics: apparent thermalization at infinite N, corresponding to the pure gravity limit, with non-thermal effects becoming visible at subleading orders. However, another ubiquitous feature of black holes is chaos: the exponential dependence of gravitational potential as a function of r near the horizon leads to rapidly growing separations between nearby perturbations of the horizon [95].

Chaos can be diagnosed by exponential growth in the commutator squared,

$$\langle [\mathcal{O}(t), \mathcal{O}(0)]^2 \rangle \sim e^{2\lambda t} \tag{2.114}$$

or equivalently in the out-of-time-order four-point correlation function. We computed the out-of-time-order four-point function in these models but found that it is oscillatory rather than exponentially growing at late times. This indicates that these systems do not in fact model black hole chaos.

2.3.1 Introduction

Matrix models are useful toy models of gauge theories and holography. Strongly coupled quantum field theories are difficult to understand directly, having a prohibitively large set of Feynman diagrams that must be summed. A good model should have a sufficiently small and well-organized set of diagrams, allowing for the computation of the full planar correlation functions. The diagrammatic structure should, however, be sufficiently nontrivial so as to capture the essential features of the bulk.

The IP model [81] is a simple large-N system of a harmonic oscillator in the $U(N)$ adjoint representation plus a harmonic oscillator in the $U(N)$ fundamental representation, coupled through a trilinear interaction. It has the same graphical structure as the 't Hooft model of two-dimensional QCD [82]. The IOP model [83] is a more tractable variant of the IP model. It possesses the same degrees of freedom, but the trilinear interaction is replaced by one that is quartic in the oscillators but quadratic in the $U(N)$ charges. Building on ideas of [84], the IP

and IOP models were introduced in [81, 83] as toy models of the gauge theory dual of an AdS black hole. These models capture a key property of black holes: the long time decay of the two-point function at infinite N, but not at finite N [85].

In this paper we compute the thermal four-point function in the IOP model in the planar limit. The motivation for studying the four-point function comes from recent work in quantum chaos and holography [86, 87, 88, 89, 90, 91, 92, 154, 94, 95, 96, 97, 98, 99, 100, 101, 150]. A signature of quantum chaos in a large-N theory is the exponential growth in time of the connected out-of-time-order four-point function [103]. The growth rate is identified as a Lyapunov exponent. A black hole has a Lyapunov exponent of $2\pi T$ [87, 89], which is the maximal possible Lyapunov exponent [88]. The significance of the out-of-time-order four-point function as a diagnostic for the viability of a model of holography was recognized in [86].

In Sec. 2.3.2 we review the role of the two-point function as a diagnostic of thermalization. In Sec. 2.3.2 we review the role of the out-of-time-order four-point function as a diagnostic of chaos. In Sec. 2.3.2 we briefly mention the Sachdev-Ye-Kitaev model [104, 87], which was recently recognized to be maximally chaotic [87]. We point out that the random coupling can, to leading order in $1/N$, be replaced by a quantum variable.

In Sec. 2.3.3 we review the calculation of the planar two-point function in the IP model. In Sec. 2.3.3 we compute the planar four-point function. This involves summing ladder diagrams, which can only be done analytically in the limit of small adjoint mass, to which we restrict ourselves.

In Sec. 2.3.4 we review the planar two-point function in the IOP model. In

Sec. 2.3.4 we compute the planar four-point function. Diagrammatically, the IOP model is more involved than the IP model. However, it has the advantage of allowing analytic computations for any adjoint mass. For both the IP and IOP models, we work in the limit that the mass of the fundamental is heavy, as compared to the temperature.

In the regimes considered, we find that the IP and IOP models are not chaotic. Some speculations on why this is so, and possible modifications of the models, are mentioned in Sec. 2.3.5.

2.3.2 Thermalization, chaos, and large N

Thermalization

Holography has provided useful insights into both strongly coupled field theories, as well as their gravity duals. A well-studied property of a black hole is its approach to equilibrium after a perturbation. A two-point function computed in a black hole background exhibits late time decay of the form [105, 106],

$$\langle \varphi(t)\varphi(0)\rangle \sim e^{-ct/\beta} \ , \tag{2.115}$$

where c is an order-one constant and β is the inverse temperature. The late time decay of the two-point function has a clear interpretation in the bulk: matter falls into the black hole, but classically nothing escapes. Computing subleading corrections in G_N to (2.115) does not prevent the late time decay.

As recognized in [85], the late time decay to zero of a two-point function is inconsistent with the properties of a finite entropy quantum mechanical system.

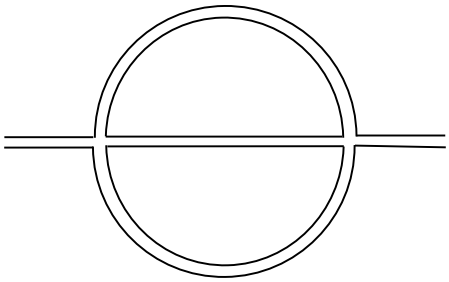


Figure 2.3: The basic graphical unit of the Hamiltonian (2.116) studied in [84].

On the field theory side, one thus has the statement that, to all orders in $1/N$, the two-point function decays to zero at late times, even though this property does not hold nonperturbatively in $1/N$. The two-point function $\langle\varphi(t)\varphi(0)\rangle_\beta$ can be regarded as the overlap between the states $\varphi(0)|\beta\rangle$ and $\varphi(t)|\beta\rangle$; its decay is a probe of thermalization. Therefore, the large N limit acts like a thermodynamic limit [84].

This late time breakdown of perturbation theory was studied in the context of matrix quantum mechanics in [84]. Reducing Yang-Mills on a sphere in terms of spherical harmonics, one obtains a Hamiltonian whose essential features can be captured by considering just two interacting matrices. For instance, [84] considers two large N matrices M_1, M_2 with a Hamiltonian,

$$H = \sum_{i=1}^{2} \frac{1}{2}\mathrm{Tr}(\dot{M_i}^2 + \omega_i^2 M_i^2) + \lambda \mathrm{Tr}(M_1 M_2 M_1 M_2) \ . \tag{2.116}$$

The relevant diagrams for the decay of the two-point function are the sunset diagrams shown in Fig. 2.3.

The model (2.116) has the drawback of still being too complicated to allow the summation of all planar Feynman diagrams. The goal of [81] was to find a

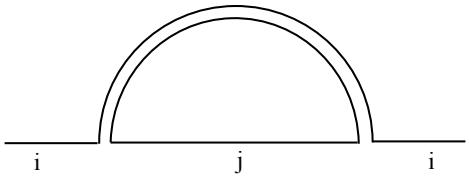


Figure 2.4: The basic graphical unit of the IP model (2.117) studied in [81]. It is like the diagram in Fig. 2.3, but cut in half. A single line is a fundamental, a double line is an adjoint.

matrix model that is more tractable, while still exhibiting the late time decay of the planar two-point function. The IP model [81] is given by the Hamiltonian,

$$H_{IP} = m\mathrm{Tr}(A^\dagger A) + Ma^\dagger a + ga^\dagger X a \ , \tag{2.117}$$

where a_i is the annihilation operator for a harmonic oscillator in the fundamental of $U(N)$, while A_{ij} is the annihilation operator for an oscillator in the adjoint, and $X_{ij} = (A_{ij} + A^\dagger_{ji})/\sqrt{2m}$. [1] As we review in Sec. 2.3.3, the planar two-point function can be found if one takes the mass of the fundamental to be large compared to the temperature, $M \gg T$. For a general mass m for the adjoint, the planar Schwinger-Dyson equation for the two-point function can be solved numerically, exhibiting the desired late time exponential decay. In the limit of small mass for the adjoint, $m \to 0$, the two-point function can be found analytically, giving late time power law decay.

A variant of the IP model, the IOP model, was introduced in [83],

$$H_{IOP} = m\mathrm{Tr}(A^\dagger A) + Ma^\dagger a + ha^\dagger_i a_l A^\dagger_{ij} A_{jl} \ . \tag{2.118}$$

[1] Since the highest term in the Hamiltonian (2.117) is cubic, there is no ground state. This is cured by adding a stabilizing term, $a^\dagger a(a^\dagger a - 1)$, which vanishes in the relevant sectors $a^\dagger a = 0, 1$ [81].

This model has the feature that analytic computations are possible for any mass m. It again exhibits power law decay of the two-point function at long times.

Chaos

Chaos in deterministic systems is understood as aperiodic long-term behavior that exhibits sensitive dependence on initial conditions. Two points in phase space, characterized by a separation $\delta x(0)$, will initially diverge at a rate,

$$\delta x(t) = \delta x(0)\, e^{\kappa t}, \tag{2.119}$$

where κ is the Lyapunov exponent.

For a number of reasons [107], there is no straightforward extension of chaos to quantum systems. In the semiclassical regime, [103] gave an intuitive definition of chaos. Replacing the variation in (2.119) by a derivative, and noting that this is given by a Poisson bracket,

$$\frac{\partial x(t)}{\partial x(0)} = \{x(t), p(0)\}\ , \tag{2.120}$$

the generalization to quantum systems consists of replacing the Poisson bracket by a commutator. The commutator is generally an operator, so seeing exponential growth requires taking an expectation value. The expectation value of the commutator in the thermal state will vanish, as a result of phase cancelations. A

simple way to cure this is to consider the square of the commutator [103], [2]

$$\langle [x(t), p(0)]^2 \rangle \sim \hbar^2 e^{2\kappa t} \ . \tag{2.121}$$

Alternatively, one can consider the thermal expectation value of the commutator times the anticommutator; this will scale as $\hbar$. Either of these consist of sums of out-of-time-order four-point functions. The important point is that a chaotic system has an out-of-time-order four-point function that exhibits exponential growth. The exponential growth persists until a time of order $-\kappa^{-1} \log \hbar$, at which point the commutator saturates at an order one value.

For a large N field theory, $1/N$ plays the role of $\hbar$, and the classical limit is the infinite N limit. For matrix models, such as the IP and IOP models, leading order in $1/N$ corresponds to keeping the planar Feynman diagrams. The criteria of chaos for evaluating the viability of a model is a powerful one, that was recognized in [86]. A good model of a strongly coupled gauge theory should having an exponentially growing out-of-time-order four-point function. Moreover, if it is to be dual to a black hole, the Lyapunov exponent must match that of a black hole [87, 89].

[2] The expectation value in (2.121), and elsewhere, is in the thermal state. The Lyapunov exponent depends on the temperature: this is the familiar statement from classical chaos that regions of phase space that do not mix have different Lyapunov exponents. If we were working in the microcanonical ensemble, then the energy would be conserved. Note also that the definition of the Lyapunov exponent that is being used is nonstandard, in that it is a local Lyapunov exponent, rather than involving a time average.

Thermalization and chaos

There is generally an intimate connection between thermalization and chaos. In the context of classical systems, there is a precise version of this statement [108], which we now review.

Letting A and B be regions of phase space, occupying phase space volumes $\mu(A)$ and $\mu(B)$, respectively, and letting φ_t denote time evolution, a dynamical system is said to be mixing if $\mu\left[\varphi_t A \cap B\right] \to \mu(A) \cdot \mu(B)$ as $t \to \infty$, for all sets A and B. In the notation of quantum mechanics, this is the statement that a system is mixing if the (connected) two-point function of any two operators decays to zero at late time. A system is defined to be ergodic if for every function f, the time mean of $f(x)$ is equal to the space mean of $f(x)$. It is shown in [108] that mixing implies ergodicity, but ergodicity does not necessarily imply mixing.

It is important to note that for a system to be mixing, the two-point function of all operators must decay. In fact, the IP and IOP models do not satisfy this criteria, as it is only the two-point functions of the fundamentals that exhibit late time decay. [3] The adjoints have a two-point function of a free harmonic oscillator; they have no self-interaction, and the interaction generated via the fundamentals is $1/N$ suppressed. Thus, exponential growth of the out-of-time-order four-point function for the fundamentals is a more refined criteria than the decay of the two-point function of the fundamentals at long times.

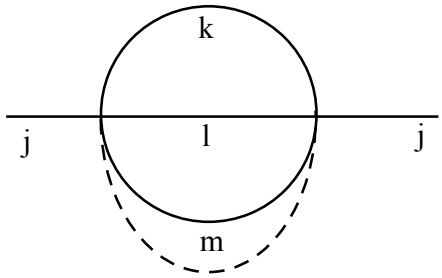


Figure 2.5: The basic graphical unit of the SYK model (2.122). The solid lines are fermions χ_i, the dotted line is the coupling J_{jklm}.

SYK model

Kitaev has proposed a variant of the Sachdev-Ye model [104] as a model of holography [87]. The SYK model consists of $N \gg 1$ Majorana fermions χ_i with a quartic interaction with random coupling J_{jklm},

$$H_{SYK} = \frac{1}{4!} \sum_{j,k,l,m=1}^{N} J_{jklm}\ \chi_j \chi_k \chi_l \chi_m \ , \tag{2.122}$$

where couplings are drawn from the distribution,

$$P(J_{jklm}) \sim \exp(-N^3 J_{jklm}^2 / 12 J^2) \ , \tag{2.123}$$

giving a disorder average of,

$$\overline{J_{jklm}^2} = \frac{3! J^2}{N^3} \ , \qquad \overline{J_{jklm}} = 0 \ . \tag{2.124}$$

Remarkably, one can analytically compute the disorder averaged large-N correlation functions in the SYK model at finite temperature and strong coupling,

[3] In other words, the IP and IOP models not fully thermalizing. If they had been, the absence of chaos in these models would have been puzzling.

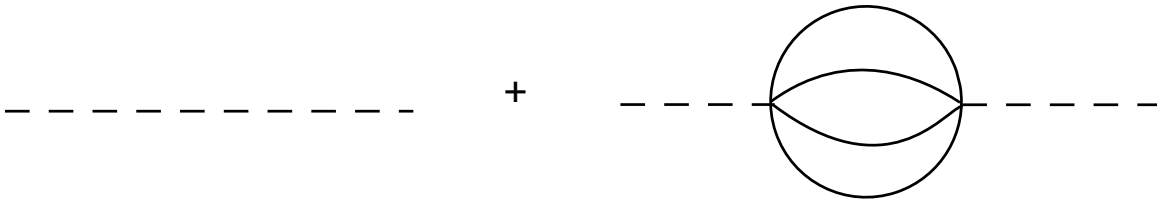


Figure 2.6: The dashed lines indicate J_{jklm}, while the sold lines are the fermions χ_i. Treating J_{jklm} as a quantum field, the quantum corrections to the two-point function are suppressed by $1/N^3$.

$\beta J \gg 1$. The two-point function exhibits exponential late time decay, see [104, 109, 87, 110]. The out-of-time-order four-point function exhibits exponential growth [87],

$$\langle \chi_i(t)\chi_j(0)\chi_i(t)\chi_j(0)\rangle \sim \frac{1}{N} e^{2\pi t/\beta} \ . \tag{2.125}$$

For studies of the four-point function, see [87, 101, 111, 112].

An important aspect of the SYK model is the quenched disorder: if the coupling J_{jklm} where instead a fixed constant, there would be additional Feynman diagrams that would contribute at leading order in $1/N$. Here we simply point out that the disorder J_{jklm} can be replaced by a quantum variable, as the quantum corrections are $1/N^3$ suppressed.

Recall that the disorder averaged expectation value of an operator O composed of the fields χ_i is,

$$\overline{\langle O\rangle} = \int DJ_{jklm}\ e^{-J_{jklm}^2 N^3/12J^2}\ \frac{\int D\chi_i\ O\ e^{-\int dt\, L}}{\int D\chi_i\ e^{-\int dt\, L}} \ . \tag{2.126}$$

The interpretation of (2.126) is that one first computes the expectation value $\langle O\rangle$ with some coupling J_{jklm} drawn from the distribution (2.123), and then averages over the J_{jklm}. If one were to instead treat J_{jklm} as a static quantum variable,

then the expectation value of O would be given by,

$$\langle O \rangle = Z^{-1} \int DJ_{jklm}\ D\chi_i\ O \exp\left(-N^3 J_{jklm}^2/12J^2 - \int dt\ L\right)\ . \tag{2.127}$$

In terms of Feynman diagrammatics, if J_{jklm} is a classical Gaussian-random parameter, then it has a two-point that is exactly $3!J^2/N^3$. If instead J_{jklm} is a quantum variable, then its leading two-point function can be chosen to be $3!J^2/N^3$, however this will receive quantum corrections, as shown in Fig. 2.6. Thus, generally (2.126) and (2.127) are different. However, for the SYK model, the first quantum correction is suppressed by a factor of $1/N^3$: the loop diagram in Fig. 2.6 has two J_{jklm} propagators, giving a factor of $(3!J^2/N^3)^2$. So, at leading order in $1/N$, (2.126) and (2.127) are the same. Equivalently, the effective action for J_{jklm} is

$$e^{-W[J_{jklm}]} = \int D\chi_i\ \exp\left(-J_{jklm}^2 N^3/12J^2 - \int dt\, L\right) = e^{-J_{jklm}^2 N^3/12J^2} + \dots\ , \tag{2.128}$$

at leading order in $1/N$. Note that the structure of the vacuum is different depending on if J_{jklm} is quenched disorder or a quantum field: the vacuum loop scales like N, and receives a correction of the same order from interactions with χ_i, as there is now a summation over the indices. This, however, is irrelevant for the purposes of connected correlation functions.

The variable J_{jklm} is still not yet a standard quantum variable, due to the constraint that it be static. There are a few somewhat artificial ways to achieve this. One could add to the action a term $\dot{J}_{jklm}\varphi$, where φ is some Lagrange multiplier field. A better option is to regard J_{jklm} as the momenta of harmonic oscillators

for which the frequency is taken to zero. Consider a harmonic oscillator with the standard Lagrangian, $(m\dot{x}^2 - m\omega^2 x^2)/2$. The Euclidean two-point function for the momentum is $\langle p(t)p(0)\rangle = m\omega e^{-\omega t}/2$. Now take the limit of $\omega \to 0$, so as to remove the time dependance. Letting $m\omega = 12J^2N^{-3}$, the momenta have the same two-point function as (2.124).

2.3.3 IP model

The IP model [81] is a quantum mechanical system, with a harmonic oscillator in the adjoint of $U(N)$ and a harmonic oscillator in the fundamental of $U(N)$, coupled through a trilinear interaction. The Hamiltonian for the IP model is given by (2.117). The two-point function is found by summing rainbow diagrams (see Fig. 2.7) and is reviewed in Sec. 2.3.3. The four-point function is given by a sum of ladder diagrams (see Fig. 2.8), which we evaluate in Sec. 2.3.3.

Two-point function

The bare zero temperature propagator for the fundamental is defined as,

$$G_0(t)\delta_{ij} \equiv \langle T a_i(t) a_j^\dagger(0)\rangle e^{iMt} \ . \tag{2.129}$$

Trivially, one has that,

$$G_0(t) = \theta(t), \qquad G_0(\omega) = \frac{i}{\omega + i\epsilon} \ . \tag{2.130}$$

It will be assumed that fundamental has a large mass, $M \gg T$, where T is the temperature. In this case, the bare finite temperature two-point function is the

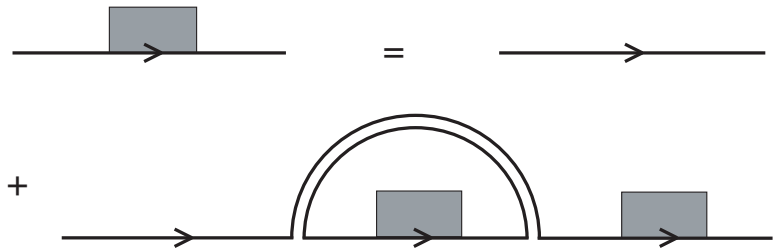


Figure 2.7: The Schwinger-Dyson equation for the propagator $G(\omega)$ in the IP model, in the planar limit. Arrows point from creation operators toward annihilation operators. A single line denotes the free propagator $G_0(\omega)$, a line with a shaded box is the dressed propagator $G(\omega)$, and a double line is the adjoint propagator $K(\omega)$. Iterating generates a sequence of nested rainbow diagrams.

same as the bare zero temperature two-point function.

The adjoints have no self-interaction, and the backreaction from interactions with the fundamental is suppressed by $1/N$. Thus, the propagator for the adjoint is that of a free oscillator in a thermal bath,

$$K(\omega) = \frac{i}{1-y}\left(\frac{1}{\omega^2 - m^2 + i\epsilon} - \frac{y}{\omega^2 - m^2 - i\epsilon}\right) , \tag{2.131}$$

where we have defined $y = e^{-m/T}$. It will be useful for later to note that in the limit that the adjoints become massless, $m \to 0$ ($y \to 1$), their two-point function becomes,

$$K(\omega) = \frac{2\pi}{1-y}\delta(\omega^2 - m^2) . \tag{2.132}$$

The planar two-point function for the fundamental is found by summing rainbow diagrams. The Schwinger-Dyson equation for the two-point function is given by (see Fig. 2.7):

$$G(\omega) = G_0(\omega) + \lambda G_0(\omega)G(\omega)\int \frac{d\omega'}{2\pi} G(\omega')K(\omega - \omega') , \tag{2.133}$$

where the 't Hooft coupling is $\lambda = g^2 N$. In general, such an integral equation is difficult. However, the assumption that $M \gg T$ implies that $G(t) = 0$ for $t < 0$. As a result, $G(\omega)$ has no poles in the upper half plane, allowing us to close the integration contour in (2.133) in the upper half plane ω' plane. Picking up the residues at $\omega' = \omega \pm m$, the Schwinger-Dyson equation turns into a difference equation,

$$G(\omega) = \frac{i}{\omega + i\epsilon}\Big(1 - \frac{\lambda}{1-y}\frac{G(\omega)}{2m}(G(\omega - m) + yG(\omega + m))\Big) \ . \tag{2.134}$$

This can be solved numerically [81], however to proceed analytically we take the limit of small adjoint mass and small 't Hooft coupling,

$$m \to 0, \qquad \nu^2 = \frac{2\lambda}{m(1-y)} = \text{const.} \tag{2.135}$$

In this limit one finds [81],

$$G(\omega) = \frac{2i}{\omega + \sqrt{\omega^2 - 2\nu^2}} \ . \tag{2.136}$$

Here the ω should really be an $\omega + i\epsilon$; we will generally suppress the $i\epsilon$, remembering that all the poles are in the lower half complex plane. The Fourier transform of the two-point function is a Bessel function,

$$G(t) = \int \frac{d\omega}{2\pi} G(\omega) e^{-i\omega t} = \frac{\sqrt{2}}{\nu t} J_1(\sqrt{2}\nu t)\, \theta(t) \ . \tag{2.137}$$

We will later encounter integrals of a similar form, so we show (2.137) in some

detail. For positive times, the ω contour in (2.137) wraps around the branch cut stretching from $-\sqrt{2}\nu < \omega < \sqrt{2}\nu$. [4] Using (2.136) and moving the square root to the numerator, we rewrite (2.137) as,

$$G(t) = \frac{i}{\nu^2}\int \frac{d\omega}{2\pi}\left(\omega - \sqrt{\omega^2 - 2\nu^2}\right)\, e^{-i\omega t}\ . \tag{2.138}$$

The integral of the first term vanishes, while the second is twice a line integral,

$$G(t) = \frac{1}{\nu^2}(1 - e^{i\pi})\int_{-\sqrt{2}\nu}^{\sqrt{2}\nu} \frac{d\omega}{2\pi}\,\sqrt{2\nu^2 - \omega^2}\, e^{-i\omega t}\ , \tag{2.139}$$

which gives (2.137). Now let us redo the calculation for the Fourier transform (2.137) slightly differently. Taking $G(\omega)$ in the form (2.136) and changing variable to,

$$x = \omega + \sqrt{\omega^2 - 2\nu^2}, \qquad \omega = \frac{x^2 + 2\nu^2}{2x}\ , \tag{2.140}$$

gives,

$$G(t) = \int \frac{dx}{2\pi}\frac{i}{x}\left(1 - \frac{2\nu^2}{x^2}\right) e^{-\frac{i}{2}\left(x + \frac{2\nu^2}{x}\right)t}\ . \tag{2.141}$$

The ω contour in (2.137) that hugs the branch cut maps into an x contour that is a circle of radius $\sqrt{2}\nu$ and centered around the essential singularity at the origin.

[4] Our ω integral was from $-\infty < \omega < \infty$. For positive time, we close the contour in the lower half plane. The branch cut is slightly below the real axis, and so is inside the contour. We can shrink the contour so that it hugs the branch cut. For negative times, the ω integral is closed in the upper half plane, and so gives zero. Also, our choice of location for the branch cut corresponds, for instance, to writing $\sqrt{\omega^2 - 2\nu^2} = \omega\exp\left(\frac{1}{2}\log(1 - 2\nu^2/\omega^2)\right)$ and taking the principal branch for the logarithm.

Using the integral representation of the Bessel function,

$$J_n(t) = \frac{i}{2\pi} \int dx \; x^{-n-1} e^{\frac{1}{2}t(x-x^{-1})} \; , \tag{2.142}$$

where the contour circles clockwise around the origin, we have,

$$G(t) = \left(J_0(\sqrt{2}\nu t) + J_2(\sqrt{2}\nu t) \right) \theta(t) \; , \tag{2.143}$$

which is equal to (2.137). At late time, $\nu t \gg 1$, the propagator decays as $G(t) \sim t^{-3/2}$.

Four-point function

We now turn to the connected four-point function. In the planar limit, it consists of a sum of ladder diagrams, as shown in Fig. 2.8. The ingoing momenta are ω_1, ω_2, while the outgoing momenta are ω_3, ω_4. [5] As in the case of the two-point function, to proceed analytically we must work in the limit specified in (2.135). In this limit, the propagator for the adjoint is given by (2.132).

Consider the ladder diagram that consists of a single rung. It is given by,

$$(-ig)^2 \int \frac{dp}{2\pi} G(\omega_1) G(\omega_1 - p) G(\omega_2) G(\omega_2 + p) K(p) \; . \tag{2.144}$$

Now inserting

$$\delta(p^2 - m^2) = \frac{1}{2m} \left[\delta(p - m) + \delta(p + m) \right] \tag{2.145}$$

[5] The ingoing momenta are drawn in Fig. 2.8 as coming from the upper left and lower right in order for the diagram to look planar.

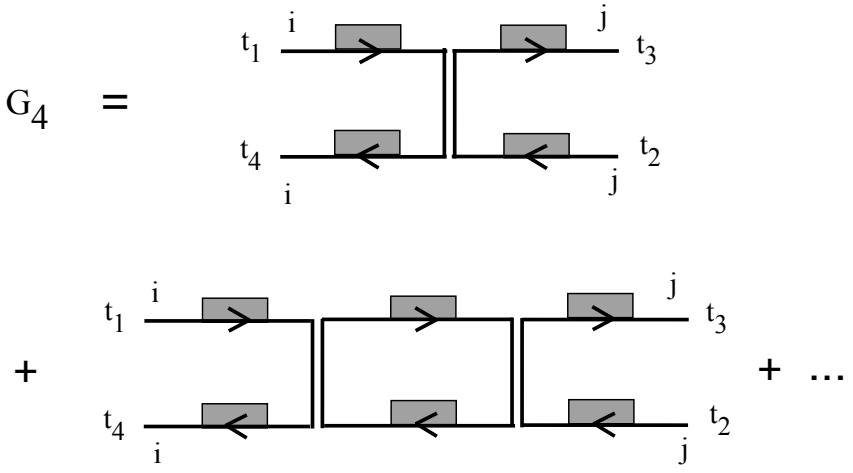


Figure 2.8: The planar four-point function G_4 (2.147) in the IP model. Ladders with $n = 1$ and $n = 2$ rungs are shown.

into (2.144), evaluating the integral, and then taking the $m \to 0$ limit, yields for (2.144),

$$\frac{(-ig)^2}{m(1-y)}\, G(\omega_1)^2 G(\omega_2)^2 \ . \tag{2.146}$$

We now sum all the ladder diagrams. As a result of the limit (2.135), all the pieces appearing in the Feynman diagrams are on-shell. Defining $G_4(\omega_1, \omega_2, \omega_3, \omega_4) = \delta(\omega_1 - \omega_3)\delta(\omega_2 - \omega_4)G_4(\omega_1, \omega_2)$, and letting n denote the number of rungs, we have

$$NG_4(\omega_1, \omega_2) = \sum_{n=1}^{\infty} \left(\frac{-\lambda}{m(1-y)} \right)^n (G(\omega_1)G(\omega_2))^{n+1} = \frac{-\frac{\nu^2}{2} G(\omega_1)^2 G(\omega_2)^2}{1 + \frac{\nu^2}{2} G(\omega_1)G(\omega_2)} \ , \tag{2.147}$$

where ν was defined in (2.135).

The Fourier transform of (2.147) gives the position space four-point function,

$$NG_4(t_{31}, t_{42}) = -\frac{\nu^2}{2} \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{G(\omega_1)G(\omega_2)}{\frac{\nu^2}{2} + G(\omega_1)^{-1}G(\omega_2)^{-1}} e^{-i\omega_1 t_{31}} e^{-i\omega_2 t_{42}} , \quad (2.148)$$

where we have defined $t_{31} \equiv t_3 - t_1$, $t_{42} \equiv t_4 - t_2$. In addition, $G(\omega)$ really denotes $G(\omega + i\epsilon)$; we suppress the $i\epsilon$, remembering that, if we are using G in a time-ordered correlator, all the poles are in the lower-half complex plane.

Free propagator

The propagator entering the four-point function (2.148) is given by (2.136). As a warmup, it is useful to study (2.148) with the free propagator (2.130), rather than the dressed one. In this case we have,

$$N\bar{G}_4(t_{31}, t_{42}) = \nu^2 \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{1}{\omega_1\omega_2} \frac{1}{\nu^2 - 2\omega_1\omega_2} e^{-i\omega_1 t_{31}} e^{-i\omega_2 t_{42}} . \quad (2.149)$$

Performing the ω_2 integral, and closing the contour in the lower half plane, we pick up poles at $\omega_2 = 0$ and $\omega_2 = \nu^2/2\omega_1$,

$$N\bar{G}_4(t_{31}, t_{42}) = -\theta(t_{42})\theta(t_{31}) + \theta(t_{42}) \int \frac{d\omega_1}{2\pi i} \frac{1}{\omega_1} e^{-i\left(\omega_1 t_{31} + \frac{\nu^2}{2\omega_1} t_{42}\right)} . \quad (2.150)$$

Using the integral representation of the Bessel function (2.142), we get,

$$N\bar{G}_4(t_{31}, t_{42}) = \left(J_0(\sqrt{2t_{31}t_{42}}\nu) - 1\right) \theta(t_{31})\theta(t_{42}) . \quad (2.151)$$

Eq. 2.151 is the time-ordered four-point function, as evidenced by the theta functions. We can obtain the out-of-time-order four-point function by dropping the theta functions. In particular, setting $t_{31} = -t_{42} = t$ gives,

$$NC(t) = I_0(\sqrt{2}\nu t) - 1 \ . \tag{2.152}$$

In the limit $\nu t \gg 1$,

$$C(t) \to \frac{1}{2^{3/4}\sqrt{\pi\nu t}N} e^{\sqrt{2}\nu t} \ . \tag{2.153}$$

By summing only a subset of the Feynman diagrams: the ladder diagrams with undressed propagators, we have found exponential growth in the out-of-time-order four-point function. While intriguing, using the free propagator is certainly not legitimate, as it violates unitarity; classically it would be equivalent to violating Liouville's theorem. However, before evaluating (2.148) with the dressed propagator, it will be instructive to study (2.149) a bit further.

Returning to (2.150), and taking the limit of large t_{31}, t_{42}, we approximate the integral via the method of steepest descent (see Appendix 2.3.6). This involves deforming the contour of integration in order for it to pass through the saddle point, at an angle so as to maintain constant phase. The saddle point of the exponent,

$$f(\omega_1) = \omega_1 t_{31} + \frac{\nu^2}{2\omega_1} t_{42} \ , \tag{2.154}$$

occurs at $\tilde{\omega}_1 = \pm\nu\sqrt{t_{42}/2t_{31}}$. As we continue from a time-ordered four-point function, to an out-of-time-order four-point function, $t_{42} \to -t_{42}$, the saddle moves off of the real axis and onto the imaginary axis. At $t_{31} = -t_{42} = t$, the saddle is at $\tilde{\omega}_1 = \pm i\nu/\sqrt{2}$. The leading exponent in the integral in (2.150) can therefore be

approximated by,

$$e^{-itf(\tilde{\omega}_1)} = e^{\sqrt{2}\nu t} \ , \tag{2.155}$$

reproducing (2.153).

Let us also reproduce (2.151) by returning to (2.147) and computing the Fourier transform of each term before taking the sum. From (2.147) and (2.130) we have,

$$\bar{G}_4(\omega_1, \omega_2) = -\sum_{n=1}^{\infty} \left(\frac{\nu^2}{2}\right)^n \frac{1}{(\omega_1\omega_2)^{n+1}} \ . \tag{2.156}$$

The Fourier transform gives,

$$\bar{G}_4(t_{31}, t_{42}) = \sum_{n=1}^{\infty} \left(\frac{\nu^2}{2}\right)^n \frac{(-t_{31}t_{42})^n}{(n!)^2} \theta(t_{31})\theta(t_{42}) = \left(J_0(\sqrt{2t_{31}t_{42}}\nu) - 1\right) \theta(t_{31})\theta(t_{42}) \ , \tag{2.157}$$

where we have made use of the series definition of the Bessel function.

The expression (2.157) is easy to see directly in time-space. Since the free two-point function for the fundamental is simply $\theta(t)$ (2.130), a ladder diagram with n rungs will have $n+1$ propagators for the fundamentals on each of the two sides. For one such side we have a product of theta functions, with the time insertions of the rungs integrated over. For the top side,

$$\int_{t_1}^{t_3} dt_{a_n} \ldots \int_{t_1}^{t_{a_3}} dt_{a_2} \int_{t_1}^{t_{a_2}} dt_{a_1} = \frac{1}{n!} t_{31}^n \ , \tag{2.158}$$

and similarly a factor of $t_{42}^n/n!$ from the bottom side. Accounting for the coupling at each vertex, $-ig$, as well as the sum over indices, and the factor of $m^{-1}(1-y)^{-1}$ coming from the adjoint propagator, we recover the sum in (2.157).

If we wish to form an out-of-time-order four-point function, for instance with

$t_{42} < 0$, then on the bottom edge of the ladder diagrams, time runs backwards: we must use a two-point function that is $\theta(-t)$ rather than $\theta(t)$. In addition, since time is running backwards on the bottom edge, the interactions come with a factor of ig, instead of $-ig$. This results in the elimination of the minus sign in the sum in (2.157), and correspondingly gives exponential growth.

Dressed propagator

We now return to the frequency-space four-point function (2.147), and evaluate the Fourier transform (2.148), this time using the full dressed propagator. Inserting the propagator $G(\omega)$ (2.136) into (2.148) gives,

$$\begin{aligned} NG_4(t_{31}, t_{42}) &= -G(t_{31})G(t_{42}) \\ &+ 4\int \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\frac{1}{2\nu^2 - (\omega_1 + \sqrt{\omega_1^2 - 2\nu^2})(\omega_2 + \sqrt{\omega_2^2 - 2\nu^2})} e^{-i\omega_1 t_{31}}\ e^{-i\omega_2 t_{42}}\ , \end{aligned} \tag{2.159}$$

where we have split off a $G(\omega_1)G(\omega_2)$ from (2.147), giving the first term in (2.159). Changing integration variables to $x_i = \omega_i + \sqrt{\omega_i^2 - 2\nu^2}$ gives,

$$\begin{aligned} NG_4(t_{31}, t_{42}) &= -G(t_{31})G(t_{42}) \\ &+ \int \frac{dx_1}{2\pi}\ \frac{dx_2}{2\pi}\ \left(1 - \frac{2\nu^2}{x_1^2}\right)\left(1 - \frac{2\nu^2}{x_2^2}\right)\ \frac{1}{2\nu^2 - x_1 x_2}\ e^{-\frac{i}{2}\left(x_1 + \frac{2\nu^2}{x_1}\right)t_{31}}\ e^{-\frac{i}{2}\left(x_2 + \frac{2\nu^2}{x_2}\right)t_{42}}\ . \end{aligned} \tag{2.160}$$

Our goal is to see if (2.160) exhibits exponential growth; if this does occur, it will be in the out-of-time-order regime, such as $t_{31} = -t_{42} = t$. We consider

the late time limit, [6] and approximate (2.160) via the saddle point method (Appendix 2.3.6): we seek to deform the contours of integration of x_1, x_2 such that they pass through a saddle, at an angle such that the phase is constant. If we are away from the poles of the integrand, the saddle point occurs at $x_i = \pm\sqrt{2}\nu$, which clearly only gives oscillatory behavior. Now consider the regions at the poles of the integrand, at $x_1 x_2 = 2\nu^2$. This a peculiar region, as

$$\omega = \frac{x}{2} + \frac{\nu^2}{x} \tag{2.161}$$

is invariant under $x \to 2\nu^2/x$. Inserting this $x_2 = 2\nu^2/x_1$ into the exponent in (2.160) , the exponent becomes,

$$\exp\left(-\frac{i}{2}\left(x_1 + \frac{2\nu^2}{x_1}\right)(t_{31} + t_{42})\right) \ , \tag{2.162}$$

which does not give rise to the exponential growth indicative of chaos. Moreover, for $t_{31} = -t_{42}$, it simply vanishes.

2.3.4 IOP model

We now turn to the IOP model [83]. Like the IP model, this is a quantum mechanical system, with a harmonic oscillator in the adjoint of $U(N)$ and a harmonic oscillator in the fundamental of $U(N)$. However, the interaction is now quartic in the oscillators (2.117), and quadratic in the $U(N)$ charges. The latter property makes the IOP model more analytically tractable than the IP model,

[6] Since we are working in the planar limit, late time is still before the scrambling time, which scales as $\log N$.

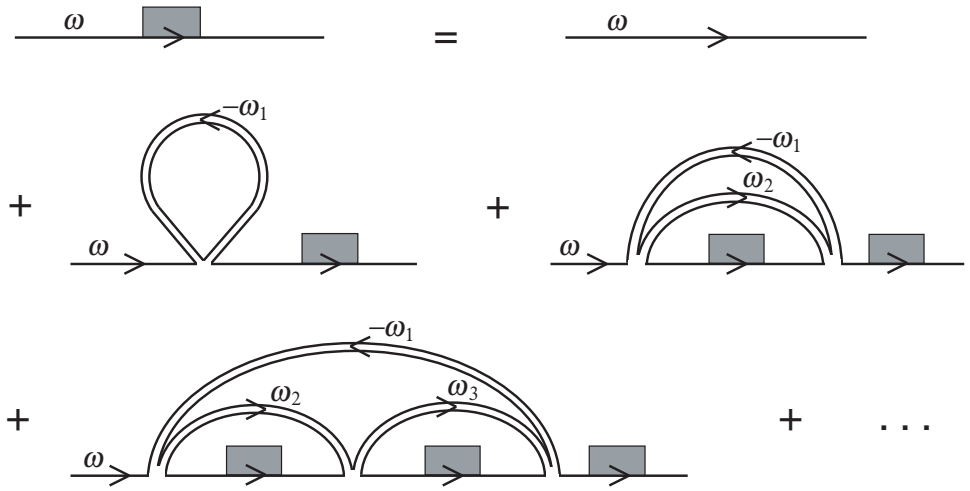


Figure 2.9: Planar Feynman graphs for the fundamental propagator $G(\omega)$ (2.166) in the IOP model. The shaded rectangles mark the full planar propagators. Arrows point from creation operators toward annihilation operators. The graphs for $n = 0$, 1, 2 are shown.

although diagrammatically it is more involved. As in the IP model, we consider the limit in which the fundamental is heavy, $M \gg T$. However, we can now obtain analytic results at any mass m for the adjoint. We review the two-point function in Sec. 2.3.4, and compute the four-point function in Sec. 2.3.4.

Two-point function

The bare propagator for the fundamental is the same as in the IP model (2.130). The propagator for the adjoint is that of free harmonic oscillator in a thermal bath, defined by $L(t)\delta_{il}\delta_{jk} = \langle T A_{ij}(t) A^\dagger_{kl}(0)\rangle$, and giving,

$$L(\omega) = \frac{i}{1-y}\left[\frac{1}{\omega - m + i\epsilon} - \frac{y}{\omega - m - i\epsilon}\right] . \tag{2.163}$$

The Schwinger-Dyson equation for the planar two-point function for the fundamental is (see Fig. 2.9),

$$G(\omega) = G_0(\omega) + G_0(\omega)G(\omega)\sum_{n=0}^{\infty} S_n(\omega) \ , \tag{2.164}$$

$$S_n(\omega) = (-ihN)^{n+1}\int \frac{d^{n+1}\vec{\omega}}{(2\pi)^{n+1}} L(-\omega_1)\prod_{l=1}^{n}[G(\omega-\omega_{l+1}-\omega_1)L(\omega_{l+1}) \tag{2.165}$$

As G only has poles in the lower-half plane, we can close the ω_i integrals in the lower-half plane and pick up residues only from L. This leads to an algebraic equation for G, with the solution [83],

$$G(\omega) = \frac{2i}{\lambda+\omega+\sqrt{(\omega-\omega_+)(\omega-\omega_-)}}, \qquad \omega_\pm = \lambda\frac{1+y\pm 2\sqrt{y}}{1-y} \ , \tag{2.166}$$

where the 't Hooft coupling is $\lambda = hN$. The propagator has a branch cut from ω_- to ω_+, leading to late-time power law decay, $t^{-3/2}$.

Four-point function

We now turn to the four-point function in the planar limit. The connected four-point function is found by summing ladder-like diagrams, shown in Fig. 2.12, where each "rung" of the ladder is found by summing an infinite number of diagrams, like the ones shown in Fig. 2.10. We warm up by computing the four-point function in the limit of small adjoint mass m, before doing the calculation for arbitrary m.

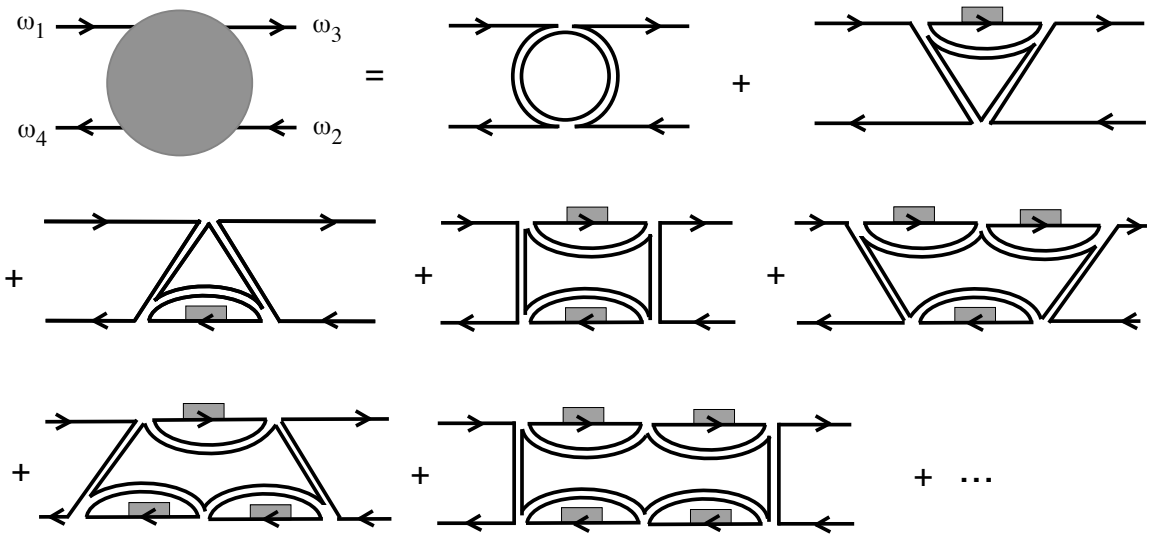


Figure 2.10: Planar diagrams contributing to a "rung" Γ in the IOP model. Diagrams with $n, m = 0, 1, 2$ are shown.

Small adjoint mass

We start with the limit $m \to 0$. In particular,

$$m \to 0, \qquad \kappa \equiv \frac{\lambda}{1-y} \ , \tag{2.167}$$

where κ is held constant. In this limit, the two-point functions for the adjoint (2.163) and the fundamental (2.166) become,

$$L(\omega) = \frac{1}{1-y} 2\pi\delta(\omega - m) \ , \tag{2.168}$$

$$G(\omega) = \frac{2i}{\omega + \sqrt{\omega(\omega - 4\kappa)}} \ . \tag{2.169}$$

To compute the four-point function, we first sum the diagrams shown in Fig. 2.10, to get

$$\Gamma(\omega_1, \omega_2, \omega_3, \omega_4) = \frac{(2\pi)^2}{N}\Gamma(\omega_1, \omega_2)\delta(\omega_{13})\delta(\omega_{24}) \ , \tag{2.170}$$

where $\omega_{ij} \equiv \omega_i - \omega_j$ and,

$$\Gamma(\omega_1, \omega_2) = \sum_{n,m=0}^{\infty} G(\omega_1)^n G(\omega_2)^m (-i\kappa)^{n+m+2} = \frac{-\kappa^2}{(1 + i\kappa G(\omega_1))(1 + i\kappa G(\omega_2))} \ , \tag{2.171}$$

where the index n/m labels the number of intermediate fundamental propagators on the top/bottom edge. As in the IP model, as a result of the $m \to 0$, all intermediate propagators are on-shell. The four-point function is given by the ladder-like sum of the Γ (see Fig. 2.12),

$$\begin{aligned} NG_4(\omega_1, \omega_2) &= \sum_{k=1}^{\infty} \Gamma(\omega_1, \omega_2)^k (G(\omega_1)G(\omega_2))^{k+1} \\ &= \frac{G(\omega_1)G(\omega_2)}{1 - \Gamma(\omega_1, \omega_2)G(\omega_1)G(\omega_2)} - G(\omega_1)G(\omega_2) \ . \end{aligned} \tag{2.172}$$

Inserting (2.171) into (2.172) gives the frequency-space four-point function $G_4(\omega_1, \omega_2, \omega_3, \omega_4) = (2\pi)^2\delta(\omega_{13})\delta(\omega_{24})G_4(\omega_1, \omega_2)$ where,

$$NG_4(\omega_1, \omega_2) = \frac{-\kappa^2 G(\omega_1)^2 G(\omega_2)^2}{1 + i\kappa(G(\omega_1) + G(\omega_2))} \ . \tag{2.173}$$

Like in the IP model, we find exponential growth in the out-of-time-order four-point function if we only sum the diagrams containing the free propagator: (2.173) with (2.130) and $t_{31} = -t_{42} = t$ gives a four-point function $\sim N^{-1}\exp(2\kappa t)$ for large t.

Now consider (2.173) with (2.169). The position-space four-point function is thus,

$$NG_4(t_{31},t_{42}) = \int \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\frac{-\kappa^2 G(\omega_1)G(\omega_2)}{G(\omega_1)^{-1}G(\omega_2)^{-1}+i\kappa(G(\omega_1)^{-1}+G(\omega_2)^{-1})}\, e^{-i\omega_1 t_{31}}e^{-i\omega_2 t_{42}} \ . \tag{2.174}$$

Changing integration variables to $x_i = \omega_i + \sqrt{\omega(\omega-4\kappa)}$, (2.174) becomes

$$\begin{aligned} NG_4(t_{31},t_{42}) \ &= \ \int \frac{dx_1}{2\pi}\frac{dx_2}{2\pi}\frac{(x_1-4\kappa)\ (x_2-4\kappa)}{(x_1-2\kappa)^2(x_2-2\kappa)^2}\frac{-4\kappa^2}{x_1x_2-2\kappa(x_1+x_2)} \\ &\quad \times\ e^{-i\frac{x_1^2}{2(x_1-2\kappa)}t_{31}}e^{-i\frac{x_2^2}{2(x_2-2\kappa)}t_{42}} \ . \end{aligned} \tag{2.175}$$

We approximate the integral by taking the limit of large time separations, and looking for saddle points which could give rise to exponential growth. Picking up the pole at $x_1x_2 = 2\kappa(x_1+x_2)$, the exponent becomes,

$$\exp\left(-i\frac{x_1^2}{2(x_1-2\kappa)}(t_{31}+t_{42})\right) \ . \tag{2.176}$$

Like in the IP model, there is no exponential growth.

Arbitrary adjoint mass

We now compute the four-point function, with the adjoints taking arbitrary mass m. The Feynman diagrams contributing to "rung" Γ are shown in Fig. 2.10. A term in this sum, having n fundamental propagators on the upper edge and m

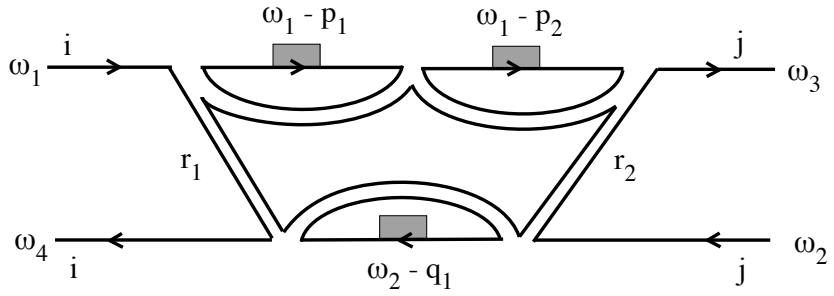


Figure 2.11: One of the diagrams entering Γ in Fig. 2.10, given by $n = 2, m = 1$ in (2.177).

on the lower, is given by,

$$(-i\lambda)^{n+m+2} \int \frac{d^n\vec{p}}{(2\pi)^n} \frac{d^m\vec{q}}{(2\pi)^m} \frac{dr_1}{2\pi} L(r_1)L(r_2) \prod_{i=1}^{n} G(\omega_1 - p_i)L(r_1 + p_i) \prod_{j=1}^{m} G(\omega_2 - q_j)L(r_2 + q_j) \,, \tag{2.177}$$

where the ingoing frequencies are ω_1, ω_2, the outgoing frequencies are ω_3, ω_4, and we have defined $r_2 = r_1 + \omega_1 - \omega_3$, and suppressed an overall factor of N^{-1}. In Fig. 2.11 the $n = 2, m = 1$ diagram from Fig. 2.10 is shown in more detail. Since $G(\omega_1 - p_i)$ has poles in the upper half p_i plane, we close the contour in the lower half plane. Similarly for the q_i integral. This gives for (2.177),

$$\frac{(-i\lambda)^{n+m+2}}{(1-y)^{n+m}} \int \frac{dr_1}{2\pi}\, L(r_1)L(r_2)\; G(\omega_1 + r_1 - m)^n\, G(\omega_2 + r_2 - m)^m \,. \tag{2.178}$$

Evaluating the integral over r_1 by closing the contour in the upper half-plane, (2.178) becomes, [7]

[7]The adjoint propagator L is given by (2.163). We denote the epsilon for $L(r_1)$ by ϵ_1, and for $L(r_2)$ by ϵ_2. Without loss of generality, we choose $\epsilon_2 > \epsilon_1$. One can equally well choose $\epsilon_2 < \epsilon_1$; this can be seen by rewriting $(\omega_1 - \omega_3 - i(\epsilon_2 - \epsilon_1))^{-1} = (\omega_1 - \omega_3 + i(\epsilon_2 - \epsilon_1))^{-1} + 2\pi i\delta(\omega_1 - \omega_3)$, and noting that $\delta(\omega_1 - \omega_3)(G(\omega_3)^n G(\omega_2)^m - G(\omega_1)^n G(\omega_4)^m) = 0$.

$$\frac{iy(-i\lambda)^{n+m+2}}{(1-y)^{n+m+2}}\left[G(\omega_1)^n G(\omega_4)^m\left(\frac{1}{\omega_1-\omega_3+i\epsilon_1+i\epsilon_2}-\frac{y}{\omega_1-\omega_3+i\epsilon_1-i\epsilon_2}\right)\right.$$
$$\left.+\,G(\omega_3)^n G(\omega_2)^m\left(\frac{1}{\omega_3-\omega_1+i\epsilon_2+i\epsilon_1}-\frac{y}{\omega_3-\omega_1-i\epsilon_1+i\epsilon_2}\right)\right]\ . \quad (2.179)$$

To sum over all the diagrams contributing to Γ (see Fig. 2.10), we must sum (2.179) over n, m from 0 to infinity. This gives $\Gamma = y\tilde{\Gamma}$ where,

$$-i\tilde{\Gamma}(1,2,3,4) = \frac{z(1,4)}{\omega_1-\omega_3+i\epsilon} - \frac{y\,z(1,4)+(1-y)\,z(2,3)}{\omega_1-\omega_3-i\epsilon}, \quad (2.180)$$

where we have defined,

$$z(j,l) = \frac{-\kappa^2}{(1+i\kappa G(\omega_j))(1+i\kappa G(\omega_l))}\ , \quad (2.181)$$

and have simplified notation to denote ω_j by j, and recall that $\kappa \equiv \lambda/(1-y)$. One can also rewrite $\tilde{\Gamma}$ in (2.180) as,

$$y\tilde{\Gamma} = y^2\, z(1,4)\, 2\pi\delta(\omega_1-\omega_3) + y(1-y)\left[\frac{i\ z(1,4)}{\omega_1-\omega_3+i\epsilon} - \frac{i\ z(2,3)}{\omega_1-\omega_3-i\epsilon}\right]\ , \quad (2.182)$$

which, recalling that $\omega_1+\omega_2 = \omega_3+\omega_4$, is manifestly symmetric under $\omega_1 \leftrightarrow \omega_2, \omega_3 \leftrightarrow \omega_4$.

Attaching external propagators to (2.180) gives the first term in the sum for the four-point function shown in Fig. 2.12. The second term requires gluing two of the $\tilde{\Gamma}$ together,

$$(\tilde{\Gamma}\times\tilde{\Gamma})(1,2,3,4) \equiv \int \frac{d\omega_a}{2\pi}\, G(a)G(\bar{a})\,\tilde{\Gamma}(1,\bar{a},a,4)\,\tilde{\Gamma}(a,2,3,\bar{a})\ , \quad (2.183)$$

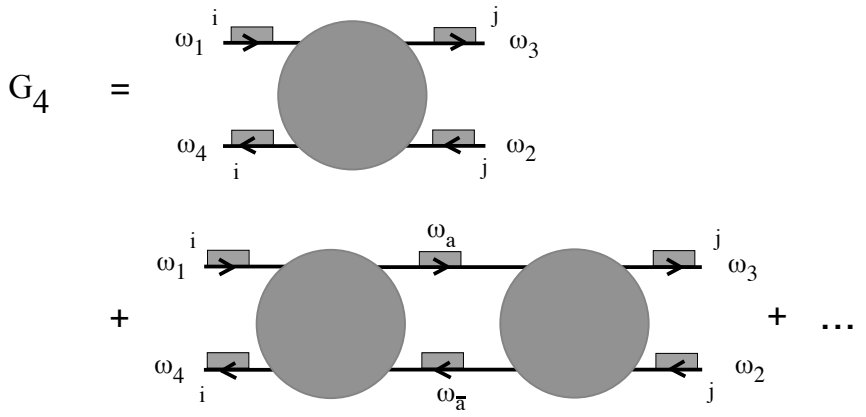


Figure 2.12: The planar four-point function consists of ladders formed by gluing together the diagrams shown in Fig. 2.10.

where $\omega_{\bar{a}} = \omega_a + \omega_4 - \omega_1$. Performing the integral in (2.183) by closing the ω_a contour in the upper-half plane gives,

$$\tilde{\Gamma} \times \tilde{\Gamma} = G(2)G(3)\, z(2,3)\, \tilde{\Gamma} + \frac{i(1-y)}{\omega_1 - \omega_3 + i\epsilon}\, G(1)G(4)\, z(1,4)\Big(z(1,4) - z(2,3)\Big)\ , \tag{2.184}$$

where both the $\tilde{\Gamma} \times \tilde{\Gamma}$ on the left, and the $\tilde{\Gamma}$ on the right, are functions of the external ω_i.

Let us simplify notation and let $(\tilde{\Gamma})^2$ denote $\tilde{\Gamma} \times \tilde{\Gamma}$, defined by (2.183). We define $(\tilde{\Gamma})^n$, arising from gluing n of the $\tilde{\Gamma}$ together, in an analogous fashion,

$$(\tilde{\Gamma})^n(1,2,3,4) \equiv \int \frac{d\omega_a}{2\pi}\, G(a)G(\bar{a})\, (\tilde{\Gamma})^{n-1}(1,\bar{a},a,4)\, \tilde{\Gamma}(a,2,3,\bar{a})\ . \tag{2.185}$$

We compute $(\tilde{\Gamma})^n$ iteratively, by gluing together $(\tilde{\Gamma})^{n-1}$ and $\tilde{\Gamma}$. The result is,

$$(\tilde{\Gamma})^n = G(2)G(3)\, z(2,3)(\tilde{\Gamma})^{n-1} + \frac{i(1-y)}{\omega_1 - \omega_3 + i\epsilon}\Big(G(1)G(4)z(1,4)\Big)^{n-1}\Big(z(1,4) - z(2,3)\Big)\ , \tag{2.186}$$

where we have for convenience expressed $(\tilde{\Gamma})^n$ in terms of $(\tilde{\Gamma})^{n-1}$. Next, we sum all the $(\tilde{\Gamma})^n$. Denoting the sum by S,

$$S = \sum_{n=1}^{\infty} (\Gamma)^n \ , \tag{2.187}$$

where recall that $\Gamma = y\tilde{\Gamma}$, and separating off the $n = 1$ term and using (2.186) for the rest, we get,

$$S\Big(1-yG(3)G(2)z(2,3)\Big) = \Gamma + \frac{i}{\omega_1 - \omega_3 + i\epsilon} \frac{y^2(1-y)G(1)G(4)z(1,4)\Big(z(1,4) - z(2,3)\Big)}{1 - yG(1)G(4)z(1,4)} \ . \tag{2.188}$$

The four-point function is given by S, with external propagators attached.

Thus, the connected four-point function for the IOP model in the planar limit is,

$$\begin{aligned} &NG_4(1,2,3,4) = \\ &A(1,2,3,4)\, 2\pi\delta(\omega_1+\omega_2-\omega_3-\omega_4) \left(yz(1,4) 2\pi\delta(\omega_1 - \omega_3) + y(1-y)\frac{iB(1,2,3,4)}{\omega_1 - \omega_3 + i\epsilon} \right) \ , \end{aligned} \tag{2.189}$$

where

$$A(1,2,3,4) = \frac{G(1)G(2)G(3)G(4)}{1 - yG(2)G(3)z(2,3)} \ , \tag{2.190}$$

$$B(1,2,3,4) = \frac{z(1,4) - z(2,3)}{1 - yG(1)G(4)z(1,4)} \ , \tag{2.191}$$

where j denotes the frequency ω_j, the propagator $G(i)$ for the fundamental is given

by (2.166), the constant y is the Boltzmann factor $y = e^{-m/T}$ where m is the mass of the adjoint and T is the temperature, and $z(j,l)$ was defined in (2.181) and is a function of $G(j)$, $G(l)$, and $\kappa = \lambda/(1-y)$, where λ is the 't Hooft coupling. In the limit of small adjoint mass m ($y \to 1$), the first term in (2.189) survives and reproduces the earlier result (2.173). The out-of-time-order four-point function does not exhibit exponential growth with time, for reasons similar to those seen in the small adjoint mass limit (2.175, 2.176); see Appendix 2.3.7.

2.3.5 Discussion

The absence of exponential growth in the out-of-time-order four-point function implies that the IOP model is not chaotic. In fact, there is a heuristic way to understand the absence of chaos in the IOP model. The interacting part of the Hamiltonian (2.118) can be written as,

$$H_{\text{int}} = -h\, q_{li} Q_{il}, \qquad q_{li} = -a_i^\dagger a_l, \quad Q_{il} = A_{ik}^\dagger A_{kl} \ . \tag{2.192}$$

As a result of the absence of self-interactions for the adjoints, combined with the assumption of large fundamental mass $M \gg T$, the number of fundamentals is time-independent and,

$$a_i(t) = e^{-ihQ_{il}t} a_l(0) \ . \tag{2.193}$$

Since Q is a hermitian matrix, it has real eigenvalues, and so the norm of the a_i operators does not grow.

If we relax the assumption that $M \gg T$, the above argument is no longer applicable, though this may not be sufficient to make the model chaotic. Heuris-

tically, chaos is associated with rapid growth. As we evolve a fundamental, it is emitting and absorbing adjoints. Since the adjoints have no self-interaction, and conversion of an adjoint into two fundamentals is suppressed by $1/N$, the only way for the adjoints to continue evolving in between emissions and absorptions is if they interact with fundamentals in the thermal bath.

It may be useful to modify the IOP model, so as to have several flavors of fundamentals. Also, the interaction (2.192) can written in terms of the quadratic Casimirs, $-hq \cdot Q = \frac{1}{2}h\mathrm{Tr}(q^2 + Q^2 - (q+Q)^2)$, allowing a computation of the two-point function at finite N through a sum over Young tableaux [83]. One could study the four-point function in this way as well.

Acknowledgements

We thank O. Aharony, D. Berenstein, T. Grover, A. Kitaev, Z. Komargodski, A. Lawrence, D. Roberts, M. Smolkin, S. Shenker, B. Shraiman, M. Srednicki, D. Stanford, A. Zhiboedov for helpful discussions. The work of BM was supported by the NSF Graduate Research Fellowship Grant No. DGE-1144085. The work of JP and VR was supported by NSF Grants PHY11-25915 and PHY13-16748. The work of JS was supported by a KITP graduate fellowship under grant PHY11-25915, by the Natural Sciences and Engineering Research Council of Canada, and by grant 376206 from the Simons Foundation.

2.3.6 Appendix A: Steepest descent

In this appendix, we review some aspects of evaluating integrals by the method of steepest descent, see e.g. [113]. Consider an integral of the form,

$$\int dz\ g(z)\, e^{-itf(z)}\ , \qquad t \gg 1\ , \tag{2.194}$$

where the integral is evaluated along some contour. For now, let $g(z), f(z)$ be smooth functions; we will discuss later how to relax this assumption. Since $t \gg 1$, the integrand generically undergoes rapid oscillations which cancel out. The idea will be to deform the contour of integration so as to follow a path for which the phase remains constant. As long as we do not cross any singularities, we are free deform the contour. Splitting $f(z)$ into a real and imaginary part,

$$f(z) = u(z) + iv(z)\ , \tag{2.195}$$

we need to deform the contour to follow a path of constant $u(z)$. The most relevant region of the integrand is one in which the real part is maximized. Letting $z = a + ib$,

$$\frac{\partial v}{\partial a} = \frac{\partial v}{\partial b} = 0\ . \tag{2.196}$$

As a result of the Cauchy-Riemann equations, this amounts to finding the saddle points, $f'(z) = 0$. Therefore, the prescription for approximating (2.194) is to focus on the vicinity of the dominant saddle point, and choose a direction for the contour that moves away from the saddle point so as to maintain constant phase $u(z)$.

As an example, consider the integral representation of the Bessel function,

$$K_0(t) = \frac{1}{2}\int_{-\infty}^{\infty} dx \frac{e^{-itx}}{\sqrt{1+x^2}} \tag{2.197}$$

This has a branch cut, $x \in (-i\infty, -i) \cup (i, i\infty)$. We perform a change of variables, $x = \sinh u$, thereby bringing (2.197) into the form (2.194),

$$\frac{1}{2}\int_{-\infty}^{\infty} du \exp(-it \sinh u) \ . \tag{2.198}$$

Extermizing $f(u) = \sinh u$, the saddle points are at $u = \pm\pi i/2$. The line of constant phase passing through the saddle points is one that runs along the imaginary axis. We deform the contour so that it runs along $-\infty < u < -i\pi/2$. Moving downward from $u_0 = -i\pi/2$ is a direction of steepest descent. In the vicinity of the saddle,

$$f(u) = f(u_0) + \frac{f''(u_0)}{2}(u-u_0)^2 + \dots \ . \tag{2.199}$$

Defining a new variable z as $u = u_0 - iz$, (2.198) becomes,

$$\int_0^{\infty} dz \exp\left(-t - t\frac{z^2}{2}\right) = \sqrt{\frac{\pi}{2t}} e^{-t} \ , \tag{2.200}$$

which is the correct large t expansion of $K_0(t)$.

We have so far discussed approximating (2.194) by the behavior near the saddle point. There are several contexts in which other regions may be relevant. If the contour has endpoints, then one must analyze the behavior near the endpoints. Additionally, if $g(z)$ has singularities, then one must analyze the integrand near those regions as well. In particular, it may happen that there is no way to deform

the contour into the relevant steepest descent contour, without passing through singularities. If the singularity of $g(z)$ is a simple pole, then we may simply deform through it, picking up the contribution of the pole. If, instead, $g(z)$ has a branch cut or an essential singularity, we must analyze the integrand in the vicinity of these regions.

For instance, consider again approximating (2.197), but without changing variables. In this case, $g(x) = (1+x^2)^{-1/2}$ and $f(x) = x$. The exponential has no saddle points, so we focus on the regions where $g(x)$ is large: near $x = \pm i$. We integrate along a direction running parallel to the imaginary axis, as we still need to maintain constant phase for the exponent. Letting $x = -i - \rho i$, with $\rho \ll 1$ so that $\sqrt{1+x^2} \approx \sqrt{2\rho}$, (2.197) is approximated by,

$$\frac{e^{-t}}{\sqrt{2}} \int_0^\infty d\rho \frac{e^{-\rho t}}{\sqrt{\rho}} \ , \tag{2.201}$$

where we have extended the range of integration to infinite ρ, as its contribution is negligible. Evaluating (2.201) gives (2.200).

2.3.7 Appendix B: Four-point integral

The four-point function for the IOP model is,

$$G_4(t_1,t_2,t_3,t_4) = \int \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi}\frac{d\omega_3}{2\pi}\, G_4(\omega_1,\omega_2,\omega_3,\omega_4)\, e^{-i\omega_1 t_{41} - i\omega_2 t_{42} - i\omega_3 t_{34}} \ , \tag{2.202}$$

where $\omega_4 = \omega_1 + \omega_2 - \omega_3$ and $G_4(\omega_1,\omega_2,\omega_3,\omega_4)$ is given by (2.189).

Our eventual interest is the out-of-time-order four-point with time separations $t_{41} = 0$, $t_{34} = -t_{42} = t$ and large t. At large t, the exponent in (2.202) undergoes

rapid oscillations as ω_2, ω_3 are varied. Since the exponent clearly has no saddle point, the only regions which could lead the four-point function to grow exponentially are those in which $G_4(\omega_1, \omega_2, \omega_3, \omega_4)$ is singular. We thus hold ω_1 fixed, and scan over ω_2, ω_3, looking for regions in which the frequency-space four-point function is divergent. The relation between ω_2 and ω_3 where this occurs then determines the form of the exponent in (2.202), which can then be written just as a function of ω_2. This function may have saddles, which will either lead to an oscillatory exponent or a growing one.

There are two terms in $G_4(\omega_1, \omega_2, \omega_3, \omega_4)$ given by (2.189). Consider the first of these,

$$y\frac{z(1,4)G(1)G(2)G(3)G(4)}{1-yG(2)G(3)z(2,3)}2\pi\delta(\omega_1-\omega_3)\ , \tag{2.203}$$

where, as before, $G(j)$ denotes $G(\omega_j)$. It is convenient to rewrite (2.203) as

$$yG(2)G(3)\frac{1}{z(2,3)^{-1}G(2)^{-1}G(3)^{-1}-y}2\pi\delta(\omega_1-\omega_3)\ , \tag{2.204}$$

where from (2.181) we have that,

$$z(j,l)^{-1}G(j)^{-1}G(l)^{-1} = -\frac{1}{\kappa^2}(G(j)^{-1}+i\kappa)(G(l)^{-1}+i\kappa)\ . \tag{2.205}$$

It is convenient to rewrite the propagator (2.166) as,

$$G(j) = \frac{2i}{x_j}, \qquad x_j = \kappa(1-y)+\omega_j+\sqrt{\omega_j^2-2(1+y)\kappa\,\omega_j+\kappa^2(1-y)^2}\ . \tag{2.206}$$

Inverting the relation between ω_j and x_j,

$$\omega_j = \frac{x_j}{2}\left(1 + \frac{2\kappa y}{x_j - 2\kappa}\right) \ . \tag{2.207}$$

Notice that (2.207) has a symmetry; ω_j is invariant under,

$$x_j - 2\kappa \to \frac{4\kappa^2 y}{x_j - 2\kappa} \ . \tag{2.208}$$

This is analogous to the invariance seen in the IP model, see (2.161), as well as in the IOP model earlier, for $y = 1$. Now, the term (2.204) is singular when the denominator vanishes. Substituting (2.205, 2.206), this occurs at $-x_2x_3+2\kappa(x_2+x_3)-4\kappa^2(1-y)=0$, which is,

$$x_3 = 2\kappa\left(1 + \frac{2\kappa y}{x_2 - 2\kappa}\right) \ . \tag{2.209}$$

As a result of the invariance (2.208), this implies $\omega_2 = \omega_3$. This is the same as what was seen for the IOP model at $y = 1$, see (2.176). Thus, the exponent in (2.202), as a function of ω_2, is oscillatory, and the same holds at the location of its saddle.

Now consider the second term in $G_4(\omega_1, \omega_2, \omega_3, \omega_4)$, which is,

$$y(1-y)\frac{G(1)G(2)G(3)G(4)}{1 - yG(2)G(3)z(2,3)}\frac{z(1,4) - z(2,3)}{1 - yG(1)G(4)z(1,4)}\frac{i}{\omega_1 - \omega_3 + i\epsilon} \ . \tag{2.210}$$

It is convenient to rewrite (2.210) as,

$$y(1-y)\frac{z(2,3)^{-1}-z(1,4)^{-1}}{(z(2,3)^{-1}G(2)^{-1}G(3)^{-1}-y)(z(1,4)^{-1}G(1)^{-1}G(4)^{-1}-y)}\frac{i}{\omega_1-\omega_3+i\epsilon}\ . \tag{2.211}$$

We regard (2.211) as a function of ω_2, ω_3, where recall that $\omega_4 = \omega_1 + \omega_2 - \omega_3$. The nontrivial singularities in (2.211) arise from $(z(2,3)^{-1}G(2)^{-1}G(3)^{-1} - y) = 0$, which as shown in (2.209) implies $\omega_2 = \omega_3$, or from $(z(1,4)^{-1}G(1)^{-1}G(4)^{-1} - y) = 0$, which again gives $\omega_2 = \omega_3$. Thus, there is no regime of exponential growth for the four-point function.

2.4 Remarks on brane and antibrane dynamics

Brane systems in string theory are of great interest since they correspond to black holes, but their uses are far more general. Restricting just to their applications in phenomenology, stacks of branes naturally give rise to gauge groups [337] and allow for the construction of compactifications that could include the Standard Model [338, 339, 340], while providing the architecture of a lower-dimensional world inside a higher-dimensional compactification in brane world scenarios [341]. They are crucial in string constructions of inflating universes [342] with a positive cosmological constant [131] and can mediate transitions between different vacua with different cosmological constants [343], which can be used (at least schematically [344, 345]) to dynamically explain the small value that we observe.

Any phenomenologically viable string construction of our universe must at minimum exhibit its gross properties: four large spacetime dimensions, the absence of supersymmetry at low scales, and a positive cosmological constant, to name a few. While the first two are relatively well-understood [346, 347, 348], explicit models with a positive cosmological constant are difficult to construct. The essential issue is the absence of supersymmetry at any scale: the temperature inherent to expanding space breaks supersymmetry completely.

The lack of supersymmetry raises the specter of instability, since there is always the possibility of decay to a supersymmetric vacuum. While some degree of metastability is acceptable – over timescales comparable to the age of our universe, for example – more violent instability will predict a rapid decay to supersymmetry, conflicting with observation and ruining the model. One approach is to begin with a stable supersymmetric model and then break supersymmetry slightly; an exam-

ple is the KKLT construction [131], where a stable compactification is modified by the addition of branes with an orientation opposite to the background ("antibranes") that break supersymmetry and raise the cosmological constant above zero.

The stability of these antibrane models has been the subject of debate since their debut, centered over the past decade largely around the question of divergent fluxes. A number of works (see [336] for references) have pointed out that solutions to Einstein's equations in the background with antibranes have p-form electromagnetic fluxes that diverge near the antibrane. It was argued that these divergent fluxes could spur brane-flux annihilation, or push the antibranes into another part of the geometry where their effects would not be visible. However, below we argued that these divergences are merely artifacts of studying the broken supergravity theory outside of its regime of validity and that brane actions should be treated as in effective field theory, with the introduction of counterterms and matching onto a full string theory calculation necessary for the correct determination of the antibrane potential. We estimate the relevant diagrams and show that they are small at weak coupling, thus the addition of a few antibranes does not destabilize the KKLT construction.

2.4.1 Introduction

Brane actions are important for understanding many aspects of string physics. However, their precise interpretation is somewhat ambiguous. A brane is a source for the bulk fields, which are singular at the brane itself. If these fields are then inserted into the brane action, the result is divergent. Many applications use a

probe approximation, in which the self-fields of a brane are not included in the brane's action. This is like a formal limit in which the number of branes goes to zero.

A more general approach is to interpret the brane action in the context of effective field theory. Here, all effects are included, and divergences are treated via the usual framework of EFT [114]. For brane actions, this has been developed in Ref. [115], which shows that renormalization is the appropriate tool even for classical divergences such as those described above.[1] This can even lead to renormalization group flows of the type usually associated with quantum loops. In this paper we develop this point of view further, and show that it is useful in resolving some vexing issues in the literature.

In §2 we present a simple model that illustrates how the framework of Ref. [114] applies to branes. We discuss the matching onto the UV theory in various cases. In §3 we apply the EFT point of view to anti-D-branes in a flux background, focusing primarily on the case of a single antibrane.[2] We recover the phenomenon [118, 119] that in a flux background both branes and antibranes are screened by a background charge of the opposite sign. Divergences of the screening cloud near the brane are resolved by matching onto string theory at short distance and are not sources of instability. We show that possible nonperturbative annihilation of the antibrane and polarization cloud, while consistent with conservation of brane charge, is inconsistent with the H_3 Bianchi identity. Further, the apparent impossibility of black branes with antibrane charge [120, 121, 122] is avoided by

[1] Related earlier work includes Refs. [116].

[2] For a review of the extensive literature on the supergravity descriptions of antibranes in flux backgrounds and a complete list of references see [136].

proper account of a Bohm-Aharonov phase. The only allowed antibrane instability is the NS5-brane instanton of Ref. [123].

2.4.2 Effective brane actions

We illustrate the principle of effective brane actions with a simple model that captures the classical divergence problem noted above, and which gives a nice illustration of the general framework of Ref. [114]. In this model, the only bulk field is a free massless scalar φ in d spacetime dimensions. For now the brane is fixed on a $p+1$ dimensional subspace $x^{p+1} = \ldots = x^{d-1} = 0$, and it interacts with the bulk field via a general function of φ and its derivatives,

$$S = -\frac{1}{2}\int d^d x\, \partial_M \varphi \partial^M \varphi + \int d^{p+1} x_\parallel\, \mathcal{L}_{\rm brane}(\varphi, \partial)\,. \tag{2.212}$$

We will use M, N for all d dimensions, μ, ν for directions tangent to the brane, and m, n for directions orthogonal to the brane. For given d and p there will be only a finite number of renormalizable interactions, but in the spirit of effective field theory we keep all interactions, with nonrenormalizable interactions suppressed by the appropriate power of a large mass scale Λ. We are imagining that the brane is described in a UV complete theory such as string theory, in which these general interactions will be generated. If we are interested in amplitudes to some specified accuracy in $1/\Lambda$, then only a finite number of interactions contribute [114].

This point of view also requires that we keep general interactions in the bulk, but for simplicity we have omitted these. The form (2.212) is stable under renormalization. To make things even simpler, we restrict the brane action to terms

quadratic in φ, but with arbitrary derivatives. Again, this form is stable under renormalization.

To begin, we consider the simple interaction $\frac{1}{2}g\varphi^2$. To first order, Fig. 1a, the amplitude for $k_1 \to k_2$ scattering in the presence of the brane is

$$\mathcal{T}^{(1)} = g(2\pi)^{p+1}\delta^{p+1}(k_{1\parallel} - k_{2\parallel}) \equiv g\delta_\parallel \,. \tag{2.213}$$

Only momenta parallel to the brane are conserved, and we abbreviate the ubiquitous δ-function as indicated.

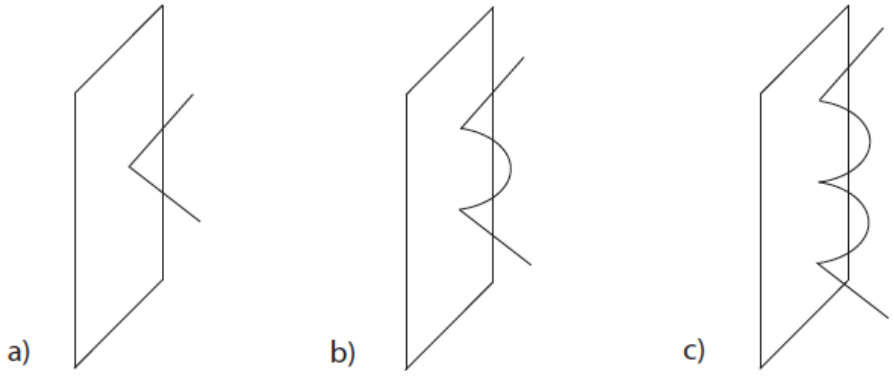


Figure 2.13: First, second, and third order terms in the amplitude for φ to scatter from the brane.

At second order, Fig. 1b, the amplitude is

$$\mathcal{T}^{(2)} = g^2\delta_\parallel \int \frac{d^r k_\perp}{(2\pi)^r} \frac{1}{k_\parallel^2 + k_\perp^2} \,. \tag{2.214}$$

Here $r = d - p - 1$ is the number of transverse dimensions. We see that this integral diverges for $r \geq 2$. To analyze this, we cut the integral off at $k_\perp = \Lambda$,

giving

$$\begin{aligned}\int^{\Lambda} \frac{d^r k_\perp}{(2\pi)^r} \frac{1}{k_\parallel^2 + k_\perp^2} &= (-1)^n \pi C_r k_\parallel^{r-2} + 2C_r \sum_{q=0}^{\infty} (-1)^q \frac{k_\parallel^{2q} \Lambda^{r-2-2q}}{r-2-q}\,, \quad r = 2n+1\,, \\ &= (-1)^n C_r k_\parallel^{r-2} \ln \frac{k_\parallel^2}{\Lambda^2} + 2C_r \sum_{\substack{q=0 \\ q \neq n-1}}^{\infty} (-1)^q \frac{k_\parallel^{2q} \Lambda^{r-2-2q}}{r-2-2q}\,, \quad r = 2n \end{aligned} \tag{2.215}$$

Here $C_r = V_{r-1}/2(2\pi)^r$ and V_{r-1} is the volume of the unit S^{r-1}. To analyze the divergences, let us note that the dimension of the interaction $\int d^{p+1}x\, \varphi^2$ is

$$\Delta = d - p - 3 = r - 2\,. \tag{2.216}$$

We include the volume element in the dimension, so negative Δ is relevant, vanishing Δ is marginal, and positive Δ is irrelevant (nonrenormalizable).

For codimension $r = 1$, the integral converges. Dropping for now terms suppressed by powers of Λ (we will return to them later), we have

$$\mathcal{T}^{(2)} = \frac{g^2}{2k_\parallel} \delta_\parallel\,. \tag{2.217}$$

This dominates the leading term (2.213) in the IR, as it should because the interaction is relevant. Further graphs form a geometric series, beginning with Fig. 1c, giving in all

$$\mathcal{T} = \frac{2gk_\parallel}{2k_\parallel - g} \delta_\parallel\,. \tag{2.218}$$

The interaction is attractive for positive g, consistent with the formation of a bound state.

For codimension $r = 2$, there is a log divergence,

$$\mathcal{T}^{(2)} = -\frac{g^2}{4\pi}\delta_\parallel \ln\frac{k_\parallel^2}{\Lambda^2}\,. \tag{2.219}$$

Again we can sum the geometric series,

$$\mathcal{T} = \frac{1}{\frac{1}{g} + \frac{1}{4\pi}\ln\frac{k_\parallel^2}{\Lambda^2}}\delta_\parallel\,. \tag{2.220}$$

The appearance of a logarithm is not surprising because for $r = 2$ the interaction is marginal. These logarithms and their RG interpretation were discussed in Ref. [115]. In conventional renormalization theory, we would take $\Lambda \to \infty$ holding fixed $g(\mu)^{-1} = g^{-1} + \frac{1}{4\pi}\ln\mu^2/\Lambda^2$. In effective field theory, Λ is a fixed UV scale. The divergence means that the effective field theory calculation is sensitive to UV physics, but only through local terms. We need to adjust g at this order, to account for the difference between our simple UV cutoff and the cutoff given by the true UV physics. We will discuss the matching onto the UV theory below. The logarithm means that the effective coupling $g(\mu)$ runs at scales below Λ. For positive g (attractive) there is again a pole in the IR, indicating a bound state. For negative g there is a Landau pole in the UV, but this is not a concern because this is only an effective theory.

For $r = 3$ the story is similar but the divergence is linear. The interaction is nonrenormalizable, so generically one would need more counterterms at higher loops, but in this simple model the higher loop graphs are just powers of the one loop graph and additional divergences do not appear.

For $r = 4$ we have

$$\mathcal{T}^{(2)} = g^2 C_4 \left(\Lambda^2 + k_\parallel^2 \ln \frac{k_\parallel^2}{\Lambda^2} \right) \delta_\parallel \,. \tag{2.221}$$

Now there are quadratic and logarithmic divergences, so the result depends on two parameters from the UV theory. The quadratic divergence requires adjustment of the original g to match onto the short distance theory. The log divergence requires a new interaction, $(\partial_\parallel \varphi)^2$. To make the power counting clearer we define a dimensionless coupling $\kappa_{00} = g\Lambda^{r-2}$, so that for $r = 4$,

$$\begin{aligned} \mathcal{T}^{(1)} &= \frac{\kappa_{00}}{\Lambda^2} \delta_\parallel \,, \\ \mathcal{T}^{(2)} &= \kappa_{00}^2 C_4 \left(\frac{1}{\Lambda^2} + \frac{k_\parallel^2}{\Lambda^4} \ln \frac{k_\parallel^2}{\Lambda^2} \right) \delta_\parallel \,. \end{aligned} \tag{2.222}$$

Because the interaction is irrelevant, $\Delta = 2$, even its leading effect is proportional to a negative power of Λ. The second order $k_\parallel$-independent term is of the same order in Λ. The $k_\parallel^2$ interaction comes with Λ^{-4}, as appropriate for a $\Delta = 4$ interaction. Its effect is suppressed relative to the $\Delta = 2$ term, but in the spirit of effective field theory we may be interested in $1/\Lambda$ corrections. Note that the first nonanalyticity in $k_\parallel^2$ comes in at order Λ^{-4}.

Note that all we are doing is solving the classical field equation

$$\partial_\parallel^2 \varphi + \partial_\perp^2 \varphi = -g\delta^r(x_\perp)\varphi \,, \tag{2.223}$$

but that this brings in the full machinery of EFT. Note also that with $i\partial_t$ replaced by $\partial_\parallel^2$ this is the same as the Schrodinger equation with a δ-function potential [124]

(this has also been noted by the authors of Ref. [115]). The bound states that we have found for $r = 1, 2$ are well-known as bound states in the Schrodinger problem. The case $r = 2$ is often used as a simple model of renormalization. Our discussion makes this connection more precise, and shows further that in general codimension this system also provides a simple model of EFT. (The cases $r = 2, 3$ are discussed in Ref. [125].)

In our toy example, where we have artificially restricted to interactions linear in φ^2, the most general brane action would be

$$S_{\rm brane} = \frac{1}{2}\int d^{p+1}x_\parallel \sum_{l,j=0}^{\infty}\sum_m \frac{\kappa_{lj}}{\Lambda^{2l+2j+r-2}} T^{jm}(\partial_\perp)\partial_{\mu_1}\ldots\partial_{\mu_l}\varphi\, T^{jm}(\partial_\perp)\partial^{\mu_1}\ldots\partial^{\mu_l}\varphi\,. \tag{2.224}$$

Here T^{jm} is a traceless polynomial of degree j, and m runs over these polynomials. In writing this we have used field redefinition to remove terms containing $\partial_\perp^2$, and have integrated by parts with respect to $\partial_\parallel$ but not $\partial_\perp$. To study amplitudes to accuracy Λ^{-s}, one would retain all terms with $\Delta \le s$.

We have omitted the brane's motion for simplicity, but this is readily included. A simple model, in which we do not try to keep the full d-dimensional Lorentz invariance, adds in a transverse collective coordinate $X^m(x_\parallel)$, beginning with the action

$$\begin{aligned} S &= -\frac{1}{2}\int d^dx\, \partial_M\varphi\partial^M\varphi + \int d^{p+1}x_\parallel \left(-\frac{\tau}{2}\partial_\mu X^m\partial^\mu X^m + \frac{g}{2}\varphi^2(x_\parallel, X_\perp(x_\parallel))\right) \\ &= -\frac{1}{2}\int d^dx\, \partial_M\varphi\partial^M\varphi \\ &\quad + \int d^{p+1}x_\parallel \left(-\frac{\tau}{2}\partial_\mu X^m\partial^\mu X^m + \frac{g}{2}\big[\varphi(x_\parallel,0) + X^m\partial_m\varphi(x_\parallel,0) + \ldots\big]^2\right). \end{aligned} \tag{2.225}$$

This now describes brane motion, and processes where scattering of scalars from the brane is accompanied by excitation of oscillations of the brane. Locality and translation invariance imply that undifferentiated X^m's appear only as arguments of φ. The same principles of renormalization apply. These principles apply further for brane and bulk actions with general fields, including all interactions allowed by symmetry.

Note that in $\mathcal{T}^{(2)}$ the brane is interacting with its own induced field, so this goes beyond the probe approximation. We have seen that this contribution is important for the leading IR physics for $r = 1, 2$. For larger r, it gives the leading nonanalytic behavior. However, in many situations only the leading behavior in $1/\Lambda$ is of interest. In particular, for branes of high codimension, the probe approximation $\mathcal{T}^{(1)}$ will be sufficient for most purposes. The point of this exercise is just to illustrate that brane actions can be sensibly interpreted in the framework of effective field theory.

Given a UV theory (we will consider some examples below), the couplings such as κ_{lj} are determined by calculating some process in both the UV theory and the effective theory with a given cutoff, and requiring that they agree.[3] After this is done, the effective theory can then be used for any other process. Note that different cutoffs will give different values for the Λ-dependent terms in integrals such as (2.215). This is compensated by different values for the couplings in the effective theory. It does not matter what cutoff we use as long as we are consistent, so in practice one often makes the simplest choice, dimensional regularization with

[3]The idea of matching is discussed in many reviews of effective field theory, e.g. [126].

minimal subtraction, for which

$$\begin{aligned}\int \frac{d^r k_\perp}{(2\pi)^r}\frac{1}{k_\parallel^2+k_\perp^2} &= (-1)^n \pi C_r k_\parallel^{r-2}\,, \quad r=2n+1\,,\\ &= (-1)^n C_r k_\parallel^{r-2}\ln\frac{k_\parallel^2}{\mu^2}\,, \quad r=2n\,. \end{aligned} \tag{2.226}$$

The absence of power law divergences does not mean that the corresponding couplings are not generated: we still need to compensate for the difference between the dimensional regulator and the true UV physics.

Once the effective action is determined, it can be applied to other situations such as a brane in a background field $\overline{\varphi}$ (we use a bar to denote the background). For example, the perturbation of the background by the brane is obtained from the same graphs as the S-matrix, in which one external state is replaced by $\overline{\varphi}$ and the other by a propagator, so the induced field is

$$\varphi_{\rm ind}(k) = \frac{1}{k^2}\int \frac{d^d k'}{(2\pi)^d}\mathcal{T}(k,k')\overline{\varphi}(k')\,. \tag{2.227}$$

(Using $\mathcal{T}$ here is a slight abuse of notation, because k' has been taken off-shell). For illustration, using the probe approximation $\mathcal{T}^{(1)}$ with the general action (2.224) gives in position space

$$\begin{aligned}\varphi_{\rm ind}(x) &= \int d^{p+1}x'_\parallel \sum_{l,j,m}\frac{\kappa_{lj}}{\Lambda^{2l+2j+r-2}}T^{jm}(\partial'_\perp)\partial'_{\mu_1}\ldots\partial'_{\mu_l}\overline{\varphi}(x')\times\\ &\qquad T^{jm}(\partial_\perp)\partial^{\mu_1}\ldots\partial^{\mu_l}\frac{1}{(d-2)V_{d-1}[(x_\parallel-x'_\parallel)^2+x_\perp^2]^{(d-2)/2}}\end{aligned} \tag{2.228}$$

This diverges at the brane, and the divergence grows with j and l. However, the

result applies only at momenta small compared to Λ and so at distances large compared to Λ^{-1}. Similarly, to study the motion of the brane in a background field, one can use the probe action or, if greater accuracy is needed, add in the higher corrections — essentially $\mathcal{T}$ again, but with both external states replaced by $\overline{\varphi}$ (we will see an example of this in Fig. 2b).

For a D-brane, the UV theory is string theory. The amplitude $\mathcal{T}^{(1)}$ is the effective description of the disk with two closed string vertex operators. By calculating this disk amplitude, and requiring that the effective field theory give the same answer, one determines the brane couplings with any number of derivatives (the equivalent to the κ_{jl}) to leading order in $g_{\rm s}$.[4] The amplitude $\mathcal{T}^{(2)}$ is the effective description of the annulus with two closed string vertex operators, and so one would need to match this amplitude to determine the effective action to order $g_{\rm s}^2$.

Another situation would be a solitonic brane, such as a magnetic monopole, vortex, or domain wall in a spontaneously broken QFT. The UV theory would be the unbroken QFT and the effective theory would describe the brane collective coordinates plus any light fields. Again one matches a UV calculation to one in the effective description. In the UV calculation, the key input is the requirement that the fields of the soliton be nonsingular.

One might try to apply the second method to the D-brane, using its supergravity description together with a condition such as [128] on allowable singularities. However, the scale of the supergravity solution for a single D-brane is smaller than the string length by a power of $g_{\rm s}$, so this is not a good description. (It is

[4] A partial list of papers on the disk effective action is given in Ref. [127].

a valid description if enough D-branes coincide, a point we will return to below.) If the supergravity approach did give an answer, it would likely not agree with the correct string theory result, because string theory knows about the scale α' and supergravity does not. Similarly, if the supergravity approach fails to give an answer due to singularities deemed bad, this has no physical significance. It is the matching onto string perturbation theory that is the correct criterion for a good singularity in the fields external to a D-brane.[5]

For sufficiently supersymmetric amplitudes, the supergravity calculation will agree with the string calculation, because of the absence of α' corrections. This does not mean that supergravity is an accurate description of a single D-brane. The magic of supersymmetry sometimes leads to complacency about the validity of effective descriptions. For example, it has sometimes led to weak/strong dualities being misunderstood as weak/weak dualities.

2.4.3 Antibranes in fluxes

Application of EFT

De Sitter vacua of string theory may be numerous but they are not simple. (Meta)stability requires the balance of several forms of energy density [130]. The KKLT construction [131] begins with a supersymmetric anti-de Sitter vacuum and excites it by adding one or more antibranes (branes having opposite supersymmetry to the background). The nature of this supersymmetry breaking has recently been understood in Ref. [132]. A body of work beginning with Refs. [133, 134]

[5]For M-branes there is no perturbative description of the UV theory, but Matrix theory [129] provides a construction of the S-matrix to which the effective theory should be matched.

has argued that the dynamics of anti-D-branes is complicated and potentially unstable.

In the KKLT model [131], a single antibrane can be sufficient to uplift an AdS vacuum to a dS vacuum, and this is the case that we focus on here. The scale of the geometry is large compared to the string length, so EFT should be valid. In the effective description the only low energy brane degrees of freedom are the gauge fields in the Poincaré directions and the collective coordinate for the brane motion. The only thing the antibrane can do to lower its energy is to move to the position of lowest potential, the bottom of the Klebanov-Strassler (KS) throat.[6]

To illustrate the use of EFT, consider a potentially problematic issue, the backreaction on H_3. A low-order contribution is shown in Fig. 2a. A bulk po-

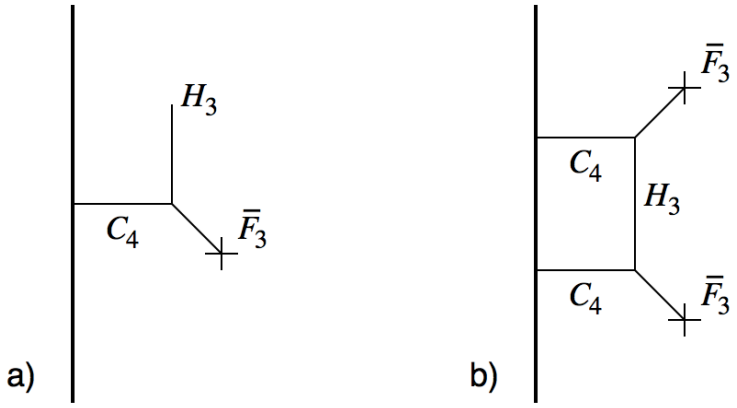


Figure 2.14: a) Lowest order backreaction on H_3. The heavy line is the anti-D3 brane, and the $\times$ denotes a background field. b) Corresponding contribution to the brane potential.

tential has scaling dimension 4. Its *engineering* dimension is 0 since we include an α'^{-4} in its kinetic term, but this is not what matters for degree of divergence;

[6]Of course, if there are massless or light moduli in the vacuum without an anti-branes, adding the antibrane could destabilize them. See for example Ref. [135]. This would be seen in the effective field theory.

henceforth 'dimension' refers to scaling unless otherwise specified. The interaction $\alpha'^{-2}\int d^4x\, C_4$ has dimension $\Delta = 4-4 = 0$. The Chern-Simons interaction $\alpha'^{-4}\int d^{10}x\, C_4\wedge F_3\wedge H_3$ has dimension $\Delta = 4+5+5-10 = 4$. The total dimension of all interactions in Fig. 2a is then 4, and it follows that

$$H_3 \propto \frac{g_s^2\alpha'^2\overline{F}_3}{x_\perp^4}\,, \tag{2.229}$$

where a bar again denotes the background field. The $x_\perp^{-4}$ is from the scaling dimension, and the α'^2 has been inserted by engineering dimensional analysis. We work in the string metric, so the g_s^2 is from the H_3 propagator.[7] There is also a contribution

$$H_3 \propto \frac{g_s\alpha'^2\overline{H}_3}{x_\perp^4}\,, \tag{2.230}$$

from a similar graph with $g_{\mu\nu}$ in place of F_5 and $\overline{H}_3$ in place of $\overline{F}_3$. Even at the limit of EFT, $x_\perp \sim \alpha'^{1/2}$, this is a small perturbation on the background at weak coupling.

The integral of the energy density $g_s^{-2}H_3^2$ diverges quadratically at the brane. The corresponding graph is Fig. 2b. The total Δ of the interactions is 8, and the leading brane counterterm $\int d^4x\, F_3^2$ has dimension 6, so the divergence again comes out quadratic. The counterterm is of order

$$g_s^2\alpha'^{-1}\int d^4x\,\sqrt{-g_4}F_3^2\,. \tag{2.231}$$

[7] The metric and B_2 have a g_s^2 in the propagator; while the RR forms do not depend on g_s. The bulk gravitational interaction contains a g_s^{-2}, while the Chern-Simons terms do not depend on g_s. The coupling of the metric to the brane is proportional to g_s^{-1}, while the coupling of the RR form to the brane does not depend on g_s.

The factors of α' from the vertices and propagators cancel, leaving an α'^{-1} from the cutoff; the net result is fixed anyway by the (engineering) dimensions. Again, the numerical coefficient would come from matching to the string annulus graph. In a similar way we get a counterterm

$$\alpha'^{-1} \int d^4x \sqrt{-g_4} H_3^2 \,. \tag{2.232}$$

The counterterms (2.231, 2.232) are each one order in g_s higher than terms that are expected in the tree level brane action,

$$g_{\rm s}\alpha'^{-1} \int d^4x \sqrt{-g_4} F_3^2 \,, \quad g_{\rm s}^{-1}\alpha'^{-1} \int d^4x \sqrt{-g_4} H_3^2 \,, \tag{2.233}$$

as expected for the annulus in comparison to the disk.

Expanding around a minimum of the potential, one gets a mass correction of order

$$\alpha'^{-1} \left(g_{\rm s}^2 (\partial \overline{F}_3)^2 + (\partial \overline{H}_3)^2 \right) X^2 \tag{2.234}$$

from the annulus corrections (2.231, 2.232). Note that this is a dimensional estimate; the signs and tensor structures are not specified. In particular, the potential will vanish along Goldstone directions such at the S^3 of the Klebanov-Strassler throat. For comparison, the leading order potential is $\alpha'^{-2} g_{\rm s}^{-1} \int d^4x \sqrt{-g_4}$. We can estimate the second derivative of this from Einstein's equation, giving a mass term of order

$$\alpha'^{-2} \left(g_{\rm s} \overline{F}_3^2 + g_{\rm s}^{-1} \overline{H}_3^2 \right) X^2 \,. \tag{2.235}$$

The mass correction (2.234) is suppressed by $g_{\rm s}$ and also by α'/L^2, where L is the

characteristic scale of the geometry. The effect of the higher derivative tree-level terms (2.233) is suppressed by α'/L^2 but not by $g_{\rm s}$.

In summary, self-consistent use of effective field theory shows no large corrections that would signal a breakdown. Again, the antibrane's only degree of freedom is its position. Energetically this is limited to a bounded space, the neighborhood of the bottom of the Klebanov-Strassler throat, and so there must be a minimum, where all perturbations have nonnegative mass-squared.

More on antibrane dynamics

When an antibrane and brane are close together, there is an open string tachyon between them that leads to their annihilation. However, when the brane dissolves into flux, its world-volume gauge field is in a confining phase, and strings cannot terminate in flux, at least perturbatively. There are no degrees of freedom within the EFT that would describe such an annihilation. But the EFT does describe the dynamics of the fluxes, and a closer look at these is warranted.

The antibrane can decay via an NS5-brane instanton [123], which mediates the process

$$\overline{\rm D3} + M \text{ units of } H_3 \wedge F_3 \;\to\; M - 1 \text{ D3}'\text{s}\,. \tag{2.236}$$

This is a nonperturbative effect. The backreaction in effective field theory does not significantly affect the instanton action: the amplitude of Fig. 1c, for example, is further reduced by the dissolving of the $\overline{\rm D3}$ in the NS5. In particular, the flux-clumping Ansatz of Refs. [119, 121, 136] does not seem to apply.

In the NS5 process [123], the initial configuration is a stack of anti-D3-branes polarized into an NS5-brane that subtends an angle $\psi = \psi_i$. The decay process

involves ψ tunneling through a potential barrier to a lower-energy final state. For a single antibrane, the initial ψ_i would be so small that the description breaks down: the initial polarization is negligible. However, the decay process for a single antibrane still requires an NS5-brane instanton in order to source the H_3 Bianchi identity, so ψ must pass through large values where the polarization picture applies, and the dominant contribution to the tunneling action comes from this region. So the KPV result still applies for the single antibrane.

In Ref. [120, 121], it is shown that there is no black antibrane solution with $\overline{\text{D6}}$ charge immersed in a background of the opposite sign.[8] This suggests that finite temperature eliminates the barrier to brane-flux annihilation so that it is rapid, rather than proceeding via tunneling. However, even if this is true, it does not provide any evidence for rapid decay at zero temperature. It is quite possible for a process to be nonperturbatively slow at low temperature and rapid at high temperature. Electroweak baryon number violation is an important example.

In an earlier version of this work we suggested a more rapid, but still nonperturbative, decay. In fact, this does not exist.[9] The remainder of this subsection deals with this. We will focus on the anti-D6 case, which has been worked out in greatest detail [137, 121]. The key field equations are

$$d(*_{10}e^{-2\varphi}H_3) \;=\; -F_0 *_{10} F_2\,, \tag{2.237}$$

[8] At zero temperature, if the antibrane δ-function in Fig. 3b is smoothed, the density has a volcano shape with a maximum at the rim. Such a maximum at positive polarization density is forbidden [137] by Eqs. (2.237-2.239), but this is at the string length and so outside the validity of effective field theory; there is no problem with the distribution in Fig. 3b in string perturbation theory. A similar argument is used in the black case, but only outside the horizon where it is valid if the Schwarzschild radius is greater than the string length.

[9] We thank Eva Silverstein and Juan Maldacena for pointing out our error.

$$dF_2 = F_0 H_3 + \delta_{\rm D6}\,, \tag{2.238}$$

$$dF_0 = 0 \tag{2.239}$$

$$dH_3 = \delta_{\rm NS5}\,. \tag{2.240}$$

We have translated to string frame for consistency with our earlier discussion. For completeness every equation should include potential brane sources, but the F1 sources in (2.237) and the D8 sources in (2.239) will play no role; thus the zero form F_0 is constant. The $\delta_{\rm D6}$ is summed with sign over brane and antibrane sources, and if the space transverse to the D6-branes is compact, it should also include negative contributions from O6 planes.

Expanding around this background we have in particular

$$d\delta F_2 = F_0 \delta H_3 + \delta\delta_{\rm D6}\,. \tag{2.241}$$

The brane induces a δF_2 via Eq. (2.238), and Eq. (2.237) then leads to a δH_3. On the RHS of Eq. (2.241) this provides a perturbation to the background D6 density due to polarization of the flux background. Eqs. (2.237, 2.238) together imply a mass-squared term of order $e^{2\overline{\varphi}}\overline{F}_0^2 \equiv \mu^2$ for the perturbations; essentially F_2 Higges H_3. The perturbations thus fall exponentially away from the brane. Integrating Eq. (2.241) over the transverse space, the LHS must then vanish, and therefore the RHS does as well: the polarization of the background screens that of the D6 or anti-D6 completely [118, 119], Fig. 3.

The total background charge contained within the volume of the screening

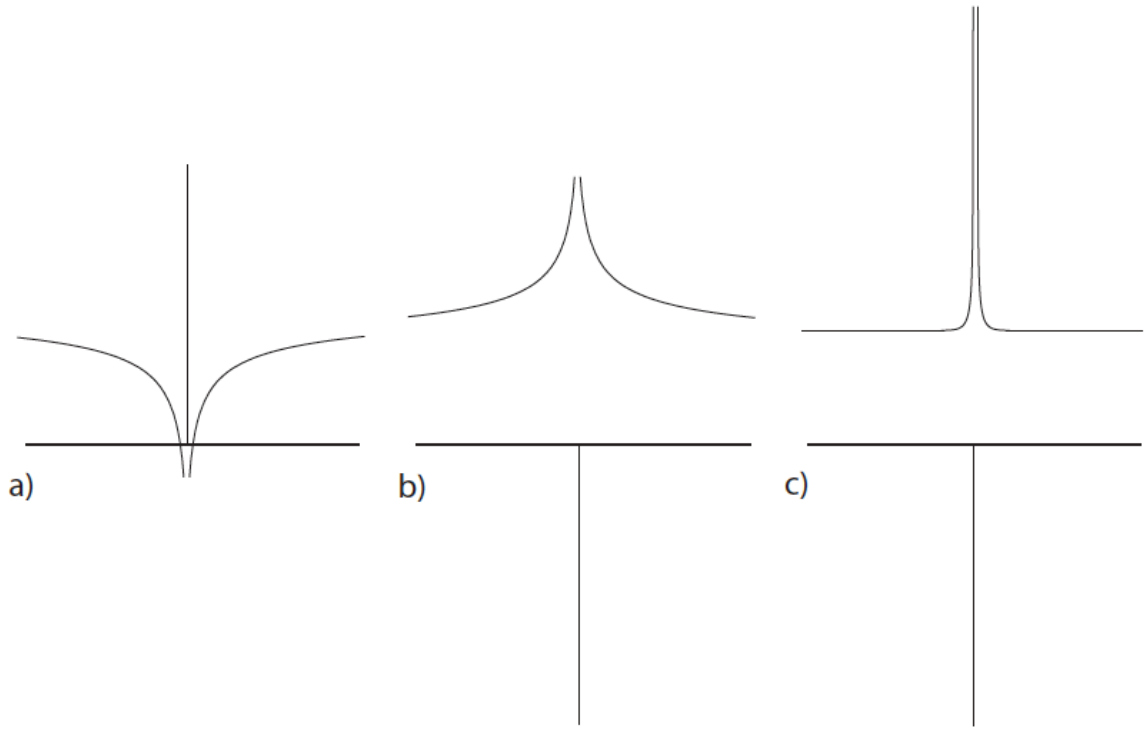


Figure 2.15: D6 densities in a flux background. In all cases the excess or deficit in the screening cloud offsets that due to the brane source. a) D6-brane in a flux background. b) Anti-D6-brane in a flux background. c) Fluctuation of the anti-D6-brane's screening cloud down to a size of order the string length.

cloud is of order

$$\overline{H}_3\overline{F}_0/\mu^3 \sim e^{\bar{\varphi}}\overline{F}_0^2/\mu^3 \sim 1/e^{2\bar{\varphi}}\overline{F}_0\,. \tag{2.242}$$

This is large if the flux is dilute and/or the coupling is weak, so we can treat the screening due to a single D6 as a perturbation. The screening cloud diverges as we approach the brane, and due to the nonlinearities of the field equation the expansion of the field near the origin will contain all negative powers of the distance. However, as in the toy model example, this is not a problem: the brane effective action gives a precise prescription for matching the fields external to the brane onto the UV physics at the string length. Further, we have seen from

Eqs. (2.229, 3.210) that even very close to the brane the screening charge density is small, so cannot drive any open string model tachyon.

This discussion suggests that the antibrane can annihilate with its polarization cloud, $\overline{\rm D3}+1$ units of $H_3\wedge F_3\to$ energy. This was suggested in Refs. [119, 136] as a means of enhancing the NS5 decay; we have argued above that any such effect is slight. The process considered in the earlier version of the present work was a local annihilation, $\overline{\rm D3}+1$ unit of $H_3\wedge F_3\to$ energy. We will see that this is forbidden by the H_3 Bianchi identity.[10]

One might have thought that something interesting could happen nonperturbatively. Consider a fluctuation of the supergravity fields like that shown in Fig. 3c, where the screening charge concentrates into a very small volume. Is it possible for the brane and the flux to mutually annihilate? This would conserve $D6$ charge, but we must also consider the H_3 Bianchi identity, which we can think of as a conservation law for the current $(*H)_7$. The unit D6 charge of the polarization cloud implies $\int_{\rm cloud} H_3 = 1/F_0$, or

$$H_3 \sim \delta_{\overline{D6}}/F_0 \tag{2.243}$$

before the annihilation. After the annihilation the cloud is gone, so there is a source

$$dH_3 \sim \delta(t)\delta_{\overline{\rm D6}}/F_0\,. \tag{2.244}$$

This is geometrically consistent with an $NS5$-brane instanton, but that would

[10] In the earlier version, the brane-flux annihilation was linked to the breaking of the heterotic string [138]. That process is consistent with the heterotic string Bianchi and quantization conditions, as shown explicitly there by a K theory construction.

give $dH_3 \sim \delta(t)\delta_{\overline{\rm D6}}$. The parametric dependence on F_0, which is an arbitrary integer in the natural units that we are using here, allows us to distinguish the known NS5-brane instanton process from the new brane-flux annihilation process suggested in version one. We see that the latter is forbidden.

It is interesting to compare the zero-temperature and high-temperature behaviors. In the absence of NS5-brane sources $H_3 = dB_2$. Integrating the F_2 Bianchi identity (2.238) on an S^2 just outside the black brane horizon, we get

$$\frac{\partial}{\partial t}\int_{S^2}(F_2 - F_0 B_2) = 0\,. \tag{2.245}$$

We omit the source term because we will consider a process during which no branes cross the S^2. The conserved quantity (2.245) was termed a Page charge in Ref. [139]. As noted there, it is localized, quantized, conserved, but *not* invariant under large gauge transformations. In particular it jumps by F_0 under $\int B_2 \to 1 + \int B_2$. Thus it is a Z_{F_0} charge, whose conservation excludes the brane-flux annihilation for a single antibrane.

Imagine starting with an antibrane at zero temperature. In the integral (2.245), the flux from the antibrane contributes -1. As the black hole forms, the argument of [120, 121] implies that it must absorb the polarization cloud in order that $\int_{S^2} F_2$ becomes positive. However, $\int_{S^2}(F_2 - F_0 B_2)$ remains negative and keeps track of the antibrane number. If we cool the system back to zero temperature, the antibrane must reappear. The integral of B_2 over the horizon is a sort of hair that can be measured in a stringy Bohm-Aharonov experiment [140]. Again, their might be a process in which an NS5-brane instanton changes the Page charge by

F_0 units, but it cannot change by a single unit. Finite temperature would be expected to reduce the barrier for the NS5-brane instanton, and might eliminate it entirely at high enough temperature.[11]

More globally, imagine an S^3 whose equator S^2 surrounds the black 6-brane, and which is elongated in time to incorporate the black brane formation and disappearance. Let the S_3 surround $N_{\rm NS5}$ NS5 instantons. Then

$$\begin{aligned} \int_{S^3} H_3 &= N_{\rm NS5}\,, \\ 0 = \int_{S^3} dF_2 &= F_0 N_{\rm NS5} + \Delta N_{\rm D6}\,. \end{aligned} \tag{2.246}$$

Thus the net $D6$ charge can only change in multiples of F_0. The integer quantization of $\int_{S^3} H_3$ follows from the Dirac quantization condition [141], independent of the dynamics internal to the S^3.

These considerations extend to other antibranes.[12] For the KKLT anti-D3, consider an $S^3_a \times S^3_b$, where S^3_a is parallel to the bottom of the KS throat and S^3_b surrounds the antibrane in the directions transverse to S^3_a, and in time. Then

$$0 = \int_{S^3_a \times S^3_b} dF_5 = \int_{S^3_a \times S^3_b} (F_3 \wedge H_3 + \delta_{\rm D3}) = M N_{\rm NS5} + \Delta N_{\rm D3}\,. \tag{2.247}$$

Here $N_{\rm D3}$ includes any D3 charge arising from flux on wrapped 7-branes. The M is the number of units of F_3 flux on S^3_a. It plays the same role as F_0 above, distinguishing the known NS5 process [123] in which $\Delta N_{\rm D3} = -M$ from a potentially new brane-flux annihilation process with $\Delta N_{\rm D3} = -1$. The latter is forbidden.

[11] We thank Don Marolf for discussions of this point.

[12] Gavin Hartnett has given a complementary argument for localized $\overline{\rm D3}$'s, that there is no positivity condition on the black brane flux in this case [142] .

Thus the metastability estimates in the original work [131] appear to be correct.[13]

Multiple antibranes

For p coincident D-branes, the effective field theory on the brane becomes non-Abelian (the notation p is standard here, not to be confused with the dimension p of the branes). When $g_s p \gg 1$, the brane theory is strongly coupled but the supergravity description is good. For (anti-)D3-branes, the geometry near the branes is described by an $AdS_5 \times S^5$ throat at the bottom of the KS throat [336]. When the background is slowly varying on the scale of the brane radius $(g_s p)^{1/4}\alpha'^{1/2}$ (meaning that p is parametrically smaller than M in the KKLT context), one can again use an effective brane description of the system as seen from the outside. In the UV, this is matched onto the supergravity description of the throat. Modes in the throat behave as $z^{\lambda_\pm}$. Most modes correspond to irrelevant interactions, where $\lambda_- = -\Delta$ is negative and $\lambda_+ = \Delta + 4$ is positive (for consistency we continue to use the somewhat nonstandard convention that Δ includes -4 from the integration). The λ_- mode goes to zero at the bottom of the throat while the λ_+ mode diverges, and we get a good boundary condition by requiring that the latter vanish. Integrating through the transition between the throat and the exterior to determine the exterior fields, and matching to an effective field theory calculation analogous to (2.228), determines the parameters in the effective action.

However, for modes corresponding to relevant interactions, both λ_+ and λ_- are positive and both modes grow down the AdS_5 throat. In this case one must

[13] The very recent work [143] has argued that the NS5 decay might be hastened by by passing through a new set of low energy configurations. The proposed configurations violate the Bianchi identity for the NS5 world-volume gauge field, and so are forbidden.

understand the nonlinearities there. The $\Delta = -1$ modes corresponding to fermion bilinears involve the 3-form fluxes. For these the singularity is resolved by brane polarization [36], giving a good UV description. The $\Delta = -2$ modes corresponding to scalar bilinears involve the five-form flux and scalar deformation of the S^5. Resolution of the resulting singularities requires the branes to move out onto the Coulomb branch [145], and because there is no $L = 0$, $\Delta = -2$, scalar bilinear the potential is always negative in some directions and the (anti-)D3 branes are expelled by the AdS throat. In either case, once the actual physics in the throat is understood, one can determine the effective field theory.

When both the $\Delta = -1$ and $\Delta = -2$ perturbations are present, there is a competition between these two effects. When the anti-D3-branes and their AdS throat are at the bottom of a KS throat, this is the case. This has recently been studied in Ref. [336]. They concluded that if a parameter $\mathrm{Im}(\mu)$ is nonzero, then it is energetically favored for the branes to be expelled from the AdS throat. They are expelled in an oblique direction (the so-called 'giant tachyon' of [336]), so they are not precisely at the bottom of the KS throat, but energetically they cannot wander too far from the bottom. The screening effect implies that antibranes attract at longer distances [118], so their precise arrangement may be intricate, but in any case our earlier discussion of the single antibrane now applies. If the parameter $\mathrm{Im}(\mu)$ vanishes (as may be required by symmetry), then the branes do polarize as in Ref. [123] if p is not too large. Again, this will be subject to nonperturbative decay via the NS5 instanton [123].

2.4.4 Conclusions

We have argued that effective field theory allows the use of brane actions beyond the probe approximation, including the treatment of both classical and quantum divergences. In all applications of brane systems, this provides a more general and physical interpretation of the results. In applying this to the antibrane in flux, this validates the assumptions of Ref. [131]: the supersymmetry-breaking antibranes can be described by effective field theory, and are metastable if their number p is not too large. It also follows that large classes of non-extremal fuzzball solutions (geometries or stringy solutions) using antibranes should exist [146].

Acknowledgments

The work of B.M. is supported by the NSF Graduate Research Fellowship Grant DGE-1144085. The work of E.M. is supported by NSF grant PHY13-16748. The work of J.P. is supported by NSF Grant PHY11-25915 (academic year) and PHY13-16748 (summer). The work of A.P. is supported by National Science Foundation Grant No. PHY12-05500. The work of P.S. was supported by a Worster Fellowship. We thank Iosif Bena, Ulf Danielsson, Mariana Graña, Gavin Hartnett, Ehson Hatefi, Gary Horowitz, Stefano Massai, Don Marolf, and Thomas Van Riet for discussions and communications. We thank Juan Maldacena and Eva Silverstein for crucial questions.

2.5 Universality of sparse $d > 2$ conformal field theory at large N

In this section we study a slightly more general question: the implications of gravitational physics for a dual conformal field theory. Such holographically dual pairs are typically realized by a brane construction, but many questions of interest (such as thermodynamics) can be studied without reference to the underlying microscopy. The implications of a gravitational holographic dual are strong: the CFT must have a very particular phase structure [335], and if the gravitational theory furthermore looks like Einstein gravity there must be a sense of bulk locality encoded in the field theory correlators [162, 349].

This section focuses on the coarser question of the reproduction of the gravitational phase structure on the conformal field theory side, originally studied in 1+1D by [156]. In that setting, AdS gravity implies that the partition function below the transition is dominated by the vacuum contribution, leading an explicit bound on the density of light states of the dual conformal field theory (the heavy states are controlled by modular invariance). Gravity further implies that the vacuum contribution is universal, in the sense that the vacuum energy density is unmodified by finite-size corrections, but since this quantity is fixed in CFT_2 by conformal invariance the bound on the light states is all that is needed for CFT_2 to reproduce the phase structure implied by gravity.

We studied the implications of the gravitational phase structure in higher dimensions. The gravitational phase structure is not known in a general background, but as in 1+1D the analysis simplifies in the presence of additional symmetry. If

one considers a higher-dimensional CFT on a spatial torus there is the following structure [187]: when the length β of the thermal circle is smaller than the smallest sidelength $L_{\min}$ of the spatial torus, the dominant gravitational saddle is a black brane, while when $\beta > L_{\min}$ the dominant saddle is an AdS-soliton geometry with the soliton wrapping the smallest spatial cycle.

On the field theory side, the cycles of the torus permute with the time circle (and amongst themselves) much as in 1+1D, leading in conformal theories to relations between partition functions at different temperatures [175, 176] that allow the derivation of a set of necessary and sufficient constraints on the CFT data. These constraints are stronger than in 1+1D since the vacuum energy is not fixed by conformal invariance: in higher dimensions one must demand vacuum dominance in the quantization along every cycle except the smallest, which then implies the universal gravitational result that the free energy of the theory is uncorrected relative to its asymptotic limit on a torus of infinite size. One can equivalently obtain universality by simply demanding that the free energy is uncorrected and that the density of light states satisfies a similar bound, but if one does not restrict the free energy in this way, the bound on the density of states must be imposed on the entire spectrum. We further showed that these constraints on the density of states are mildly sub-saturated by free orbifold theories, much like in CFT_2 where such theories saturate the universality constraint exactly.

2.5.1 Introduction

The strongest form of the AdS/CFT correspondence states that every conformal field theory (CFT_d) is dual to a theory of quantum gravity living in a

higher-dimensional anti-de Sitter space (AdS_{d+1}). For a generic CFT, the dual theory of quantum gravity at low energies will look nothing like semi-classical Einstein gravity. One of the most interesting questions in the context of holography is then to understand which CFTs – when interpreted as theories of quantum gravity in AdS – have a semi-classical Einstein gravity limit.

The most straightforward constraint emerging from the AdS/CFT dictionary for a semi-classical bulk is that the CFT should have a large number of degrees of freedom, usually parameterized by N. Large N in the field theory implies a semi-classical bulk since its inverse scales as a positive power of the Planck length in AdS units: $N^{-1} \sim (\ell_P/\ell_{\mathrm{AdS}})^{\#}$ for $\# > 0$. This is the bulk expansion parameter controlling AdS-scale quantum gravitational effects.[14]

Besides large N, a semi-classical theory of gravity in anti-de Sitter space has many universal features that must be encoded in any putative dual CFT. To explore the emergence of gravity from field-theoretic degrees of freedom, it is natural to try to reproduce these universal features by implementing some additional assumptions on a generic large-N CFT. There has been tremendous progress in this direction for the case of three-dimensional gravity [147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 358, 158, 159, 160], throughout which large central charge and a sparse low-energy spectrum play a prominent role. These powerful methods for the most part rely on the fact that all stress tensor interactions in the CFT are captured by the Virasoro block of the identity, which is assumed to domi-

[14]To have a theory that looks like Einstein gravity at low energies, we also need an expansion parameter that can suppress higher-spin fields. The 't Hooft coupling in gauge theory usually plays the role of this expansion parameter. Interestingly, like in the D1-D5 duality, certain features of Einstein gravity can be reproduced without explicitly invoking this assumption. We will not explicitly implement any constraints on our field theories with the purpose of decoupling higher-spin fields.

nate. The success of this particular approach is related to the topological nature of gravity in three dimensions, which precludes obvious generalizations to higher dimensions. Nevertheless, it is a compelling problem to reproduce features of higher-dimensional AdS gravity purely from the CFT. A small sample of work in this direction includes [161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 349, 172].

In this paper, we will focus on a technical tool that has received little exposure in higher dimensions: modular invariance. For 2d CFTs, modular invariance can be used to precisely determine how sparse the spectrum should be to reproduce the thermal phase structure of 3d gravity [156] (see [173] for a similar consideration in supersymmetric theories). For theories obeying this sparseness constraint, the Cardy formula [174] – which is usually only valid asymptotically as $\Delta/c \to \infty$ – has an extended regime of validity down to energies $\Delta \sim c$. This precisely matches the bulk phase structure since the black holes begin dominating the ensemble at $\Delta \sim c$.

The relevance of modular invariance in higher-dimensional holographic CFTs has been much less explored. In [175, 176], it was shown that modular invariance of the torus partition function implies the existence of an asymptotic formula that correctly reproduces the Bekenstein-Hawking entropy of the dual black brane. This formula is the higher-dimensional generalization of the Cardy formula and only holds in the limit of large energy for generic CFTs. Holographic CFTs, on the other hand, must have an extended range of validity of this formula as implied by the bulk phase structure. The goal of this paper is to further exploit modular invariance and place constraints on CFTs such that they have this extended range of validity. We also want to match the precise phase structure of gravity, which

is much richer than in two dimensions and exhibits both quantum and thermal phase transitions. One of the key challenges that we will face is that the functional form of the vacuum energy in higher dimensions is not uniquely fixed by conformal invariance, although we will discover several nontrivial constraints due to modular invariance.

We can summarize our results as follows. A general CFT on $\mathbb{T}^d$ will have an extended Cardy formula and a universal phase structure if and only if the partition function is dominated by the vacuum contribution when quantizing along any cycle but the shortest one. Proving this will require using the modular constraints on the vacuum energy alluded to above. From here, we will consider large-N theories and exhibit distinct sets of necessary and sufficient sparseness conditions on the spectrum to achieve this vacuum domination.

In analyzing calculable theories that satisfy these necessary and sufficient conditions, and which therefore have a universal free energy, we are led to the construction of symmetric orbifold theories in higher dimensions. Symmetric orbifolds have been analyzed in great depth in two dimensions [177, 178, 179, 180, 181, 182, 183], and play an explicit role in the D1-D5 duality [184, 58, 59]. Still, they have not explicitly appeared in holographic dualities in higher dimensions nor, to the best of our knowledge, have they been constructed. For their construction, we use a similar procedure as in two dimensions to build a modular invariant partition function. This includes both untwisted and twisted sectors. For large-N symmetric product orbifolds, the density of states of the untwisted sector is shown to be slightly sub-Hagedorn, whereas for the twisted sector it is precisely Hagedorn. Saturation of the necessary and sufficient conditions for universality is then

guaranteed by assuming that the subextensive parts of the vacuum energy vanish. This assumption constrains the choice of seed theory we can pick. This is somewhat of a loss of generality compared to two dimensions but can be expected by the increasing richness of CFTs in higher dimensions. Provided we pick the seed accordingly, the symmetric orbifolds reproduce the phase structure of higher-dimensional AdS gravity: they have an extended regime of validity of the Cardy formula and a Hagedorn transition at precisely the same temperature as the Hawking-Page transition in the bulk.

The paper is organized as follows. We start in section 2.5.2 with a general discussion of CFTs on d-dimensional tori and modular invariance. In section 2.5.3 we summarize the phase structure of toroidally compactified gravity in anti-de Sitter spacetime. These two sections set the stage for the meat of the paper. Section 2.5.4 is dedicated to a detailed discussion of the necessary and sufficient conditions that are required to have a universal free energy. The implementation of these conditions is then explored in section 2.5.5. We discuss the construction of orbifold theories on d-dimensional tori and show that symmetric product orbifolds have a universal free energy. We conclude with a discussion and outlook in section 2.5.6. The appendices contain additional material, including extensions to the case with angular momentum and calculations translating the results from the canonical partition function to the microcanonical density of states.

2.5.2 Generalities of CFT$_d$

We now introduce some of the basic technology of modular invariance that we will use to derive our general CFT results. For more details see [175, 176]. In this paper we will study conformal field theories defined on a Euclidean d-torus $\mathbb{T}^d$. We fix the metric on this torus to be

$$ds^2 = dx_0^2 + dx_1^2 + \cdots + dx_{d-1}^2 \tag{2.248}$$

with identifications

$$(x_0, x_1, .., x_{d-1}) \sim (x_0, x_1, .., x_{d-1}) + \sum_{i=0}^{d-1} n_i U_i \,. \tag{2.249}$$

where U_i are vectors defining the torus $\mathbb{T}^d$ and the n_i are integers. These vectors can be conveniently organized in a matrix as

$$U = (U_0 \quad \cdots \quad U_{d-1})^T = \begin{pmatrix} L_0 & \theta_{01} & \cdots & \theta_{0,(d-2)} & \theta_{0,(d-1)} \\ 0 & L_1 & \cdots & \theta_{1,(d-2)} & \theta_{1,(d-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & L_{d-2} & \theta_{(d-2),(d-1)} \\ 0 & 0 & \cdots & 0 & L_{d-1} \end{pmatrix} \tag{2.250}$$

and define a d-dimensional lattice of identifications. This matrix contains the lengths of the cycles along its diagonal and the θ_{ij} capture all possible twists of the torus $\mathbb{T}^d$. Modular invariance of the torus partition function for conformal field theories is a powerful constraint on the theory. The invariance can be stated

as the action of large conformal transformations on the lattice spanned by the set $\{U_i\}$. These large conformal transformations form the group $SL(d,\mathbb{Z})$ and act on the matrix U in (2.250) by left multiplication. $SL(d,\mathbb{Z})$ is generated by two elements [185]

$$S = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ (-1)^{d+1} & 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \tag{2.251}$$

They can be shown to generate any pairwise swap and a twist along any direction. For even d, we quotient by the center of the group $\{-1,1\}$ to obtain $PSL(d,\mathbb{Z})$, but for simplicity we will universally refer to the group as $SL(d,\mathbb{Z})$. Using scale invariance to unit-normalize one of the cycle lengths shows that we have $(d-1)(d+2)/2$ real moduli captured by the matrix U.

In spacetime dimension greater than two, modular transformations generically change the spatial background of the theory (i.e. change the Hilbert space), making it difficult to relate the low-lying states to the high-lying states on a fixed background. However, as discussed in [175] there exist two choices of torus which allow for a high-temperature/low-temperature duality to be considered. The first is the background $S^1_\beta \times S^1_L \times \mathbb{T}^{d-2}_{L_\infty}$, where $L_\infty \gg \beta, L, \beta^2/L$. In this case by appealing to extensivity in the large directions we have the approximate invariance

$$\log Z(\beta) \approx (L/\beta)^{d-2} \log Z(L^2/\beta)\,. \tag{2.252}$$

This can be transformed into an exact high-temperature/low-temperature duality by passing to a density defined by dividing $\log Z(\beta)$ by the volume of the large torus as it decompactifies, but we will not pursue that here.

To produce an exact invariance on a compact manifold, we can also consider a special torus given by $S^1_\beta \times S^1_L \times S^1_{L^2/\beta} \times \cdots \times S^1_{L^{d-1}/\beta^{d-2}}$, for which

$$Z(\beta) = Z(L^d/\beta^{d-1})\,. \tag{2.253}$$

This invariance is obtained by an $SL(d,\mathbb{Z})$ transformation and a scale transformation. It will play an important role in our CFT analysis.

To deal with the case of a general torus where there is no high-temperature/low-temperature duality, we will find it useful to define some notation. For a d-dimensional torus of side lengths $L_0, L_1, \ldots, L_{d-1}$, where $\beta = L_0$, we will denote the partition function quantized in an arbitrary channel as:

$$Z[\mathcal{M}^d] = Z(L_i)_{\mathcal{M}_i} = \sum e^{-L_i E_{\mathcal{M}_i}}\,. \tag{2.254}$$

$Z[\mathcal{M}^d]$ denotes the Euclidean path-integral representation of our partition function, which treats space and time democratically. The next form of the partition function picks direction i as time and gives a Hilbert space interpretation of the path integral. Since the spatial manifold will change depending on which direction is chosen as time, we use the notation $\mathcal{M}_i$ to explicitly denote the spatial manifold. It is defined as $\mathcal{M}^d = \mathcal{M}_i \times S^1_{L_i}$. Brackets will always imply a Euclidean path-integral representation while parentheses will imply a Hilbert-space representation.

Review of higher-dimensional Cardy formulas

Now we will provide a derivation of the higher-dimensional Cardy formula on an arbitrary spatial manifold $S^1_\beta \times X$. We will only need the result for a spatial torus, but we will keep the discussion general. The fact that modular transformations generically change the Hilbert space of the torus partition function will not provide an obstruction, although we will see in the resulting formulas that our high-temperature partition function and asymptotic density of states refer to the vacuum energy on a *different* spatial background in general.

We assume our theory to be local, modular invariant, and to have a spectrum of real energies on the torus that is bounded below by an energy that is discretely gapped from the rest of the spectrum. At asymptotically high temperature $\beta/V_X^{1/(d-1)} \to 0$, we can use extensivity of the free energy to replace our spatial manifold X with a torus $\mathbb{T}^{d-1}$ of cycle lengths $L_1 \leq L_2 \leq \cdots \leq L_{d-1}$ and no twists, with $V_X = L_1 \cdots L_{d-1} \equiv V_{\mathcal{M}_0}$. We therefore have

$$Z[S^1_\beta \times X] = Z(\beta)_X \approx Z(\beta)_{\mathcal{M}_0} = \sum e^{-\beta E_{\mathcal{M}_0}} \approx e^{\tilde{c} V_{\mathcal{M}_0}/\beta^{d-1}} \tag{2.255}$$

at asymptotically small β for some thermal coefficient $\tilde{c} > 0$. This thermal coefficient is not a priori related to any anomalies except in two dimensions. Considering a quantization along L_{d-1} gives us

$$Z(L_{d-1})_{\mathcal{M}_{d-1}} = \sum e^{-L_{d-1} E_{\mathcal{M}_{d-1}}} = e^{-L_{d-1} E_{\text{vac},\mathcal{M}_{d-1}}} \sum e^{-L_{d-1}(E-E_{\text{vac}})_{\mathcal{M}_{d-1}}} . \tag{2.256}$$

For $d = 2$ in a scale-invariant theory, β becoming asymptotically small is equiva-

lent to L_{d-1} becoming asymptotically large, since only the ratio L_{d-1}/β is meaningful. However, for $d > 2$ we have the additional directions L_i which may prevent us from interpreting the quantization in the L_{d-1} channel as a low-temperature partition function which projects to the vacuum. To deal with this, consider the limit $L_{d-1} \to \infty$ where we indeed project efficiently to the vacuum:

$$\lim_{L_{d-1}\to\infty} \frac{\log Z(L_{d-1})_{\mathcal{M}_{d-1}}}{L_{d-1}} = -E_{\text{vac},\mathcal{M}_{d-1}} \,. \tag{2.257}$$

Using $Z(\beta)_{\mathcal{M}_0} = Z(L_{d-1})_{\mathcal{M}_{d-1}}$ gives us $E_{\text{vac},\mathcal{M}_{d-1}} = -\tilde{c}V_{\mathcal{M}_{d-1}}/\beta^d$. We are therefore able to extract the scaling of the vacuum energy as $E_{\text{vac},\mathcal{M}_{d-1}} \propto -V_{\mathcal{M}_{d-1}}/\beta^d$ as $\beta \to 0$. The proportionality coefficient, which we define as ε_{vac}, is $\varepsilon_{\text{vac}} = \tilde{c}$. Furthermore, notice that $E_{\text{vac},\mathcal{M}_{d-1}}$ is clearly independent of L_{d-1}, so this result is general even though we took the limit $L_{d-1} \to \infty$ to obtain it. In the general case of arbitrary L_{d-1} we can therefore write for $\beta \to 0$

$$Z(L_{d-1})_{\mathcal{M}_{d-1}} = e^{\tilde{c}V_{\mathcal{M}_0}/\beta^{d-1}} \sum e^{-L_{d-1}(E-E_{\text{vac}})_{\mathcal{M}_{d-1}}} \,. \tag{2.258}$$

Again equating with $Z(\beta)_{\mathcal{M}_0}$, we see that the excited states must contribute at subleading order, since the vacuum contribution is sufficient to obtain $Z(L_{d-1})_{\mathcal{M}_{d-1}} = Z(\beta)_{\mathcal{M}_0}$ at leading order in small β. The concern over the directions L_i and poor projection to the vacuum alluded to earlier is therefore not a problem at leading order. We are finally left with

$$S(\beta) = (1 - \beta\partial_\beta) \log Z(\beta)_X \approx dV_X \varepsilon_{\text{vac}}/\beta^{d-1} \tag{2.259}$$

for the high-temperature entropy of a modular-invariant CFT on an arbitrary spatial background X.

Now we consider the implications for the density of states:

$$\rho(E_s) = \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} d\beta\, Z(\beta)_X e^{\beta E_s} \tag{2.260}$$

$$= \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} d\beta \left(e^{-\varepsilon_{\rm vac}V_X/\beta^{d-1}}\sum e^{-\beta E}\right) e^{\varepsilon_{\rm vac}V_X/\beta^{d-1}+\beta E_s}\,, \tag{2.261}$$

for some $\alpha > 0$. Performing a saddle-point on the part of the integrand outside of the parentheses and evaluating the integrand on this saddle $\beta_s \propto E_s^{-1/d}$ gives us the higher-dimensional Cardy formula:

$$\log\rho(E_s) = \frac{d}{(d-1)^{\frac{d-1}{d}}}(\varepsilon_{\rm vac}V_X)^{\frac{1}{d}}E_s^{\frac{d-1}{d}}\,. \tag{2.262}$$

The saddle point implies $\beta_s \to 0$ as $E_s \to \infty$. To ensure that this saddle point is valid, we need to check that the part of the integrand in the parentheses, which we call $\tilde{Z}_X(\beta)$, does not give a big contribution on the saddle:

$$\tilde{Z}_X(\beta_s) = e^{-\varepsilon_{\rm vac}V_X/\beta_s^{d-1}}\sum e^{-\beta_s E}\,. \tag{2.263}$$

From high-temperature ($\beta_s \to 0$) extensivity (2.255), we know that we can write this as

$$\tilde{Z}_X(\beta_s) \approx e^{-\varepsilon_{\rm vac}V_X/\beta_s^{d-1}}e^{\tilde{c}V_X/\beta_s^{d-1}} = 1\,, \tag{2.264}$$

where we used $\tilde{c} = \varepsilon_{\rm vac}$ (and one notices $\tilde{c}$ is independent of spatial background by replacing the high-temperature partition function on the given manifold with

the high-temperature partition function on a torus of spatial lengths $L_1, \ldots L_{d-1}$ with $V_{\mathcal{M}_0} = V_X$). Our saddle-point approximation is therefore justified, and we have the higher-dimensional Cardy formula as advertised.

In particular, considering the spatial background to be $X = S^{d-1}$ gives the asymptotic density of local operators by the state-operator correspondence. In the rest of this paper we will only be interested in the CFT on $\mathbb{T}^d$.

Review of vacuum energies in CFT

Normalization of vacuum energy

In a generic field theory, one is always free to shift the Hamiltonian by an arbitrary constant. This therefore shifts what we call the vacuum energy. Indeed, the well-known Casimir effect demonstrates that derivatives with respect to spatial directions dE_{vac}/dL_i are the physical observables, leaving an ambiguity in the normalization of E_{vac}. Additional structure, such as supersymmetry or modular invariance, disallows such an ambiguity. Even in a purely scale-invariant theory one can fix the normalization of the vacuum energy. Scale invariance requires that energies, and in particular the ground state energy, scale as inverse lengths under a rescaling of the spatial manifold: $E_{\text{vac}}(\lambda L_1, \lambda L_2, \ldots) = \lambda^{-1} E_{\text{vac}}(L_1, L_2, \ldots)$. This fixes the shift ambiguity in E_{vac}.

Subextensive corrections to the vacuum energy

The higher-dimensional Cardy formulas involves the vacuum energy density on $S^1 \times \mathbb{R}^{d-2}$, which by its relation to the extensive free energy density in a different channel is negative and has a fixed functional form. If we compactify

more directions and make them comparable to the size of the original S^1, then we will in general get corrections to the asymptotic formula. For two-dimensional CFT there is only one spatial cycle so no such corrections can enter. To capture the essence of what happens, let us consider a three-dimensional CFT on $S^1_\beta \times S^1_{L_1} \times S^1_{L_2}$ with $L_1 < L_2$. The low-temperature partition function will project to the vacuum state on $S^1_{L_1} \times S^1_{L_2}$, which can be parameterized as

$$E_{\text{vac},L_1\times L_2} = -\frac{\varepsilon_{\text{vac}} L_2}{L_1^2}(1 + f(L_1/L_2))\,. \tag{2.265}$$

Let us define $y = L_1/L_2$. The function $f(y)$ is capturing all of the corrections beyond the asymptotic formula, so we have $f(0) = 0$ and $f(y \to \infty) = -1 + y^3$. In general, $f(y)$ is a nontrivial function of y. Later in the text we will derive some positivity and monotonicity constraints on $f(y)$ by using modular invariance, but for now let us exhibit its functional form for the free boson theory, shown in figure 2.16.

In higher dimensions, there are more independent ratios that can be varied, and in general the corrections beyond the asymptotic formula are given by some nontrivial function of $d-2$ dimensionless ratios $y_i = L_1/L_i$ which for simplicity we will often write as $f(\mathbf{y})$ with $\mathbf{y} = (y_2, y_3, \ldots, y_{d-1})$.

We will also find it useful to consider the parameterization of the vacuum energy in arbitrary dimension as

$$E_{\text{vac}} = -\frac{\varepsilon_{\text{vac}} V_{d-1}}{L^d_{\text{min}}}\left(1 + \tilde{f}(\mathbf{y})\right). \tag{2.266}$$

which always has the smallest cycle in the denominator. The key difference be-

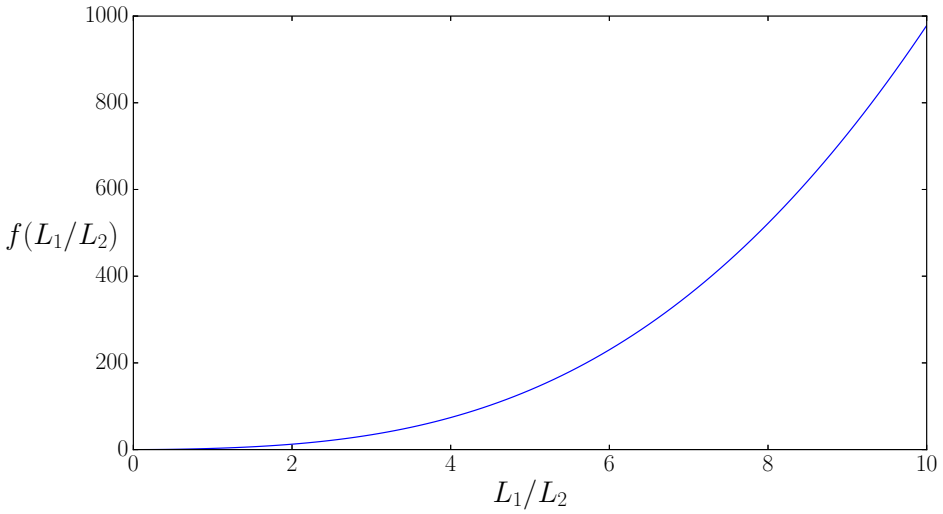


Figure 2.16: The functional form of $f(L_1/L_2)$ in the vacuum energy (defined in (2.265)) of a free boson in $2+1$ dimensions on a two-torus $\mathbb{T}^2$ with sides L_1 and L_2. As can be seen in the plot, $f(L_1/L_2)$ is positive and monotonically increasing.

tween $\tilde{f}(\mathbf{y})$ and $f(\mathbf{y})$ is that it is possible for $\tilde{f}(\mathbf{y})$ to be identically zero for all values of its arguments, whereas this is not the case for $f(\mathbf{y})$ as discussed in three dimensions above. We will find, for example, that gravity implies a vacuum energy structure with $\tilde{f}(\mathbf{y}) = 0$ up to $1/N$ corrections. We will often just write $\tilde{f}(\mathbf{y}) = 0$, by which we mean the equality up to $1/N$ corrections.

2.5.3 Phase structure of toroidally compactified AdS gravity

In this section we will recap what is known about the phase structure of gravity in AdS with a toroidally compactified boundary. This phase structure is easy to deduce for pure gravity without spontaneous breaking of translation invariance, which is the case we will restrict ourselves to. The most remarkable feature of

this phase structure is the absence of any nontrivial finite-size corrections to the vacuum energy and free energy, up to sharp phase transitions as circles become comparably sized. In other words, the function $\tilde{f}(\mathbf{y})$ defined in the previous section *vanishes* for all values of its arguments. As usual there will be nonzero contributions suppressed by $1/N$. Note that weakly coupled theories, including e.g. $\mathcal{N} = 4$ super Yang-Mills, do not realize this sort of structure [186]. We will not consider the possibility that the singular solutions used in [186] are relevant for the phase structure. An argument against them is as follows. Assume that such a singular solution provides the vacuum energy of the theory under multiple compactifications. By the higher-dimensional Cardy formula, there must therefore exist a black brane with higher entropy than AdS-Schwarzschild. Any such black brane should be modular S-related to the singular solution. But that means the "black brane" will be horizonless and singular, and if e.g. α' effects resolve the singularity and pop out a horizon, then the entropy should be proportional to some power of α'. But the ground state energy is a boundary term and is not proportional to α'. This is inconsistent, by the Cardy formula which relates the two.

We consider our theory at inverse temperature β on a spatial torus of side lengths L_i. The Euclidean solutions with the correct periodicity conditions are the toroidally compactified Poincaré patch, black brane, and $d-1$ AdS solitons

$$ds^2_{\text{pp}} = r^2 dx_0^2 + \frac{dr^2}{r^2} + r^2 d\varphi_i d\varphi^i \,, \tag{2.267}$$

$$ds^2_{\text{bb}} = r^2 \left(1 - (r_h/r)^d\right) dx_0^2 + \frac{dr^2}{r^2 \left(1 - (r_h/r)^d\right)} + r^2 d\varphi_i d\varphi^i \,, \tag{2.268}$$

$$ds^2_{\rm sol,k} = r^2 dx_0^2 + \frac{dr^2}{r^2\left(1-(r_{0,k}/r)^d\right)} + r^2\left(1-(r_{0,k}/r)^d\right)d\varphi_k^2 + r^2 d\varphi_j d\varphi^j\,, \quad (2.269)$$

all of which have the identification $x_0 \sim x_0 + \beta$. There are $d-1$ AdS solitons since there are $d-1$ circles that are allowed to pinch off in the interior. This means that we are picking supersymmetry-breaking boundary conditions around all cycles, which is motivated by maintaining S-invariance of our thermal partition function.

The parameter r_h ($r_{0,k}$) is fixed by demanding the x_0 (φ_k) circle caps off smoothly:

$$r_h = \frac{4\pi}{d\beta}\,, \qquad r_{0,k} = \frac{4\pi}{dL_k}\,. \quad (2.270)$$

Considering the ensemble at finite temperature and zero angular velocity, we need to compare the free energy of these solutions:

$$F_{\rm bb} = -\frac{r_h^d V_{d-1}}{16\pi G}\,, \qquad F_{\rm sol,k} = -\frac{r_{0,k}^d V_{d-1}}{16\pi G}\,, \qquad F_{\rm pp} = 0\,. \quad (2.271)$$

The Poincaré patch solution never dominates so we will not consider it in what follows. We will also assume that the AdS soliton of minimal energy gives the vacuum energy of the theory under a toroidal compactification [187].

Thermal phase structure

We will first consider the thermal phase structure, which can be illustrated by fixing a spatial torus and varying the inverse temperature β. The AdS soliton with the cycle of smallest length $L_{\rm min}$ pinching off has minimal free energy and dominates all the other ones. We will denote this as the $k = \min$ soliton. Thus,

the two relevant solutions are this $k = \min$ soliton and the black brane. These two exhibit a thermal phase transition at $\beta = L_{\min}$ with the black brane dominating the ensemble at high temperature $\beta < L_{\min}$. The energy at the phase transition is

$$E\big|_{r_h=\frac{4\pi}{dL_{\min}}} = -\partial_\beta \log Z = -(d-1)E_{\text{vac}}\,, \tag{2.272}$$

where $E_{\text{sol},k=\min} = F_{\text{sol},k=\min} = E_{\text{vac}}$ is the vacuum energy of the theory.

Quantum phase structure

A very important new feature in the phase structure of higher-dimensional toroidally compactified AdS spacetime is the existence of quantum phase transitions. These are phase transitions that can occur at zero temperature and are therefore driven by quantum fluctuations and not thermal fluctuations. They occur as we vary the spatial cycle sizes and reach a point where two spatial cycle sizes coincide and are minimal with respect to the rest. Let us call these cycle lengths L_1 and L_2 and pass from $L_1 < L_2$ to $L_1 > L_2$. In this case the vacuum energy exhibits a sharp transition from the $k = 1$ soliton to the $k = 2$ soliton. This is precisely the behavior that fixes $\tilde{f}(\mathbf{y}) = 0$, as alluded to earlier. To exhibit a phase transition in the free energy instead of the vacuum energy, we need to restrict ourselves to the low-temperature phase $\beta > L_{\min}$ where the black brane does not dominate.

2.5.4 Necessary and sufficient conditions for universality

In this section we would like to highlight a few difficulties in generalizing a discussion from two dimensions to higher dimensions. Let us first consider a two-dimensional CFT with cycle lengths β and L. For such a theory, vacuum domination of the torus partition function in channel L, for arbitrary cycle size $L > \beta$, is necessary and sufficient for universality of the partition function for all β. To see this, we write vacuum domination in the L channel as

$$Z(\beta)_L = Z(L)_\beta = \sum_{E_\beta} \exp\left(-LE_\beta\right) \approx \exp\left(-LE_{\text{vac},\beta}\right) = \exp\left(\frac{\pi cL}{6\beta}\right) . \quad (2.273)$$

Due to the fact that the vacuum energy for two-dimensional CFT is uniquely fixed by conformal invariance, we get a universal answer for the partition function. In the β channel, this form is that of an extensive free energy, and gives the Cardy formula in the canonical ensemble $S(\beta) = \pi cL/(3\beta)$.

In higher dimensions, vacuum domination of the torus partition function in one channel seems neither necessary nor sufficient for extensive Cardy growth in a different channel. This is because the vacuum energy on a generic torus is not uniquely fixed by conformal invariance. But it turns out we can use $SL(d, \mathbb{Z})$ invariance to show that a slightly modified version of the statement is valid. In particular, we will show that vacuum domination in all channels except that of the smallest cycle is necessary and sufficient for universality of the partition function for all β. Before we begin, we will prove some useful properties of the function $f(\mathbf{y})$ which characterizes the subextensive corrections to the vacuum energy and will play a starring role in our general CFT and symmetric orbifold analyses.

Sections 2.5.4 and 2.5.4 will contain results about generic modular-invariant CFTs. Sections 2.5.4 and 2.5.4 will then specify to large-N theories.

Modular constraints on vacuum energy

We now utilize the connection between the vacuum energy and the excited states implied by modular invariance, as first pointed out in appendix A of [175]. We will find that, somewhat surprisingly, modular invariance constrains all subextensive corrections to the vacuum energy to have a fixed sign and monotonic behavior.

Consider a spatial torus with side lengths $L_1 \leq \cdots \leq L_{d-1}$ and take the quantization along β at low temperature, which efficiently projects to the vacuum:

$$\lim_{\beta\to\infty} \frac{\log Z(\beta)_{\mathcal{M}_0}}{\beta} = -E_{\text{vac},\mathcal{M}_0} = \frac{\varepsilon_{\text{vac}} V_{\mathcal{M}_0}}{L_1^d}(1 + f(\mathbf{y}))\,. \tag{2.274}$$

We also consider the $d-2$ quantizations $L_2, \ldots, L_{d-1}$, which give

$$\lim_{\beta\to\infty} \frac{\log Z(L_i)_{\mathcal{M}_i}}{\beta} = \frac{\varepsilon_{\text{vac}} V_{\mathcal{M}_0}}{L_1^d}(1 + f(\mathbf{y} \setminus y_i, 0)) + \lim_{\beta\to\infty} \frac{1}{\beta} \log\left(\sum_E e^{-L_i(E-E_{\text{vac}})_{\mathcal{M}_i}}\right), \tag{2.275}$$

where $\mathbf{y} \setminus y_i$ is the vector $\mathbf{y}$ without the y_i-th element. The reason for the different arguments of f is that in the L_i quantization, instead of the ratio L_1/L_i we have $L_1/\beta = 0$ as $\beta \to \infty$. The second term on the right-hand-side does not vanish since the logarithm of the shifted partition function becomes linear in β at large β due to extensivity.

We want to analyze the monotonicity properties of $f(\mathbf{y})$ with respect to its $d-2$ arguments. To analyze any given ratio y_i, we can equate the quantization

along β with the quantization along L_i. This gives

$$\frac{\varepsilon_{\text{vac}} V_{\mathcal{M}_0}}{L_1^d} \left(f(\mathbf{y}) - f(\mathbf{y} \setminus y_i, 0)\right) = \lim_{\beta\to\infty} \frac{1}{\beta} \log\left(\sum_E \exp\left(-L_i (E - E_{\text{vac}})_{\mathcal{M}_i}\right)\right). \tag{2.276}$$

By unitarity, the right-hand-side is manifestly non-negative, so we conclude

$$f(\mathbf{y}) - f(\mathbf{y} \setminus y_i, 0) \geq 0\,. \tag{2.277}$$

Furthermore, the right-hand-side of (2.276) is a monotonically decreasing function of L_i. This means we can differentiate the left-hand-side with respect to L_i and obtain

$$f(\mathbf{y}) - f(\mathbf{y} \setminus y_i, 0) + L_i \partial_{L_i} f(\mathbf{y}) \leq 0 \implies \partial_{L_i} f(\mathbf{y}) \leq 0 \implies \partial_{y_i} f(\mathbf{y}) \geq 0\,, \tag{2.278}$$

where the first implication follows from the previous positivity property. The second implication follows from the fact that increasing L_i is the same as keeping all ratios y_j fixed except for the ratio $y_i = L_1/L_i$, which is decreased. In particular, this means that the function increases under any possible variation. Furthermore, since $f(\mathbf{0}) = 0$ this means that $f(\mathbf{y}) \geq 0$. These facts will be used heavily in what follows.

Modular invariance can also be used to constrain the behavior of the vacuum energy under spatial twists. By re-interpreting the spatial twist as an angular potential in a different channel, we can see that the vacuum energy cannot increase due to a spatial twist. The proof goes as follows. Consider the following partition function in the low-temperature limit with twist θ_{kj} between two spatial directions

k and j:

$$\lim_{\beta\to\infty}\frac{\log Z(\beta;\theta_{kj})_{\mathcal{M}_0}}{\beta}=-E_{\rm vac}(L_1,...,L_{d-1};\theta_{kj})\,. \tag{2.279}$$

Since the spatial directions are twisted, we may quantize along direction k, in which case the twist becomes an angular potential:

$$\lim_{\beta\to\infty}\frac{\log Z(L_k;\theta_{kj})_{\mathcal{M}_k}}{\beta}=\lim_{\beta\to\infty}\frac{1}{\beta}\log\left(\sum_E\exp\left(-L_kE_{\mathcal{M}_k}+iP_j\theta_{kj}\right)\right)\,. \tag{2.280}$$

The introduction of θ_{kj} only adds phases to the partition function in this channel, which decreases its real part. The vacuum energy is always manifestly real, so when equating the two quantizations it will be the case that the partition function with angular potential will evaluate to a real number. This means that the vacuum energy, which is negative, will be strictly greater or equal to its value without twists. This will be used in section 2.5.5.

Necessary and sufficient conditions

With the properties of the vacuum energy in hand, we are now ready to show that vacuum domination in all but the smallest channel is necessary and sufficient to have a universal free energy.

First we show *sufficiency*. We consider an ordering $\beta<L_1\leq\cdots\leq L_{d-1}$. Vacuum domination in the channels L_i means

$$Z(L_i)_{\mathcal{M}_i}=\exp\left(-L_iE_{{\rm vac},\mathcal{M}_i}\right)\,. \tag{2.281}$$

As we saw in the previous section, the vacuum energy is not uniquely fixed for

higher-dimensional CFTs. However, equating the $d-2$ quantizations lets us extract the vacuum energy:

$$Z(L_1)_{\mathcal{M}_1} = Z(L_2)_{\mathcal{M}_2} = \cdots = Z(L_{d-1})_{\mathcal{M}_{d-1}} \tag{2.282}$$

$$\implies -L_1 E_{\text{vac},\mathcal{M}_1} = -L_2 E_{\text{vac},\mathcal{M}_2} = \cdots = -L_{d-1} E_{\text{vac},\mathcal{M}_{d-1}}\,. \tag{2.283}$$

Since $E_{\text{vac},\mathcal{M}_i}$ is independent of L_i, we conclude that $E_{\text{vac},\mathcal{M}_i}$ is linear in the cycle lengths $L_{j\neq i}$. The β dependence is then fixed by dimensional analysis, and the coefficient is fixed by matching onto the asymptotic case of small β:

$$E_{\text{vac},\mathcal{M}_i} = -\varepsilon_{\text{vac}} V_{\mathcal{M}_i}/\beta^d. \tag{2.284}$$

Thus, we see that vacuum domination in all but the smallest channel determines the functional form of the vacuum energy. We can now use $Z(\beta)_{\mathcal{M}_0} = Z(L_i)_{\mathcal{M}_i}$ to get

$$Z(\beta)_{\mathcal{M}_0} = \exp\left(\varepsilon_{\text{vac}} L_i V_{\mathcal{M}_i}/\beta^d\right) = \exp\left(\varepsilon_{\text{vac}} V_{\mathcal{M}_0}/\beta^{d-1}\right)\,. \tag{2.285}$$

This is just the Cardy formula. In a regular CFT it holds only asymptotically in small β, but here we have shown that vacuum domination in the spatial channels L_i is sufficient to make it valid for all temperatures $\beta < L_i$. For $\beta > L_1$ we again have a universal expression for $Z(\beta)_{\mathcal{M}_0}$, which by assumption is given by the contribution of the vacuum only.

Showing that vacuum domination in all but the smallest cycle is *necessary* for universality requires the properties of $f(\mathbf{y})$ proven in the previous subsection.

Consider the quantization along an arbitrary channel of cycle size L_i

$$Z(L_i)_{\mathcal{M}_i} = \sum e^{-L_i E_{\mathcal{M}_i}} = e^{-L_i E_{\text{vac},\mathcal{M}_i}} \sum e^{-L_i (E-E_0)_{\mathcal{M}_i}} \,. \tag{2.286}$$

In two spacetime dimensions, it is the vacuum contribution in this channel that gives Cardy behavior in the β channel and therefore universality. The excited states contribute as positive numbers, and would ruin the Cardy behavior. Therefore it is necessary that they not contribute, i.e. necessary that we are vacuum dominated in this channel. In higher dimensions, one may worry that the excited state contributions cancel against the non-universal pieces of the vacuum energy, precluding the necessity of vacuum domination. However, by the positivity of $f(\mathbf{y})$ this can never happen. Thus, to get the correct Cardy behavior in the β channel it is necessary that the excited states do not contribute. This is true for arbitrary channel i. We conclude that it is necessary to be vacuum dominated in all but the smallest cycle.

It is interesting that for a universal free energy it is necessary and sufficient to have vacuum domination in all but the smallest channel. One could have suspected that explicit assumptions about the subextensive corrections to the vacuum energy would have to enter, but they do not.

We can state an equivalent set of necessary and sufficient conditions. To obtain a universal free energy for all β on an arbitrary rectangular torus, it is necessary and sufficient to have vacuum domination in the largest spatial cycle, with the vacuum energy taking the universal form with no subleading corrections. In fact, by using the non-negativity and monotonicity of the subextensive corrections,

we can state the necessary and sufficient condition as vacuum domination in the largest spatial cycle, with the vacuum energy on a square torus of side length L equal to $\varepsilon_{\rm vac}/L$.

In the rest of this section we will restrict attention to large-N theories.

Sparseness constraints without assuming $\tilde{f}(\mathbf{y}) = 0$

It is difficult to make progress in the case where we make no explicit assumptions about the functional form of the vacuum energy. To achieve vacuum domination in all but the smallest channel of a large-N theory, we can bound the entire spectrum on an arbitrary spatial torus of side lengths $L_1 \leq L_2 \leq \cdots \leq L_{d-1}$ as

$$\rho(\Delta_{\mathcal{M}_0}) \lesssim \exp\left(L_1 \Delta_{\mathcal{M}_0}\right) , \qquad \Delta_{\mathcal{M}_0} \equiv (E - E_{\rm vac})_{\mathcal{M}_0} . \tag{2.287}$$

This is a necessary and sufficient condition, although it is possible that it is implied by a more minimal set of necessary and sufficient conditions. To see how this condition arises, one writes the partition function as

$$Z(\beta)_{\mathcal{M}_0} = \exp\left(-\beta E_{\rm vac}\right) \sum \exp\left(-\beta \Delta_{\mathcal{M}_0}\right) \rho(\Delta_{\mathcal{M}_0}) \tag{2.288}$$

and bounds the density of states as (2.287) for the entire spectrum. At large N, with a vacuum contribution that scales exponentially in N, this suppresses all excited state contributions as soon as $\beta > L_1$. This means all cycles except the smallest will be vacuum-dominated, as required. We give another method of proof for vacuum domination in appendix 2.5.7 which restricts the sparseness bound to only the light states, but requires an additional assumption on the field theory.

We can also show that it is necessary and sufficient to solve the problem on a spatial square torus, i.e. that the free energy is universal for all β on a spatial square torus of side length L. The necessary direction is obvious. To show sufficiency, consider the quantization along L:

$$Z(L)_{\mathcal{M}_{d-1}} = \exp\left(-LE_{\text{vac},\mathcal{M}_{d-1}}\right) \sum_{\Delta} \exp\left(-L\Delta_{\mathcal{M}_{d-1}}\right) \tag{2.289}$$

$$= Z(\beta)_{\mathcal{M}_0} \approx \exp\left(\varepsilon_{\text{vac}} L^{d-1}/\beta^{d-1}\right) . \tag{2.290}$$

where the final expression is by assumption of universality. The only way to satisfy this equality is for the contribution of the excited states and the subextensive corrections to the vacuum energy in the L channel to vanish. In particular we are vacuum dominated in the L channel. Taking arbitrary $L_{d-1} > L$ keeps us vacuum dominated since it is at even lower temperature:

$$Z(L_{d-1})_{\mathcal{M}_{d-1}} \approx \exp\left(\varepsilon_{\text{vac}} L_{d-1} L^{d-2}/\beta^{d-1}\right) . \tag{2.291}$$

In the β channel this gives us the ordinary Cardy formula with no subextensive corrections, and in another L channel we have

$$Z(L)_{\mathcal{M}_{d-2}} = \exp\left(-LE_{\text{vac},\mathcal{M}_{d-2}}\right) \sum_{\Delta} \exp\left(-L\Delta_{\mathcal{M}_{d-2}}\right) \tag{2.292}$$

$$= Z(L_{d-1})_{\mathcal{M}_{d-1}} \approx \exp\left(\varepsilon_{\text{vac}} L_{d-1} L^{d-2}/\beta^{d-1}\right) . \tag{2.293}$$

Again, this means that we are vacuum dominated in the L channel. Now we can consider arbitrary L_{d-2} satisfying $L < L_{d-2} < L_{d-1}$, for which we will remain

vacuum dominated:

$$Z(L_{d-2})_{\mathcal{M}_{d-2}} \approx \exp\left(\varepsilon_{\text{vac}} L_{d-1} L_{d-2} L^{d-3}/\beta^{d-1}\right) . \tag{2.294}$$

By equating this expression with the partition function in the L_{d-1} channel, we see that we are still vacuum dominated in that channel. By continuing this procedure we are able to generalize to an arbitrary torus $\beta < L_1 < \cdots < L_{d-1}$, and we obtain

$$\log Z(\beta) = \begin{cases} \varepsilon_{\text{vac}} V_{\mathcal{M}_0}/\beta^{d-1}, & \beta < L_1 \\ \varepsilon_{\text{vac}} V_{\mathcal{M}_1}/L_1^{d-1}, & \beta > L_1 \end{cases} . \tag{2.295}$$

Altogether, we have that the free energy is universal at all temperatures on an arbitrary spatial torus. So solving the problem on a spatial square torus is both necessary and sufficient to solving the general problem, thanks to properties of the positivity of $f(\mathbf{y})$.

Sparseness constraints assuming $\tilde{f}(\mathbf{y}) = 0$

In this section we will show that assuming $\tilde{f}(\mathbf{y}) = 0$ (up to $1/N$ corrections) allows us to exhibit a constraint on the light spectrum that naturally generalizes the two-dimensional case. This is not too surprising, as $\tilde{f}(\mathbf{y}) = 0$ is automatically true in two dimensions, although some more work will be required in higher dimensions.

We start by considering the special torus with ordering $\beta < L < L^2/\beta < \cdots < L^{d-1}/\beta^{d-2}$. As discussed in the introduction, this special torus has an exact low-temperature/high-temperature duality $Z(\beta)_{\mathcal{M}_0} = Z(L^d/\beta^{d-1})_{\mathcal{M}_0}$. This will allow

us to uplift the arguments of [156] to our case. In the upcoming manipulations, we will not keep explicitly the specification of the spatial manifold $\mathcal{M}_0$, since this duality allows us to keep our spatial manifold fixed once and for all.

By following the steps in [156], one can show that the partition function is dominated by the light states up to a theory-independent error. We will denote light states as those with energy $E < \epsilon$ for some arbitrary ϵ. We have

$$\log Z_{\text{light}}(L^d/\beta^{d-1}) \leq \log Z(\beta) \leq \log Z_{\text{light}}(L^d/\beta^{d-1}) - \log\left(1 - e^{\varepsilon(\beta - L^d/\beta^{d-1})}\right). \tag{2.296}$$

This error grows arbitrarily large as $\beta \to L$ or $\epsilon \to 0$. For $\beta > L$ we can derive a similar upper and lower bound.

For a family of CFTs labeled by N, we assume that the vacuum energy also scales with N. This will be true in all examples we consider. When taking N large, we can scale $\epsilon \to 0$, in which case the partition function is squeezed by its bounds and given just by the light states up to $\mathcal{O}(1)$ corrections. In the context of assuming $\tilde{f}(\mathbf{y}) = 0$, we then obtain universality

$$\log Z(\beta) = \begin{cases} \log Z_{\text{light}}(L^d/\beta^{d-1}) = -\frac{L^d}{\beta^{d-1}} E_{\text{vac}} & \beta < L \\ \log Z_{\text{light}}(\beta) = -\beta E_{\text{vac}} & \beta > L \end{cases}, \tag{2.297}$$

if and only if the density of light states is bounded as

$$\rho(\Delta) \lesssim \exp\left(\frac{L^d}{\beta^{d-1}}\Delta\right), \qquad \Delta \leq -E_{\text{vac}}\,, \tag{2.298}$$

where $\Delta = E - E_{\text{vac}}$. Notice that if we did not assume a universal form for

the vacuum energy with $\tilde{f}(\mathbf{y}) = 0$, the free energy would still be very theory-dependent.

To generalize the argument above to an arbitrary d-torus, the idea will be to push the special torus very close to the square torus. From here, we can use the fact that whenever a partition function is dominated by the vacuum contribution at some inverse temperature β, then it will also be dominated by that contribution for larger β. Channel by channel, we will see that we will be able to generalize to an arbitrary torus. Assuming a universal form of the vacuum energy will be crucial for this argument.

It will be convenient to consider starting with a quantization along the L^{d-1}/β^{d-2} channel, because it is the largest cycle when $\beta < L$. We will now restore the explicit spatial manifold dependence since we will be considering quantizations along different channels. We have

$$Z(L^{d-1}/\beta^{d-2})_{\mathcal{M}_{d-1}} = Z(L^d/\beta^{d-1})_{\mathcal{M}_0} = Z(\beta)_{\mathcal{M}_0}\,. \tag{2.299}$$

By using (2.297) we can write this as

$$Z(L^{d-1}/\beta^{d-2})_{\mathcal{M}_{d-1}} = \exp\left(-\frac{L^d}{\beta^{d-1}}E_{\text{vac},\mathcal{M}_0}\right) = \exp\left(-\frac{L^{d-1}}{\beta^{d-2}}E_{\text{vac},\mathcal{M}_{d-1}}\right)\,. \tag{2.300}$$

This means that we are vacuum-dominated in the L^{d-1}/β^{d-2} channel.

Let us now take a larger cycle $L_{d-1} > L^{d-1}/\beta^{d-2}$, for which we will remain vacuum-dominated:

$$Z(L_{d-1})_{\mathcal{M}_{d-1}} = \exp\left(-L_{d-1}E_{\text{vac},\mathcal{M}_{d-1}}\right)\,. \tag{2.301}$$

Quantizing now along the the second largest cycle $L^{d-2}/\beta^{d-3} < L_{d-1}$ gives us

$$Z(L^{d-2}/\beta^{d-3}) = \exp\left(-L^{d-2}/\beta^{d-3} E_{\text{vac},\mathcal{M}_{d-2}}\right) \sum_{\Delta} \exp\left(-L^{d-2}\Delta_{\mathcal{M}_{d-2}}/\beta^{d-3}\right). \tag{2.302}$$

But by our assumption $\tilde{f}(\mathbf{y}) = 0$, we have

$$L_{d-1} E_{\text{vac},\mathcal{M}_{d-1}} = L^{d-2} E_{\text{vac},\mathcal{M}_{d-2}}/\beta^{d-3}, \tag{2.303}$$

which means that $Z(L^{d-2}/\beta^{d-3})$ is given by its vacuum contribution only. One can now consider $L_{d-2} > L^{d-2}/\beta^{d-3}$, for which we will remain vacuum-dominated in the L_{d-2} channel. By comparing to the L_{d-1} channel, we can verify that we remain vacuum-dominated there as well. We can now move to the L^{d-3}/β^{d-4} channel and continue this procedure up to and including the L channel. In a final step, we can compare to the β channel and see that it indeed has universal Cardy behavior:

$$\log Z(\beta) = \frac{\varepsilon_{\text{vac}} V_{\mathcal{M}_0}}{\beta^d}. \tag{2.304}$$

There is no need now to consider smaller β since we have already considered general variations of the other $d-1$ cycles. Since the partition function is a function of $d-1$ independent dimensionless ratios, we have already captured all possible variations.

The generality of the torus that results from this procedure is restricted by the special torus with which we began. But notice that the special torus can be arbitrarily close to a d-dimensional square torus, which means this procedure results in a universal free energy on an arbitrary torus. From this argument it

is clear that the only assumption made on the spectrum is the bound in (2.298). In fact, it is enough to impose this constraint for the square torus, since our procedure begins from that case (or arbitrarily close to it) and generalizes to an arbitrary torus. The sparseness constraint is therefore

$$\rho(\Delta) \lesssim \exp\left(L\Delta_{L\times L\times\cdots\times L}\right) \tag{2.305}$$

and is imposed only on the states with energies $E = \Delta + E_{\text{vac}} < 0$.

2.5.5 Symmetric Product Orbifolds in $d > 2$

In this section we construct orbifold conformal field theories in higher dimensions using a procedure analogous to the one in two dimensions. We will see that these theories contain both twisted and untwisted sector states and will give an estimate for the density of states within these sectors. Finally, we will show that under the assumption that $\tilde{f}(\mathbf{y}) = 0$, the free energy has a universal behavior at large N which agrees with Einstein gravity.

A review of permutation orbifolds in two dimensions

In two dimensions, symmetric product orbifolds (or the more general permutation orbifolds) provide a vast landscape of two-dimensional CFTs with large central charge that have a potentially sparse spectrum and are thus of interest in the context of holography [188, 189, 190, 191]. The goal of this section will be to extend these constructions to higher dimensions. We start by a review of permutation orbifolds in two dimensions which will set most of the notation that we will

then carry over to higher dimensions. Permutation orbifolds are defined by the choice of two parameters: a "seed" CFT $\mathcal{C}$ and a permutation group $G_N \subset S_N$. A permutation orbifold $\mathcal{C}_N$ is then defined to be

$$\mathcal{C}_N \equiv \frac{\mathcal{C}^{\otimes N}}{G_N} \,. \tag{2.306}$$

The procedure by which we take this quotient is called an orbifold. It projects out all states of the product theory that are not invariant under the action of the group. The Hilbert space thus gets restricted to

$$\mathcal{H}^{\otimes N} \;\longrightarrow\; \frac{\mathcal{H}^{\otimes N}}{G_N} \,, \tag{2.307}$$

where $\mathcal{H}$ is the Hilbert space of $\mathcal{C}$. This projection onto invariant states is crucial as it gets rid of most of the low-lying states and hence provides some hope of obtaining a sparse spectrum. When computing the torus partition function, this projection onto invariant states is implemented by a sum over all possible insertions of group elements in the Euclidean time direction. This is summarized by the following formula

$$Z_{\text{untw}} = \frac{1}{|G_N|} \sum_{g \in G_N} \; g \boxed{} \tag{2.308}$$

where the box represents the torus with the vertical direction being Euclidean time.

However, (2.308) is obviously not modular invariant as it singles out the time

direction. Modular invariance is restored in the following way

$$Z_{\text{tot}} = \frac{1}{|G_N|} \sum_{g,h \in G_N | gh=hg} g \boxed{}_{\,h} \,. \tag{2.309}$$

The requirement that the two group elements must commute comes from demanding that the fields have well-defined boundary conditions [192]. The insertion of elements h in the spatial direction are interpreted as twisted sectors, where the boundary conditions of the fields are twisted by group elements. There is one twisted sector per conjugacy class of the group, which in the case of $G_N = S_N$ gives one twisted sector per Young diagram. In [188, 189, 190], the space of permutation orbifolds was explored and a criterion was given for these theories to have a well-defined large N limit (and thus a potential holographic dual). It was found that many properties of the spectrum depends solely on the group G_N and not on the choice of the seed theory. Groups that give a good large-N limit are called oligomorphic permutation groups [193, 194, 195]. Although a complete proof is still missing, it is believed permutation orbifolds by oligomorphic groups all have at least a Hagedorn density of light states, but the growth may be even faster [188, 190]. For the symmetric group, it was shown in [183, 156] that the growth is exactly Hagedorn with the precise coefficient saturating the bound on the density of light states produced in [156].

Symmetric product orbifolds thus reproduce the phase structure of 3d gravity. Note that they are still far from local theories of gravity such as supergravity on $AdS_3 \times S^3$, as their low-lying spectrum is Hagedorn and so they look more like

classical string theories. The D1-D5 CFT has a moduli space that is proposed to contain a point, known as the orbifold point, where the theory becomes a free symmetric product orbifold theory. According to this proposal, the orbifold point is connected to the point where the supergravity description is valid by an exactly marginal deformation. It is only the strongly coupled theory that is dual to supergravity, and from this point of view it is surprising that the free theory realizes the phase structure of gravity.

Symmetric product orbifolds in higher dimensions

In two dimensions, we saw that symmetric product orbifolds are examples of theories with a sparse enough spectrum to satisfy the bound from [156] and thus have a universal phase structure at large N. We would now like to construct weakly coupled examples of theories satisfying our new criteria in higher dimensions. In dimensions greater than two, it is in general much harder to construct large-N CFTs. One may of course take tensor products but these will never have a sparse enough spectrum. In fact, the spectrum below some fixed energy level will not even converge as $N \to \infty$. Imposing some form of Gauss' law to project out many of the low-lying states is usually done by introducing some coupling to a gauge field, which makes preserving conformal invariance highly non-trivial. A natural way to achieve this same projection is through the construction of orbifold conformal field theories familiar from two dimensions. To the best of our knowledge, there is no construction of orbifold conformal field theories in higher dimensions, which as explained in the previous subsection is perhaps the most natural way of obtaining theories that are conformal, have a large number of degrees

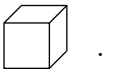
of freedom, but also a sparse low-lying spectrum.

We will now describe the construction of symmetric product orbifolds in d dimensions.[15] We will construct the partition function, i.e. the Hilbert space and the spectrum of the Hamiltonian on $\mathbb{T}^{d-1}$. We comment on other properties of the theory such as correlation functions in the discussion section.

The starting point is again to consider a seed $\mathrm{CFT}_d\,\mathcal{C}$ and to define the orbifold theory $\mathcal{C}_N$ as

$$\mathcal{C}_N \equiv \frac{\mathcal{C}^{\otimes N}}{S_N} \tag{2.310}$$

The orbifolding procedure goes as follows. We start by projecting onto invariant states by inserting all elements of the group in the time direction. This gives

$$Z_{\text{untw}} = \frac{1}{N!}\sum_{g\in S_N} \; g \; \text{  } \quad . \tag{2.311}$$

The box of the 2d case has now been lifted to a d-dimensional hypercube which again describes the torus. We will represent it by a $3d$ cube and leave the other dimensions implicit. Again, the mere projection is obviously not modular invariant. By applying elements of $SL(d,\mathbb{Z})$ (for instance the S element given in (2.251)), we quickly see that group elements must also be inserted in the space directions. Having well-defined boundary conditions for the fields constrains the d group elements to be commuting. The partition function of the orbifold theory is then

[15] Here we will assume that the group is S_N but the generalization to other permutation groups follows trivially from our construction.

defined as

$$Z_{\rm orb} = \frac{1}{N!} \sum_{\substack{g_0,\ldots,g_{d-1}\in S_N \\ g_i g_j = g_j g_i \forall i,j}} g_0 \square^{g_1}_{g_{d-1}} \quad . \tag{2.312}$$

Twisted sector states will correspond to any states with non-trivial insertions in any of the space directions. The different twisted sectors are no longer labeled just by conjugacy classes, but by sets of $d-1$ commuting elements, up to overall conjugation. This orbifolding procedure describes a well-defined $SL(d,\mathbb{Z})$-invariant partition function.

Spectrum of the theory

The untwisted sector states

We now turn our attention to the spectrum of these orbifold theories. Other properties will depend strongly on the choice of seed. We start by considering the untwisted sector states. These are given by states of the product theory, up to symmetrization. From the point of view of the partition function, their contribution consists of all elements in the sum (2.312) where $g_1 = ... = g_{d-1} = 1$. Consider the contribution of a K-tuple to the density of states. A K-tuple is a state where K of the N CFTs are excited, while the other $N-K$ are in the vacuum. The contributions of all possible K-tuples of distinct states are encapsulated by the following expression:

$$\rho(\Delta) = \int dK \int d\Delta_1 .... d\Delta_K \frac{1}{K!} \rho_0(\Delta_1) .... \rho_0(\Delta_K) \delta(\Delta - \sum_{j=1}^{K} \Delta_j) \,, \tag{2.313}$$

where $\Delta = E - NE_{\rm vac}$, $\Delta_i = E_i - E_{\rm vac}$ and ρ_0 is the density of states of the seed theory.[16] It can be shown that the contribution of K-tuples with subsets of identical states do not give a larger contribution than the one considered here, so it is sufficient to focus on this case. The combinatorial prefactor $1/K!$ was introduced to remove the equivalent permutations of the K states. One way to understand its inclusion is to consider how the orbifold projection is done. A given K-tuple in the product theory is made S_N invariant by summing over all of its possible permutations. For example, the 3-tuples $\{a,b,c\}, \{a,c,b\}, \{b,a,c\}, \{b,c,a\}, \{c,a,b\}, \{c,b,a\}$ of the pre-orbifolded theory lead to the same orbifolded 3-tuple and thus should only be counted once. The triple integral giving $\rho(\Delta)$, when left to its own devices without combinatorial prefactor, would count all six configurations.

Along with the states being distinct, let us first assume that each of the individual degeneracies can be approximated by the Cardy formula of the seed theory. The Cardy formula in higher dimensions was given in (2.262) and reads

$$\log \rho(E) = \frac{d}{(d-1)^{\frac{d-1}{d}}} (\varepsilon_{\rm vac} V_{d-1})^{\frac{1}{d}} E^{\frac{d-1}{d}} \,. \tag{2.314}$$

Now let us proceed as in [188] to find the density of states. Performing the integrals over energies E_i by a saddle-point approximation where the large parameter is the total energy E, we find saddle-point values $E_i = E/K$ for all i. To assure that the state in each copy is distinct, we need the degeneracy to pick from to be much larger than K. Thus the validity of this assumption and the validity of the Cardy

[16] Here we use the notation that Δ is a shifted energy that satisfies $\Delta \geq 0$, but we wish to emphasize that it is *not* in any way related to the scaling dimension of a local operator.

formula in each seed theory require, respectively,

$$\exp\left[\frac{d}{(d-1)^{\frac{d-1}{d}}}(\varepsilon_{\rm vac}V_{d-1})^{\frac{1}{d}}(\Delta/K+E_{\rm vac})^{\frac{d-1}{d}}\right]\gg K,\quad \Delta/K\gg|E_{\rm vac}|\,. \tag{2.315}$$

We will check whether these conditions are satisfied at the end. Note that the second constraint implies that we can drop $E_{\rm vac}$ in the Cardy formula when expressed in terms of Δ. We thus have

$$\rho(\Delta)\sim\int dK\exp\left[daK^{\frac{1}{d}}\Delta^{\frac{d-1}{d}}-K\log K+K\right] \tag{2.316}$$

with

$$a\equiv\frac{1}{(d-1)^{\frac{d-1}{d}}}(\varepsilon_{\rm vac}V_{d-1})^{\frac{1}{d}}\,. \tag{2.317}$$

We can now do a second saddle-point approximation to evaluate the integral over K. The large parameter is again given by the total shifted energy Δ. The saddle point equation is

$$a\Delta^{\frac{d-1}{d}}K_s^{\frac{1-d}{d}}-\log K_s=0\,, \tag{2.318}$$

which gives

$$K_s\sim\frac{a^{\frac{d}{d-1}}\Delta}{\left(\log\left[a^{\frac{d}{d-1}}\Delta\right]\right)^{\frac{d}{d-1}}} \tag{2.319}$$

at large Δ. Plugging this back in the density of states we find

$$\rho(E)\sim\exp\left[(d-1)\frac{a^{\frac{d}{d-1}}\Delta}{\left(\log\left[a^{\frac{d}{d-1}}\Delta\right]\right)^{\frac{1}{d-1}}}\right], \tag{2.320}$$

where we have used large Δ to drop subleading pieces which either have a larger

power of the logarithm in the denominator or are terms proportional to $\log\log\Delta$. We find a growth of states that is slightly sub-Hagedorn and the growth increases with the dimension of the field theory. Inserting K_s in our necessary assumptions shows that they can be satisfied for large enough Δ. In particular, the second condition becomes

$$a^{\frac{d}{d-1}}\Delta \gg \exp\left[a\left|E_{\text{vac}}\right|^{\frac{d-1}{d}}\right] \tag{2.321}$$

which is then sufficient to satisfy the first condition. Here E_{vac} is the vacuum energy of the seed theory and does not scale with N. Notice also that K_s grows with Δ and must not violate the bound $K_s \leq N$. This implies a bound on our energies from the saddle:

$$a^{\frac{d}{d-1}}\Delta \lesssim N\left[\log(N)\right]^{\frac{d}{d-1}} \ . \tag{2.322}$$

So altogether our density of states formula is reliable in the range

$$\exp\left[a\left|E_{\text{vac}}\right|^{\frac{d-1}{d}}\right] \ll a^{\frac{d}{d-1}}\Delta \lesssim N\left[\log(N)\right]^{\frac{d}{d-1}} \ . \tag{2.323}$$

In particular we can consider energies that scale with N. However, as we will shortly see, the density of states quickly becomes dominated by the twisted sectors. Note that this growth of states is also a lower bound for any permutation orbifold as orbifolding by a subgroup of S_N always projects out fewer states.

The twisted sector states

We will now give a lower bound on the density of states coming from the twisted sectors. If the intuition from two dimensions carries over, it will be the twisted sectors that give the dominant contribution to the density of states. Indeed, this is the result we will find. We start by a more general discussion of twisted sector states and their contribution to the partition function.

A twisted sector is given by $d-1$ commuting elements $g_1, ..., g_{d-1}$ of S_N, up to overall conjugation. There is also a projection onto S_N-invariant states by summing over elements in the time direction but at this point we only focus on the identity contribution in that direction. We define $\mathbb{T}$ to be the original d-torus used to compute the partition function. We leave the dependence on the vectors $U_0, ..., U_{d-1}$ implicit. Let us consider the action of the subgroup $G_{g_1,...,g_{d-1}}$ of S_N (defined to be the group generated by $g_1, ..., g_{d-1}$) on the N copies of the CFT. The action of this group will be to glue certain copies of the CFT together. Concretely, let Φ^k denote a field on $\mathbb{T}$ of the k-th CFT, then in the twisted sector defined by $G_{g_1,...g_{d-1}}$ this field has boundary conditions

$$\Phi^k(x_0, x_1, ..., x_j + L_j, ..., x_{d-1}) = \Phi^{g_j(k)}(x_0, x_1, ..., x_j, ..., x_{d-1})\,. \tag{2.324}$$

Tracking the orbit of the k-th copy under $G_{g_1,...g_{d-1}}$ allows us to define a single field $\tilde{\Phi}$ with modified boundary conditions. In particular it will have larger periods. A field $\tilde{\Phi}^i$ can be defined for each orbit of the group $G_{g_1,...g_{d-1}}$ and we will denote

the set of these orbits by

$$\{O_i\}\,, \quad i = 1, ...i_{\max}\,, \tag{2.325}$$

where $i_{\max}$ depends on the precise choice of $g_1, ..., g_{d-1}$. As the different orbits do not talk to each other, the path integral will split into a product of $i_{\max}$ independent path integrals, one over each field $\tilde{\Phi}^i$. The new boundary conditions of the fields in a given O_i under the action of $G_{g_1,...,g_{d-1}}$ enable us to rewrite that particular contribution to the path integral as a torus partition function, but now with $\mathbb{T}$ replaced by a new torus $\tilde{\mathbb{T}}_i$. The original identifications coming from (2.250) were

$$(x_0, x_1, .., x_{d-1}) \sim (x_0, x_1, .., x_{d-1}) + \sum_{i=0}^{d-1} n_i U_i\,. \tag{2.326}$$

for any integers n_i. Once the elements $g_1, ..., g_{d-1}$ are inserted the identifications are changed and they are encoded in a new torus. As these boundary conditions follow from the orbits, the identifications from the new torus are given by the elements in $G_{g_1,...,g_{d-1}}$ that leave the orbit invariant, i.e.

$$g_1^{m_1}...g_{d-1}^{m_{d-1}} O_i = O_i\,. \tag{2.327}$$

This means that the identifications become

$$(x_0, x_1, .., x_{d-1}) \sim (x_0, x_1, .., x_{d-1}) + \sum_{i=0}^{d-1} m_i U_i\,. \tag{2.328}$$

with the m_i such that (2.327) is satisfied. Alternatively, one can define new vectors

in the following way

$$\begin{aligned}
\tilde{U}_1 &= m_1^{\min} U_1 + m_{1,2} U_2 + ... + m_{1,d-1} U_{d-1} \,, \\
&\vdots \\
\tilde{U}_{d-2} &= m_{d-2}^{\min} U_{d-2} + m_{d-2,d-1} U_{d-1} \,, \\
\tilde{U}_{d-1} &= m_{d-1}^{\min} U_{d-1} \,,
\end{aligned} \tag{2.329}$$

where $m_{d-1}^{\min}$ is the smallest integer m_{d-1} such that $g_{d-1}^{m_{d-1}} O_i = O_i$, $(m_{d-2,d-1}, m_{d-2}^{\min})$ are the pair with smallest non-zero m_{d-2} such that $g_{d-1}^{m_{d-2,d-1}} g_{d-2}^{m_{d-2}^{\min}} O_i = O_i$ and the $(m_1^{\min}, ..., m_{1,d-1})$ are the set of integers with minimal non-zero m_1 such that (2.327) is satisfied. These vectors define a new torus $\tilde{\mathbb{T}}_i$ with volume

$$\mathrm{Vol}(\tilde{\mathbb{T}}_i) = \left(\prod_j m_j^{\min} \right) \mathrm{Vol}(\mathbb{T}) \equiv |O_i| \mathrm{Vol}(\mathbb{T}) \,. \tag{2.330}$$

Since the g_i commute, $|O_i|$ is just the number of elements in the orbit O_i.

A twisted sector will thus give a set of new tori $\tilde{\mathbb{T}}_i$ whose different volumes depend on the orbits of the action of $G_{g_1,...,g_{d-1}}$. For each orbit of that action, we will get a separate torus and schematically, this will give a contribution to the partition function of the form

$$Z_{\mathrm{tot}} \sim \prod_i Z(\tilde{\mathbb{T}}_i) \,, \tag{2.331}$$

where the product over i is a product over the orbits. This is a generalization of

Bantay's formula [179] to higher dimensions. For every orbit O_i we have

$$\mathrm{Vol}(\tilde{\mathbb{T}}_i) = |O_i|\mathrm{Vol}(\mathbb{T})\,, \tag{2.332}$$

where $|O_i|$ is the length of the orbit. We will now calculate the contribution to the partition function from a single non-trivial orbit of length $L = M^{d-1}$ giving a torus with equal rescaling M in all spatial directions. For simplicity, we also consider a case with $m_{i,j} = 0\ \forall i \neq j$. The torus $\tilde{\mathbb{T}}_i$ corresponding to this orbit is then

$$(\tilde{U}_0, ..., \tilde{U}_{d-1}) = (U_0, MU_1, ..., MU_{d-1})\,. \tag{2.333}$$

We can always find elements $g_1, ..., g_{d-1}$ that produce the desired torus with equal scaling of the spatial cycles. To produce the new torus given in (2.333), we use for example the following elements:

$$\begin{aligned}
g_1 &= (1\ ...\ M)\,(M+1\ ...\ 2M)\,...\,(M^{d-1}-M+1\ ...\ M^{d-1})(M^{d-1}+1)\ ...\ (N)\\
g_2 &= (1\ \ M+1\ ...\ M(M-1)+1)\,...\\
&\quad\ (M^{d-1}-M(M-1)\ \ \ M^{d-1}-M(M-2)\ ...\ M^{d-1})(M^{d-1}+1)\ ...\ (N)\\
&\ \ \vdots \\
g_{d-1} &= (1\ \ M^{d-2}+1\ ...\ M^{d-2}(M-1)+1)...(M^{d-2}\ \ 2M^{d-2}\ ...\ M^{d-1})\\
&\quad\ (M^{d-1}+1)\ ...\ (N)
\end{aligned} \tag{2.334}$$

for $L = M^{d-1}$. For example in $d = 3$ and for $L = 9$, we get

$$g_1 = (1\ 2\ 3)(4\ 5\ 6)(7\ 8\ 9)(10)...(N)\,,$$

$$g_2 \;=\; (1\ 4\ 7)(2\ 5\ 8)(3\ 6\ 9)(10)...(N)\,. \tag{2.335}$$

One can quickly check that all these elements commute and that they define an orbit of length L as well as $N-L$ singlets. One can also check that $m_1^{\min} = ... = m_{d-1}^{\min} = M$. We will call $Z_{\rm sq}$ this particular contribution to the partition function, and it reads

$$\begin{aligned} Z_{\rm sq} &= Z(U_0, U_1, ..., U_{d-1})^{N-L} Z(U_0, MU_1, ..., MU_{d-1}) \\ &= Z(U_0, U_1, ..., U_{d-1})^{N-L} Z(U_0/L^{\frac{1}{d-1}}, U_1, ..., U_{d-1})\,, \end{aligned} \tag{2.336}$$

where we uniformly rescaled the torus and used $L = M^{d-1}$. From this, we can infer the behaviour of the density of states:

$$Z_{\rm sq} = \sum_E \rho_{\rm sq}(E) e^{-\beta E} = e^{-\beta E_{\rm vac}(N-L)}\,(1 + \dots) \sum_E \rho_0(E) e^{-\beta E/L^{\frac{1}{d-1}}}\,. \tag{2.337}$$

We can ignore the excited states encapsulated in "..." as they will only increase $\rho_{\rm sq}(E)$, which will increase our final answer. In this section, we are only after a lower bound for the density of states so we can ignore such terms. Shifting E to $L^{\frac{1}{d-1}}(E - E_{\rm vac}(N-L))$ gives us

$$\rho_{\rm sq}(E) = \rho_0(L^{\frac{1}{d-1}}(E - E_{\rm vac}(N-L)))\,. \tag{2.338}$$

This will be the key formula to derive the final result.

In the full partition function we sum over all $L \leq N$ and for large L, we are in a regime where we may use the Cardy formula of the seed theory given in (2.262).

To find the twisted sector that gives the maximal contribution at energy E, we evaluate the sum over L using a saddle point approximation. The resulting saddle point equation for L is solved by

$$L_s = \frac{(E_{\rm vac} N - E)}{d E_{\rm vac}} \,, \tag{2.339}$$

which will be a good approximation provided $L_s \gg 1$. We now plug this back in (2.338) and use the Cardy formula (2.262) to obtain

$$\rho(E) \sim \exp\left[a \frac{(d-1)^{\frac{d-1}{d}}}{|E_{\rm vac}|^{1/d}} (E - N E_{\rm vac}) \right] . \tag{2.340}$$

Note that this is a Hagedorn growth as in two dimensions but the coefficient of the Hagedorn growth depends on the vacuum energy of the seed theory. This is somewhat a loss of universality compared to two dimensions and it will be very important in what follows to understand precisely the properties of the vacuum energy of the orbifold theory. This will be the task of the next subsection. The regime in which this expression is reliable is for $1 \ll L_s \leq N$ which in terms of energies is

$$1 \ll \frac{E - N E_{\rm vac}}{|E_{\rm vac}|} \leq dN \,. \tag{2.341}$$

Finally, it is important to emphasize that this is merely a lower bound on the density of states[17]. We have only given the contribution from one type of twisted sectors and other sectors might dominate. We have also not taken into account

[17] In fact, the method used in this section only gives an estimate for the lower bound. We have only inserted one element - the identity - in the time direction and have not taken into account the projection to S_N invariant states. Following the method we will use in section 2.5.5 one can show that this estimate is actually precise.

the projection onto S_N invariant states by inserting commuting elements of the group in the time direction. In two dimensions, one can show that the estimate coming from this particular twisted sector (called long strings in $2d$) actually gives the dominant contribution. We will discuss this further when analyzing the free energy but we first turn our attention to the vacuum energy.

Vacuum energy of the orbifold theory

We want to understand precisely the properties of the vacuum energy of the orbifold theory. In two dimensions, it is clear that the central charge gets multiplied by N when going from the seed theory to the product (or orbifold) theory. Since the vacuum energy is fixed by the central charge, it also gets multiplied by N. Naively, one would expect a similar behavior in higher dimensions. The all-vacuum contribution in the untwisted sector indeed has energy NE_{vac}, but it may be possible that other twisted sectors give even more negative contributions. We will now address this possibility and show that it is impossible, so that the vacuum energy of the orbifold theory is in fact given as

$$E_{\text{vac}}^{\text{orbi}} = NE_{\text{vac}}\,. \tag{2.342}$$

To prove this, first recall that it is not necessary to consider twisted sectors inducing twists between any of the dimensions because they always increase the vacuum energy, as explained in section 2.5.4. The only thing we need to check is that rescalings of the torus do not give a contribution that is more negative than (2.342). A twisted sector in principle gives a product of partition functions

if there is more than one orbit, but it will suffice to consider the case of a single orbit. This is because if there are different orbits, the vacuum energy is simply the sum of the vacuum energy for each orbit. In the case of a single orbit, the partition function looks like

$$Z = \sum_E e^{-\beta E} \,. \tag{2.343}$$

For a generic torus there can be angular potentials, but we have suppressed them since they will not influence the vacuum energy. Note that these values E are not directly the energy on the new spatial torus as there may have been a rescaling of the time direction. The vacuum energy of the orbifold theory $E_{\rm vac}^{\rm orbi}$ is simply the smallest such value of E. Now consider a twisted sector giving an arbitrary rescaling $U_i \to M_i U_i$ such that

$$\prod_{i=0}^{d-1} M_i = N \,. \tag{2.344}$$

This is needed as the scaling of the full torus must be equal to N if there is only one orbit. On such a torus, the vacuum contribution will be of the form

$$\begin{aligned} E_{\rm vac}^{\rm orbi}(M_i) &= -M_0 \frac{\varepsilon_{\rm vac} V_{d-1} \prod_{i>0} M_i}{M_1^d L_1^d} \left(1 + f(\mathbf{y}_1)\right) \\ &= -\frac{N}{M_1^d} \frac{\varepsilon_{\rm vac} V_{d-1}}{L_1^d} \left(1 + f(\mathbf{y}_1)\right) , \end{aligned} \tag{2.345}$$

where we used (2.344) and

$$\mathbf{y}_1 = \left(\frac{M_1 L_1}{M_2 L_2}, \ldots, \frac{M_1 L_1}{M_{d-1} L_{d-1}} \right) . \tag{2.346}$$

From (2.345) and using the monotonicity property of $f(\mathbf{y})$ under the increase of any of its arguments, it is clear that this expression is maximized for all $M_i = 1$ except for M_1. At first glance, it is not clear if increasing M_1 increases or decreases the energy as it appears both in the denominator and in $f(\mathbf{y})$ which change in opposite directions. However, one can alternatively write the vacuum energy as

$$E_{\text{vac}}(M_i) = -\frac{N}{M_2^d}\frac{\varepsilon_{\text{vac}} V_{d-1}}{L_2^d}\left(1 + f(\mathbf{y}_2)\right), \tag{2.347}$$

with

$$\mathbf{y}_2 = \left(\frac{M_2 L_2}{M_1 L_1}, \ldots, \frac{M_2 L_2}{M_{d-1} L_{d-1}}\right). \tag{2.348}$$

In this form, it is clear that $M_1 > 1$ would only give a less negative value to the free energy. We have thus showed that to get the minimal contribution, we need

$$M_0 = N, \quad M_i = 1 \ \ \forall\, i\,, \tag{2.349}$$

which then gives precisely the vacuum energy (2.342).

Although this might appear as good news for the orbifold theory to be a "nice" theory, it is very bad news for any chance of universality at large N. We have shown in the previous section that having $\tilde{f}(\mathbf{y}) = 0$ is a necessary condition for a universal free energy and an extended regime of the Cardy formula. Here, we see that the orbifold theory has $\tilde{f}(\mathbf{y}) = 0$ only if the seed theory does. The choice of seed becomes crucial to reproduce the phase structure of gravity. In fact, this result is not so surprising. In two dimensions, we could consider ourselves lucky that the S_N orbifold theory, which is a free theory, reproduces the phase

structure of Einstein gravity. It is only the strong coupling deformation of the orbifold theory that is dual to Einstein gravity so there is no a priori reason why one should have expected the orbifold theory to reproduce the phase structure of gravity. In higher dimensions, it appears that for a general seed, some form of coupling between the N CFTs must be introduced to force $\tilde{f}(\mathbf{y})$ to vanish. One might consider deforming the orbifold theory by some operator to achieve this effect. In particular, the existence of any exactly marginal deformations might allow reducing the Hagedorn density of light states to something compatible with Einstein gravity, as is proposed to occur in the D1-D5 duality. This could be directly connected to the vanishing of $\tilde{f}(\mathbf{y})$.

In the following subsection, we will show that choosing a seed theory with $\tilde{f}(\mathbf{y}) = 0$ both gives a theory that saturates the sparseness bound and reproduces the phase structure of gravity.

Universality for $\tilde{f}(\mathbf{y}) = 0$ and free energy at large N

If $\tilde{f}(\mathbf{y}) = 0$, we have $E_{\text{vac}} = -\varepsilon_{\text{vac}} V_{d-1}/L_1^d$ where L_1 is the length of the smallest cycle. Inserting this expression in (2.340), we obtain

$$\rho(E) \sim \exp\left(L_1(E - NE_{\text{vac}})\right) \tag{2.350}$$

for the growth coming from the specific twisted sector we previously considered. Note that the coefficient of the Hagedorn growth precisely saturates the bound on the light states given in (2.305) if we put the theory on the square torus. At the upper end of the range of validity of (2.340) where $E = -(d-1)NE_{\text{vac}}$, we

precisely recover the Cardy growth at the same energy. This indicates that the spectrum transitions sharply from Hagedorn to Cardy exactly where expected. However, we have only given a lower bound for the density of states as we only computed the contribution coming from a particular twisted sector. We will now show that for $\tilde{f}(\mathbf{y}) = 0$ it is also an upper bound. We will do so by computing the free energy and see that it precisely reproduces the universal behavior discussed in section 2.5.4. This implies that the density of low-lying states is bounded above by (2.350), which becomes both a lower and upper bound. This means that no other twisted sector can give a bigger contribution and the density of states is well-approximated by (2.350).

To compute the free energy at large N, we will follow a similar procedure as that in two dimensions [183]. The starting point is a combinatorics formula first introduced by Bantay [180]. Let G be a finitely generated group and Z a function on the finite index subgroups of G that takes values in a commutative ring and is constant on conjugacy classes of subgroups. We have the following identity

$$\sum_{N=0}^{\infty} \frac{p^N}{N!} \sum_{\varphi: G \to S_N} \prod_{\xi \in \mathcal{O}(\varphi)} Z(G_\xi) = \exp\left(\sum_{H<G} p^{[G:H]} \frac{Z(H)}{[G:H]} \right), \tag{2.351}$$

where φ is an homomorphism from G to S_N and H are subgroups of G with finite index given by $[G : H]$. In our case, Z will be the partition function and $G = \pi_1(\mathbb{T}^d) = \mathbb{Z}^d$. This group is abelian and the sum over homomorphisms φ is equivalent to the sum over commuting elements introduced earlier. The image of φ acts on N letters (momentarily this will be the N copies of the CFT) by the usual S_N action and its orbit is denoted by $\mathcal{O}(\varphi)$. The subgroup G_ξ consists

of those elements of g such that $\varphi(g)$ leaves ξ invariant. In fact, the left hand side is simply the generating function for the partition functions of the symmetric product orbifolds. It corresponds to

$$\mathcal{Z} = \sum_N p^N Z_N \,, \tag{2.352}$$

where Z_N is the partition function of $\mathcal{C}^{\otimes N}/S_N$ and thus the action of φ can be thought of as permuting the copies in $\mathcal{C}^{\otimes N}$. Just like in two dimensions, it is often more convenient to work with this generating function and to later find the coefficient of the term p^N to extract Z_N.

Bantay's formula equates the generating function to an exponential of a sum over new partition functions. This sum over partition functions really corresponds to a sum over new tori, and for a given index, the volume of the new tori will be the original volume times the index. Just as for $SL(2,\mathbb{Z})$, there is a very natural way to include all tori of a given index by using Hecke operators. Consider a torus to be described by the matrix U given in (2.250), which is upper triangular. Now consider the following set of matrices

$$\Omega_L = \left\{ \begin{bmatrix} a_0 & a_{01} & \cdots & a_{0,(d-2)} & a_{0,(d-1)} \\ 0 & a_1 & \cdots & a_{1,(d-2)} & a_{1,(d-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{d-2} & a_{(d-2),(d-1)} \\ 0 & 0 & \cdots & 0 & a_{d-1} \end{bmatrix} \;\middle|\; \prod_i a_i = L, 0 \le a_{j,i} < a_i \;\forall\; i,j \right\} \tag{2.353}$$

with L fixed. These matrices are elements of $GL(d,\mathbf{Z})$ and act on the lattice

vectors U_i defining the torus according to $\tilde{U} = HU$ with H an element of Ω_L. These new tori will have a volume L times larger than the original torus U. Consequently, the new lattice defined by the new torus is a sublattice H of $\mathbb{Z}^d$ and the index $[G : H]$ of H in $G = \mathbb{Z}^d$ is L. The purpose of these matrices is to parameterize the finite index subgroups of G so that we can write

$$\sum_{H<G} p^{[G:H]} \frac{Z(H)}{[G:H]} = \sum_{L>0} \frac{p^L}{L} \sum_{A\in\Omega_L} Z(AU)\,. \tag{2.354}$$

Fortunately, the right hand side can be rewritten in terms of Hecke operators for $SL(d,\mathbb{Z})$,

$$T_L Z(U) \equiv \sum_{A\in\Omega_L} Z(AU)\,, \tag{2.355}$$

which encapsulate the sum over different tori mentioned earlier. Note that the Hecke transform of Z is also an $SL(d,\mathbb{Z})$ modular invariant. Bantay's formula then becomes

$$\mathcal{Z}(U) = \exp\left(\sum_{L>0} \frac{p^L}{L} T_L Z(U)\right). \tag{2.356}$$

Because $T_L Z(U)$ is a function invariant under $SL(d,\mathbb{Z})$ [196], and it has a corresponding extensive free energy, its asymptotic growth is also given by the higher-dimensional Cardy formula. To see this directly, notice that $T_L Z(U)$ is a sum over partition functions of different tori. Each of these obeys the higher-dimensional Cardy formula, although the explicit dependence on the volume of the torus in our higher-dimensional Cardy formula may seem confusing. Note however that at asymptotically large energies we have $E \propto V_{d-1}^{-1/(d-1)}$, so the volume of the torus cancels out and the formula can be written in terms of a dimensionless energy.

Thus, there is no confusion as to "which volume" enters into the Cardy formula for $T_L Z(U)$. In fact, the situation is even better. The gap between the first excited state and the vacuum grows with L indicating that at large L, the Cardy formula will become a good estimate for the Hecke transformed partition function.

We are now ready to estimate the free energy. Let us take a rectangular d-torus with sides $\beta, L_1, ..., L_{d-1}$, i.e

$$U = \begin{pmatrix} \beta & 0 & \cdots & 0 & 0 \\ 0 & L_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & L_{d-2} & 0 \\ 0 & 0 & \cdots & 0 & L_{d-1} \end{pmatrix}, \tag{2.357}$$

and let us assume L_1 is the smallest spatial cycle. Writing $\tilde{p} = pe^{\beta E_{\rm vac}}$,

$$\begin{aligned} \mathcal{Z} &= \exp\left(\sum_{L>0} \frac{\tilde{p}^L}{L} + \sum_{L>0} \frac{\tilde{p}^L}{L} \sum_{E>0} \tilde{\rho}_{T_L}(E) e^{-\beta E}\right) \\ &= \left(\sum_{K=0}^{\infty} \tilde{p}^K\right) \exp\left(\sum_{L>0} \frac{\tilde{p}^L}{L} \sum_{E>0} \tilde{\rho}_{T_L}(E) e^{-\beta E}\right), \end{aligned} \tag{2.358}$$

where we have defined $\tilde{\rho}_{T_L}(E)$ such that

$$e^{L\beta E_{\rm vac}} T_L Z(U) = 1 + \sum_{E>0} \tilde{\rho}_{T_L}(E) e^{-\beta E} \tag{2.359}$$

Using the Cardy formula, the sum over energies in (2.358) becomes

$$\sum_{E>0} e^{\left(daL^{\frac{1}{d}}(E+E_{\rm vac}L)^{\frac{d-1}{d}}\right)} e^{-\beta E} \quad \sim \quad \exp\left(L|E_{\rm vac}|\left(\frac{L_1^d}{\beta^{d-1}}-\beta\right)\right), \quad (2.360)$$

where we assumed L_1 to be the smallest cycle and used (2.317) as well as

$$E_{\rm vac} = \frac{-\varepsilon_{\rm vac} V_{d-1}}{L_1^d}\,. \quad (2.361)$$

The saddle point value for E is

$$E_s = |E_{\rm vac}|L\left(1+\frac{(d-1)L_1^d}{\beta^d}\right), \quad (2.362)$$

which will be large for large L. This justifies the use of the Cardy formula. The terms with low E will of course not be in the Cardy regime but these will only give a subleading contribution. Overall, the error on the each term in the sum over L will be of order $e^{-uL/\beta^{d-1}}$ for some positive order one number u that is theory dependent. Plugging (2.360) into (2.358) we get

$$\begin{aligned}
\mathcal{Z} &= \left(\sum_{K=0}^{\infty}\tilde{p}^K\right)\exp\left(\sum_{L>0}\frac{1}{L}\left(\tilde{p}\exp\left(|E_{\rm vac}|\left(\frac{L_1^d}{\beta^{d-1}}-\beta\right)\right)\right)^L\right) \\
&= \left(\sum_{K=0}^{\infty}\tilde{p}^K\right)\exp\left(-\log\left(1-\tilde{p}e^{|E_{\rm vac}|\beta(L_1^d/\beta^d-1)}\right)\right) \\
&= \left(\sum_{K=0}^{\infty}\tilde{p}^K\right)\frac{1}{1-\tilde{p}e^{|E_{\rm vac}|\beta(L_1^d/\beta^d-1)}}\,. \quad (2.363)
\end{aligned}$$

We can now extract the free energy. Note that because the vacuum energy is negative and proportional to N, the partition function diverges as $N\to\infty$ so we

need to consider the shifted partition function and shifted free energy

$$\begin{aligned} \tilde{Z} &\equiv e^{E_{\rm vac}\beta} Z\,, \\ \tilde{F} &\equiv -\frac{\log \tilde{Z}}{\beta}\,. \end{aligned} \tag{2.364}$$

The shifted partition function will then simply be the term $\tilde{p}^N$ in (2.363), which is given by

$$\tilde{Z}_N = \frac{\exp\left((N+1)|E_{\rm vac}|\beta\left(\frac{L_1^d}{\beta^d}-1\right)\right)-1}{\exp\left(|E_{\rm vac}|\beta\left(\frac{L_1^d}{\beta^d}-1\right)\right)-1}\,. \tag{2.365}$$

The free energy as $N\to\infty$ for $\beta < L_1$ is thus

$$\tilde{F}_N(U) = -N|E_{\rm vac}|\left(\frac{L_1^d}{\beta^d}-1\right)\,. \tag{2.366}$$

For $\beta > L_1$, we get

$$\tilde{F}_N(U) = \frac{1}{\beta}\log\left(1-\exp\left(|E_{\rm vac}|\beta\left(\frac{L_1^d}{\beta^d}-1\right)\right)\right) + F_{\rm cor}(\beta)\,, \tag{2.367}$$

where the $F_{\rm cor}(\beta)$ corresponds to another $\mathcal{O}(1)$ contribution coming from subleading corrections to the saddle point as well as the low energy contributions. The free energy thus has a phase transition at $\beta = L_1$ and goes from being $\mathcal{O}(1)$ to $\mathcal{O}(N)$. This precisely matches the phase structure of the bulk gravitational theory.

Modular invariance is not manifest in the shifted free energy above. In order

to recover it, we consider the quantity

$$\mathscr{F}(U) = \lim_{N\to\infty} \frac{1}{N} F_N(U)\,, \tag{2.368}$$

where $F_N(U)$ is the unshifted free energy and $\mathscr{F}(U) = \mathscr{F}(\beta, L_1, ..., L_{d-1})$. Using the results obtained above,

$$\mathscr{F}(U) = \begin{cases} -\frac{\varepsilon_{\rm vac} V_{d-1}}{\beta^d} & \beta < L_1 \\ -\frac{\varepsilon_{\rm vac} V_{d-1}}{L_1^d} & \beta > L_1 \end{cases}\,, \tag{2.369}$$

where L_1 is the smallest cycle. The free energy is a modular covariant quantity which transforms under the S transform of $SL(d,\mathbb{Z})$ as

$$\mathscr{F}(\beta, L_1, ..., L_{d-1}) = \frac{L_1}{\beta} \mathscr{F}(L_1, ..., L_{d-1}, \beta). \tag{2.370}$$

Upon checking this transformation rule for (2.369), we see that in both regimes the free energy transforms as expected.

2.5.6 Discussion

In this paper we have studied conformal field theories in dimensions $d > 2$ compactified on tori. The main goal was to explore the implications of the assumed invariance under the $SL(d,\mathbb{Z})$ modular group and see what additional constraints on the spectrum would reproduce the phase diagram of gravity in anti-de Sitter space. We have uncovered both similarities and differences with the two-dimensional case. We have presumably only scratched the surface of this

interesting subject and many issues and open questions remain, some of which we list below.

Modular invariance

The modular group $SL(d,\mathbb{Z})$ consists of the large diffeomorphisms (i.e. not continuously connected to the identity element) which map a d-dimensional torus to itself. In two dimensions, there are well-known systems, such as the chiral fermion, whose partition function is not modular invariant. However, such theories have gravitational anomalies and can therefore a priori not be consistently defined on arbitrary manifolds. Moreover, when such theories appear in nature, as in the edge modes in the quantum Hall effect, the relevant anomalies are canceled due to an anomaly inflow mechanism which crucially relies on the existence of a higher-dimensional system to which the theory is coupled (for a higher-dimensional version of this statement see e.g. [197]). We are not aware of a local and unitary conformal field theory which is free of local gravitational anomalies and not modular invariant. But modular invariance is weaker than the absence of local gravitational anomalies. There are many modular invariant CFTs with $c_L - c_R \neq 0$ which have gravitational anomalies, while modular invariance only implies that $c_L - c_R$ must be an integer multiple of 24. It would be interesting to explore the generalizations of these statements to higher dimensions.

Another approach to using modular invariance to learn about conformal field theories on tori is to consider bounds coming from the fixed points of $SL(d,\mathbb{Z})$. This would be a generalization of the "modular bootstrap" [198, 199, 200, 201, 202, 203, 204, 205] to higher dimensions. This is valid for general conformal field

theories, and taking a large-N limit may give insight into holographic theories.

State-operator correspondence

The usual arguments for the state-operator correspondence in conformal field theory rely on radial quantization and apply to the theory on the spatial sphere S^{d-1} times time. The local operators obtained in this way can be inserted on other manifolds as well but the one-to-one correspondence with states in the Hilbert space no longer applies. The main problem in applying radial quantization to the torus is that, as opposed to spheres, one can not smoothly shrink a torus of dimension larger than one to a point. Stated more precisely, the metric $ds^2 = dr^2 + r^2 d\Omega^2$ is not smooth at $r = 0$ unless Ω is the round unit sphere.

One cannot even apply the standard radial quantization argument to the conformal field theory on $S^1 \times \mathbb{R}^{d-2}$ times time. At $r = 0$, the metric $ds^2 = dr^2 + r^2 d\varphi^2 + r^2 dx_i dx_i$ looks like a singular $\mathbb{R}^{d-2}$-dimensional plane, suggesting that some sort of surface operators might be relevant. That such operators are generically needed can for example be seen using the orbifold theories we studied in this paper. Orbifold theories can be thought of as theories with a discrete gauge symmetry, and in case the theory lives on $S^1 \times \mathbb{R}^{d-2}$ we should include twisted sectors which involve twisted boundary conditions when going around the S^1. These twisted boundary conditions can be detected by a Wilson line operator for the discrete gauge field around the S^1. To create a non-trivial expectation value for the Wilson line operator, we need an operator which creates non-contractible loops, and for this we need an operator localized along a $(d-2)$-dimensional surface. One can think of such operators as a higher-dimensional generalization of

the 't Hooft line operators. A local operator in $d > 2$ is unable to generate a non-trivial vev for the discrete Wilson line operator and can therefore not create twisted sector states. Surface operators of dimension $d-2$ which create twisted boundary conditions also feature prominently in the replica trick computations of entanglement entropy in dimensions $d > 2$; they are the generalized twist fields associated to the boundary of the entangling area.

If $(d-2)$-dimensional surface operators are the right operators for the theory on $S^1 \times \mathbb{R}^{d-2}$, it is plausible they are also relevant for CFT's on tori. One can for example consider the surface operators dual to periodic field configurations on $\mathbb{R}^{d-2}$, but it is not clear the resulting surface operator will have the right periodicity as well. Alternatively, one can study the Euclidean theory on an annulus times $\mathbb{T}^{d-2}$, with the annulus having inner radius R_1 and outer radius R_2. The Euclidean path integral in principle provides a map from states on the torus $S^1_{R_1} \times \mathbb{T}^{d-2}$ to $S^1_{R_2} \times \mathbb{T}^{d-2}$, and by taking the limit $R_1 \to 0$ one can imagine obtaining singular boundary conditions for a surface operator localized along a $(d-2)$-torus.

Clearly, more work is required to understand whether the above construction provides a useful version of the state-operator correspondence for field theories on tori, and if it does, what a useful basis for the space of surface operators could possibly be. There seems to be a significant overcounting, as one can construct a surface operator for any choice of state on the torus and for any choice of one-cycle on the torus. Currently, we do not even have a compelling compact Euclidean path integral representation of the ground state of the theory on the torus.

It might also be interesting to explore the state-operator correspondence from

an AdS/CFT point of view. One would then need to glue Euclidean caps to the Lorentzian solutions discussed in section 2.5.3. Since the Lorentzian solutions require a choice of one-cycle which is smoothly being contracted in the interior, a similar choice will be needed for the Euclidean caps, leading apparently once more to the same overcounting as we observed above. It would still be interesting to construct the explicit form of the geometry where a Euclidean cap without the insertion of surface operators is smoothly glued to the Lorentzian AdS solutions. If such solutions could be found, its boundary geometry would provide a Euclidean path integral description of the ground state of the corresponding CFT, at least in the large N and strong coupling approximation.

Defining the orbifold theory

In section 2.5.5, we defined a prescription to compute the partition function of the orbifold theory. This prescription describes both the Hilbert space and the spectrum of the Hamiltonian on the torus. In two dimensions, the orbifolding prescription also fully describes the procedure to compute arbitrary correlation functions of (un)twisted sector local operators, at least in principle. In higher dimensions, because of the lack of a precise state-operator correspondence, it is not clear wether we have really fully specified a theory. For that, we need to determine the full set of correlation functions and hence know the set of operators in the theory. It is clear that all untwisted sector correlation functions make sense in the orbifold theory so all local correlation functions are well-defined and calculable. Furthermore, the theory possesses a stress tensor as the stress tensor is always in the untwisted sector. Nevertheless, the questions touching the twisted

sector states and/or line operators is much more obscure and it would be very interesting to understand the extent to which the orbifolding prescription fully determines these.

One way orbifold theories in higher dimensions can potentially appear (and therefore inherit a natural definition) are as discrete gauge theories that arise in the infrared limit of a gauge theory with spontaneously broken continuous gauge symmetry (e.g. $SU(N) \to S_N$). This would also explain how to couple the theory to other manifolds, an issue we turn to in the next section.

Orbifold theories on other manifolds

The orbifold theories we studied are most easily defined on tori. However, if we have fully defined a theory we should be able to put it on any manifold. Viewing them as theories with a discrete gauge group also provides a prescription for the sum over twisted sectors when computing the path integral for other manifolds. The sum over twisted sectors is the same as the sum over the space of flat connections modulo an overall conjugation, and for a manifold M this space is $\mathrm{Hom}(\pi_1(M), G)/G$. But even for flat space, where no sum over twisted sectors needs to be taken, there are still signs of the discrete gauge symmetry. In particular, one can consider surface operators which create twisted sector states even on the plane, and their correlation functions contain interesting new information. Such operators naturally arise in the context of Renyi entropy calculation in higher dimensions [206, 207].

Outlook

We have only begun to explore the properties of modular-invariant field theories on tori and their role in AdS/CFT. The interesting relations between the form of the ground state energy, universal free energy at high-temperature, sparseness conditions on the spectrum and vacuum dominance in the partition function beg for a deeper understanding. Is there a more precise relation between the low- and high-energy spectrum that can be rigorously established? Can subleading corrections be systematically analyzed? How much of the rich structure in $d = 2$ and the mathematics of $SL(2, \mathbb{Z})$ can be carried over to $d > 2$? Does all this shed any new light on which theories can have weakly coupled gravitational duals?

Acknowledgements

The authors would like to acknowledge helpful conversations with Tom Hartman, Diego Hofman, Christoph Keller, and Alex Maloney. AB is supported by the Foundation for Fundamental Research on Matter (FOM). JK is supported by the Delta ITP consortium, a program of the Netherlands Organisation for Scientific Research (NWO) that is funded by the Dutch Ministry of Education, Culture and Science (OCW). BM and ES are supported by NSF Grant PHY13-16748. MS is supported by the Yzurdiaga Chair at UCSB and the Simons Qubit Collaboration. MS would like to thank the Stanford Institute of Theoretical Physics for hospitality during part of the completion of this work.

2.5.7 Appendix A: Alternate proof without assuming $\tilde{f}(\mathbf{y}) = 0$

In this section we will try to generalize the proof in section 2.5.4 to the case where we make no assumptions on the form of the vacuum energy. To illustrate the point, we will work in three spacetime dimensions and make the spatial manifold explicit. We will again be using a proof like that of [156], but this time we will take $N \to \infty$ from the start.[18]

Consider a rectangular three-torus with side lengths β, L_1, and L_2 with $\beta < L_1 < L_2$. We have the relations

$$Z(\beta)_{L_1\times L_2} - Z(L_1)_{\beta\times L_2} = Z(\beta)_{L_1\times L_2} - Z(L_2)_{\beta\times L_1} = 0 \implies \tag{2.371}$$

$$\left(\sum_L e^{-\beta E_{L_1\times L_2}} - \sum_L e^{-L_1 E_{\beta\times L_2}}\right) + \left(\sum_H e^{-\beta E_{L_1\times L_2}} - \sum_H e^{-L_1 E_{\beta\times L_2}}\right) = 0\,, \tag{2.372}$$

$$\left(\sum_L e^{-\beta E_{L_1\times L_2}} - \sum_L e^{-L_2 E_{\beta\times L_1}}\right) + \left(\sum_H e^{-\beta E_{L_1\times L_2}} - \sum_H e^{-L_2 E_{\beta\times L_1}}\right) = 0\,. \tag{2.373}$$

Notice that light states L and heavy states H are playing triple duty, since the spatial background changes in the different quantizations. In any given quantization, the states L refer to negative energy states that scale with a positive power of N while H refers to positive energy states that scale with a positive power of N. We eliminate the consideration of states with $\mathcal{O}(1)$ energies by bounding their density of states so that their contribution is $\mathcal{O}(1)$ and therefore subleading.

[18] Thanks to Tom Hartman for discussions about this simpler form of proof.

We now assume that for $\beta < L_1$ and $\beta < L_2$, we have

$$\sum_L e^{-L_1 E_{\beta\times L_2}} \gg \sum_L e^{-\beta E_{L_1\times L_2}}, \qquad \sum_H e^{-L_1 E_{\beta\times L_2}} \ll \sum_H e^{-\beta E_{L_1\times L_2}}, \tag{2.374}$$

$$\sum_L e^{-L_2 E_{\beta\times L_1}} \gg \sum_L e^{-\beta E_{L_1\times L_2}}, \qquad \sum_H e^{-L_2 E_{\beta\times L_1}} \ll \sum_H e^{-\beta E_{L_1\times L_2}}\,. \tag{2.375}$$

These inequalities can be proven to be true in two spacetime dimensions and for the special torus in a general number of dimensions. In fact, it is what makes a proof like that of [156] work.

Using these inequalities, we can approximate the above equalities as

$$\sum_L e^{-L_1 E_{\beta\times L_2}} \approx \sum_H e^{-\beta E_{L_1\times L_2}}, \qquad \sum_L e^{-L_2 E_{\beta\times L_1}} \approx \sum_H e^{-\beta E_{L_1\times L_2}}. \tag{2.376}$$

Then we can use $Z_H(L_1)_{\beta\times L_2} \ll Z_H(\beta)_{L_1\times L_2} \approx Z_L(L_1)_{\beta\times L_2}$ and $Z_H(L_2)_{\beta\times L_1} \ll Z_H(\beta)_{L_1\times L_2} \approx Z_L(L_2)_{\beta\times L_1}$ to approximate the partition function in the L_1 and L_2 channels as

$$Z(L_1)_{\beta\times L_2} = Z_L(L_1)_{\beta\times L_2} + Z_H(L_1)_{\beta\times L_2} \approx Z_L(L_1)_{\beta\times L_2}\,, \tag{2.377}$$

$$Z(L_2)_{\beta\times L_1} = Z_L(L_2)_{\beta\times L_1} + Z_H(L_2)_{\beta\times L_1} \approx Z_L(L_2)_{\beta\times L_1}\,. \tag{2.378}$$

We see that under the assumptions (2.374) and (2.375), the partition function is vacuum dominated in the L_1 and L_2 channels if and only if

$$\rho(E_{L_1\times L_2} < 0) \lesssim e^{L_1 (E - E_{\text{vac}})_{L_1\times L_2}}. \tag{2.379}$$

As explained in section 2.5.4 this is necessary and sufficient for a universal free

energy at all temperatures on an arbitrary spatial torus.

In general dimension, the sufficient conditions for a universal free energy are the $d-1$ inequalities that generalize (2.375) and a sparse light spectrum:

$$\rho(\Delta) \lesssim \exp\left(L_{\min}\Delta\right) , \tag{2.380}$$

where $L_{\min}$ is the minimum cycle size of the spatial torus.

2.5.8 Appendix B: Microcanonical density of states

The results derived in the main text are phrased in terms of the canonical partition function $Z(\beta)$. In general such results do not immediately translate into statements about the microcanonical density of states. However, as discussed carefully for two dimensions in [156], the limit $N \to \infty$ is a good thermodynamic limit which allows us to conclude $\rho(\langle E\rangle) \approx e^{S(\langle E\rangle)}$ for $\langle E\rangle = -\partial_\beta \log Z(\beta)$. Large N suppresses the fluctuations in $\langle E\rangle$ and unambiguously defines an energy $E \equiv \langle E\rangle$. The arguments of [156] carry over straightforwardly and imply that the Cardy density of states has an extended range of validity that holds down to $E = -(d-1)E_{\text{vac}}$, which is the energy corresponding to $\beta = L_1$. Instead of repeating those arguments we will give an alternative way of understanding the thermodynamic limit which gives additional intuition. In the next sections we will evaluate the inverse Laplace transform connecting the canonical partition function to the microcanonical density of states in several examples. The existence of a stable saddle point is the statement of an equivalence between the two ensembles.

In section 2.5.8 we will consider the case of two-dimensional CFT and use the

fact that the partition function is dominated by the light states [156]. In section 2.5.8 we will assume $\tilde{f}(\mathbf{y}) = 0$, in which case the special torus allows us to get away with only bounding the density of light states, just as in two dimensions and as we saw in the main text. Finally, in section 2.5.8 we will consider an extension to angular momentum. Here, to extend the regime of validity we will have to bound the density of states for the entire spectrum, again as we saw in the main text.

$d = 2$

We will begin by considering the case of two-dimensional conformal field theories, treated in [156]. Here we will directly evaluate the inverse Laplace transform connecting the canonical partition function to the density of states.

We begin with the expression for the degeneracy

$$\rho(h_s, \bar{h}_s) = \int_{i\alpha-\infty}^{i\alpha+\infty} d\tau \int_{-i\alpha-\infty}^{-i\alpha+\infty} d\bar{\tau}\, I(\tau, \bar{\tau})\, \tilde{Z}(-1/\tau, -1/\bar{\tau})\,, \tag{2.381}$$

for $\alpha > 0$, where

$$I(\tau, \bar{\tau}) \equiv e^{-2\pi i\tau(h_s - c/24)} e^{2\pi i\bar{\tau}(\bar{h}_s - \bar{c}/24)} \times\ e^{-2\pi i/\tau(-c/24)} e^{2\pi i/\bar{\tau}(-\bar{c}/24)}\ . \tag{2.382}$$

Evaluating the integral using the saddle point approximation for large h_s and $\bar{h}_s$ requires solving the saddle equations $I^{(1,0)}(\tau, \bar{\tau}) = I^{(0,1)}(\tau, \bar{\tau}) = 0$, which gives the dominant saddle

$$\tau_s = \tau_1^s + i\tau_2^s = +i\sqrt{\frac{c}{24h_s - c}}\,, \tag{2.383}$$

$$\bar{\tau}_s = \tau_1^s - i\tau_2^s = -i\sqrt{\frac{\bar{c}}{24\bar{h}_s - \bar{c}}}\,. \tag{2.384}$$

Evaluating the integrand on this saddle gives the entropy

$$S = \log \rho(h_s, \bar{h}_s) = 2\pi \left(\sqrt{\frac{c}{6}\left(h_s - \frac{c}{24}\right)} + \sqrt{\frac{\bar{c}}{6}\left(\bar{h}_s - \frac{\bar{c}}{24}\right)} \right). \tag{2.385}$$

To ensure the saddle point approximation is justified, we have to check that $\tilde{Z}$ does not make big contributions on the saddle:

$$\tilde{Z}(-1/\tau_s, -1/\bar{\tau}_s) = \sum_{h,\bar{h}} \rho(h,\bar{h}) \exp\left(\frac{-2\pi h}{\sqrt{\frac{c}{24h_s - c}}} - \frac{2\pi \bar{h}}{\sqrt{\frac{\bar{c}}{24\bar{h}_s - \bar{c}}}} \right). \tag{2.386}$$

As $h_s \to \infty$ and $\bar{h}_s \to \infty$ with c finite, all terms except for the vacuum contribution are infinitely exponentially suppressed, justifying our saddle point approximation. This is the ordinary Cardy formula.

Now let us consider the limit $c \to \infty$ with $h_s = mc$. For simplicity we set $c = \bar{c}$ and $\tau_1 = 0$. So we have the canonical partition function at inverse temperature $\beta = \tau_2/(2\pi)$. We will have the same saddle as before but need to check again that $\tilde{Z}$ does not give a big contribution on the saddle. If we take $m \to \infty$, then again all terms except the vacuum contribution are infinitely exponentially suppressed and our saddle is justified. But now we want to see how small we can make m. We will use the fact that $Z(\beta)$ is dominated by the light states as long as $\beta > 2\pi$. This means that $\tilde{Z}(\beta)$ is also dominated by the light states. We can therefore write

$$\tilde{Z}(4\pi^2/\beta_s) \approx \sum_{\Delta \le c/12 + \epsilon} \rho(\Delta) \exp\left(\frac{2\pi\Delta}{\sqrt{\frac{c}{12\Delta_s - c}}} \right) \tag{2.387}$$

for $\Delta = h + \bar{h}$. We need all terms on the right-hand-side to contribute exponential suppressions, except for the identity operator which will contribute $+1$. To push the validity of the saddle down to $\Delta_s = c/6$, which is the result expected from gravity, we need to bound the degeneracy as

$$\rho(\Delta \leq c/12 + \epsilon) \lesssim \exp(2\pi\Delta) = \exp(2\pi(E + c/12)) \,. \tag{2.388}$$

This is the same bound on the light states as in [156].

$d > 2$

In this case, we consider the high-temperature/low-temperature duality on the special torus, for which $Z(\beta) = Z(L^d/\beta^{d-1})$. We have

$$\rho(E_s) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} d\beta \, Z(\beta) \, e^{\beta E_s} \tag{2.389}$$

for $\alpha > 0$. Performing a modular transformation and multiplying and dividing by a common factor gives (omitting the integration limits and $1/2\pi i$)

$$\rho(E_s) = \int d\beta \, \left(e^{-\varepsilon_{\text{vac}} V_{d-1}/\beta^{d-1}} Z(L^d/\beta^{d-1}) \right) e^{\varepsilon_{\text{vac}} V_{d-1}/\beta^{d-1} + \beta E_s} \,. \tag{2.390}$$

We will hold off on defining V_{d-1} for the moment, which will actually be defined to be independent of β. At large N with E_s scaling as a positive power of N the saddle point (ignoring the term in parentheses) occurs at

$$\beta_s = \left(\frac{(d-1)\varepsilon_{\text{vac}} V_{d-1}}{E_s} \right)^{\frac{1}{d}} , \tag{2.391}$$

which gives an on-shell entropy of

$$\rho(E_s) = \exp\left(\frac{d}{(d-1)^{\frac{d-1}{d}}}(\varepsilon_{\rm vac}V_{d-1})^{\frac{1}{d}}E_s^{\frac{d-1}{d}}\right) \tag{2.392}$$

To make sure the saddle is controlled, we again want the term in parentheses to contribute as $1 + e^{-(\dots)}$ on the saddle. To show this, we will use the fact from the main text that for $Z(L^d/\beta^{d-1})$ is dominated by the contribution of the light states for $L > \beta$. This lets us write the term in parentheses on the saddle as

$$\sum_{E<\epsilon} \exp\left(-(L^d E + \varepsilon_{\rm vac}V_{d-1})\left(\frac{E_s}{(d-1)\varepsilon_{\rm vac}V_{d-1}}\right)^{\frac{d-1}{d}}\right). \tag{2.393}$$

We now want to define $V_{d-1} = V_{\mathcal{M}_0} = L\cdots L^{d-1}/\beta_s^{d-2}$ as the volume of the special torus. Since β_s depends on V_{d-1}, this is an equation that can easily be solved for V_{d-1} in terms of only $\varepsilon_{\rm vac}$ and E_s, but all we need to know is that it gives the volume of the special torus on the saddle. Now using our assumption that the subextensive corrections to the vacuum energy vanish, we see that the vacuum state contributes as $+1$. To approach the square torus as in the main text we want to push E_s down to $-(d-1)E_{\rm vac}$, which will require bounding the density of light states as

$$\rho(E) \lesssim \exp(L(E - E_{\rm vac})), \qquad E \leq -(d-1)\,E_{\rm vac}\,, \tag{2.394}$$

where the energies are taken to be on a square torus. This is the same bound as we saw in the main text. At this point we can perform a similar bootstrapping procedure to obtain this density of states on an arbitrary spatial torus and at

arbitrarily higher energies.

Cardy extension with angular momentum on $\mathbb{T}^2 \times \mathbb{R}^{d-2}$

We will show in this section that similar manipulations can be performed once angular momentum is included. In particular, assuming sparseness on the low-lying spectrum, we can extend the generalized Cardy formula with angular momentum to include the entire range

$$J^2 < (E - E_{\rm vac})(E + (d-1)E_{\rm vac})\,. \tag{2.395}$$

Note that this has the correct limits. For $d = 2$ we recover $E_L E_R > c^2/576$, and for $J = 0$, $d > 2$ we get $E > -(d-1)E_{\rm vac}$.

Before we perform our CFT analysis, we should analyze the phase structure of gravity with the appropriate boundary conditions. We are introducing a chemical potential for angular momentum, which corresponds to adding a twist in the periodicity of Euclidean time. The solutions are the same as in the main text, but with angular velocity added. The Poincaré patch and soliton geometries can be written as before except with the new identification $t \sim t + i\beta + \theta$, while the black brane is written as

$$ds^2 = \left((r_h/r)^d u_\mu u_\nu + r^2 \eta_{\mu\nu}\right) dx^\mu dx^\nu + \frac{dr^2}{r^2\left(1 - (r_h/r)^d\right)}\,, \tag{2.396}$$

$$u_\mu = \left(\frac{-1}{\sqrt{1-a^2}}\,, \frac{a}{\sqrt{1-a^2}}, \vec{0}\right)\,. \tag{2.397}$$

The free energies of the solutions are given by

$$F_{\rm bb} = -\frac{r_h^d L L_\infty^{d-2}}{16\pi G}\,, \qquad F_{\rm sol} = -\frac{r_0^d L L_\infty^{d-2}}{16\pi G}\,, \qquad F_{\rm pp} = 0\,, \tag{2.398}$$

with

$$r_h = \frac{4\pi}{d}\sqrt{\frac{1}{\beta^2+\theta^2}}\,, \qquad r_0 = \frac{4\pi}{dL}\,. \tag{2.399}$$

The energy and the angular momentum of the black brane are given by the usual thermodynamic relations in terms of a Euclidean partition function $\mathcal{Z}(\beta,\theta) = \Sigma\; e^{-\beta H + i\theta J} = \Sigma\; e^{-\beta(H+aJ)}$:

$$E = -\frac{\partial}{\partial\beta}\bigg|_\theta \log Z = \frac{r_h^d L L_\infty^{d-2}}{16\pi G}\frac{d-1+a^2}{1-a^2}\,, \tag{2.400}$$

$$J = -i\frac{\partial}{\partial\theta}\bigg|_\beta \log Z = \frac{r_h^d L L_\infty^{d-2}}{16\pi G}\frac{da}{(1-a^2)}\,. \tag{2.401}$$

From the expressions for the free energies, we see that the soliton dominates the ensemble for $r_0 > r_h$. At the phase transition $r_h = \frac{4\pi}{dL}$, the energy and angular momentum are related by

$$J^2 = (E - E_{\rm vac})(E + (d-1)E_{\rm vac})\,. \tag{2.402}$$

We now turn to our CFT analysis. The canonical partition function at finite temperature and angular velocity is defined as

$$Z(\tau,\bar\tau) = {\rm Tr}\left(e^{2\pi i\tau E_R} e^{-2\pi i\bar\tau E_L}\right),$$

$$E_R + E_L = E\,, \qquad E_R - E_L = J\,,$$

where $\tau = re^{i\varphi}$ is the modular parameter whose imaginary part acts as the inverse temperature, and the real part acts as the chemical potential for angular momentum. We have only turned on a single angular momentum generator. The microcanonical density of states is given by the usual inverse Laplace transform (up to subleading Jacobian factors which we ignore):

$$\rho(E_s, J_s) = \int dr d\varphi \; Z(r,\varphi) \; \exp\left[-\pi i r e^{i\varphi}(E_s + J_s) + \pi i r e^{-i\varphi}(E_s - J_s)\right]. \quad (2.403)$$

For simplicity, we will work in the special case of $\mathbb{T}^2 \times \mathbb{T}^{d-2}_\infty$ and consider the angular momentum to be along the spatial cycle of the $\mathbb{T}^2$. On this background, modular invariance gives

$$\log\, Z(r,\varphi) \approx r^{2-d} \log\, Z(-r^{-1}, -\varphi)\,. \quad (2.404)$$

As before we define a shifted partition function as

$$\tilde{Z}(r,\varphi) \equiv \mathrm{Tr}\, \exp\left[\pi i r e^{i\varphi}(E_s + J_s - E_{\rm vac}) - \pi i r e^{-i\varphi}(E_s - J_s - E_{\rm vac})\right]. \quad (2.405)$$

Using the above we write the density of states as

$$\rho(E_s, J_s) = \int dr \; d\varphi \; \tilde{Z}\left(-\frac{1}{r}, -\varphi\right)^{r^{2-d}} \exp\left[\frac{-\pi i E_{\rm vac}}{r^{d-1}e^{i\varphi}} + \frac{\pi i E_{\rm vac}}{r^{d-1}e^{-i\varphi}}\right] \times \exp\left[-\pi i r e^{i\varphi}(E_s + J_s) + \pi i r e^{-i\varphi}(E_s - J_s)\right]. \quad (2.406)$$

At large N, we can approximate the above integral by its saddle-point value, which

gives

$$\rho(E_s, J_s) = \exp\Bigg[\pi\sqrt{d}\left(\frac{2}{(d-1)^{d-1}}\right)^{\frac{1}{d}}\left(\sqrt{d^2E_s^2 - 4(d-1)J_s^2} - (d-2)E_s\right)^{\frac{d-2}{2d}} \times \left(\sqrt{d^2E_s^2 - 4(d-1)J_s^2} + dE_s\right)^{1/2}(-E_{\rm vac})^{\frac{1}{d}}\Bigg]. \tag{2.407}$$

This is the higher-dimensional Cardy formula with angular momentum. To ensure that our saddle is controlled and this formula is valid, we need to check that the neglected piece $\tilde{Z}$ is not large on the saddle. By definition

$$\tilde{Z}\left(-\frac{1}{r_s}, -\varphi_s\right) = \int_{\rm light} dEdJ\ \rho(E,J)\ \exp\left[-\frac{\pi i}{r_s}\left(e^{-i\varphi_s}(\Delta+J) - e^{i\varphi_s}(\Delta-J)\right)\right] + \int_{\rm heavy} dEdJ\ \rho_{\rm Cardy}(E,J)\ \exp\left[-\frac{\pi i}{r_s}\left(e^{-i\varphi_s}(\Delta+J) - e^{i\varphi_s}(\Delta-J)\right)\right], \tag{2.408}$$

where we have used $\Delta = E - E_{\rm vac}$. We would like to find and maximize the range in the spectrum where the heavy states lie. The first line stands for the contribution of light states and is $\mathcal{O}(1)$ as long as the density of light states obeys a Hagedorn bound. The second line is small if

$$\log\rho_{\rm Cardy} - \frac{\pi i}{r_s}\left(e^{-i\varphi_s}(\Delta+J) - e^{i\varphi_s}(\Delta-J)\right) < 0\,. \tag{2.409}$$

Let us denote the left hand side of this expression by $\mathcal{T}(E_s, J_s, E, J)$. The dependence of $\mathcal{T}$ on E_s and J_s comes through r_s and φ_s. Using the values of the saddle and the Cardy formula gives a messy expression for $\mathcal{T}(E_s, J_s, E, J)$.

We would like to find the region in the E, J plane where $\mathcal{T}(E_s, J_s, E, J) < 0$. Note that since $\mathcal{T}(E_s, J_s, E, J)$ is also a function of E_s and J_s, this region will depend on the values of E_s and J_s. This means we need to find the values of E_s and J_s for which the region in the E, J plane is maximized. To guarantee that $\mathcal{T}$ is less than zero in a given region in the E, J plane, it will be sufficient to show that the maximum value of $\mathcal{T}$ with respect to E and J is smaller than zero in that region. Saturating this bound will give us the extended range of validity of the Cardy formula. In other words, maximization of $\mathcal{T}$ with respect to E and J will give us E and J in terms of E_s and J_s. Then demanding the maximum of $\mathcal{T}$ to be smaller than zero will give a constraint on how small we can make E_s and J_s.

Let's see this in the simpler case of $d = 2$ and $J = 0$:

$$\mathcal{T}(\Delta, \Delta_s) = 2\pi\sqrt{\frac{c}{3}\left(\Delta - \frac{c}{12}\right)} - \frac{2\pi\Delta}{\sqrt{\frac{c}{12\Delta_s - c}}}\,. \tag{2.410}$$

Extremizing with respect to Δ gives

$$\Delta_\star = \frac{c\Delta_s}{12\Delta_s - c}\,, \qquad \mathcal{T}_{max} = \frac{\pi}{3}(c - 6\Delta_s)\sqrt{\frac{c}{12\Delta_s - c}}\,. \tag{2.411}$$

Imposing $\mathcal{T}_{max} \leq 0$ gives

$$\Delta_s \geq c/6\,. \tag{2.412}$$

Hence, we find that using this method we can safely extend the validity of the Cardy formula to the range $\Delta_s \geq c/6$. For energies smaller than that $\tilde{Z}$ stops being $\mathcal{O}(1)$ and the saddle point analysis is not valid. Note that in this method

the contribution of light states ($\Delta < c/6$) was made $\mathcal{O}(1)$ by imposing a Hagedorn bound $\rho(\Delta < c/6) \lesssim \exp(2\pi\Delta)$. Here we have not used the result from [156] that the partition function at large N is dominated by the states with $\Delta \lesssim c/12$, which would allow us to only place a Hagedorn bound on those states.

Proceeding similarly for arbitrary d and nonzero J, we find Cardy behavior for the range

$$(d-1)EE_{\rm vac} - dE_{\rm vac}^2 + E^2 > J^2\,, \tag{2.413}$$

which is identical to the bulk result (2.402).

Chapter 3

Entanglement

3.1 Introduction

In recent years there has been a surge of activity surrounding entanglement as a key quantity of interest in quantum systems. It has an astonishing array of applications: it can be used to diagnose a wide variety of new phases of matter [355, 356]; it is a resource for quantum computation [354, 350, 351], and enables quantum error correction [352, 353, 354]; its growth reflects the dispersion of information [357, 358]; it is intimately linked to the connectedness of spacetime [359], and its dynamics can even be used to derive Einstein's equations [360, 361].

However a number of much more basic issues related to entanglement have not been completely addressed. One such question is its very definition: a rigorous computation of entanglement entropy (without resort to the replica trick) requires a choice of boundary conditions at the edge of the chosen entangling region, but on the other hand, imposing boundary conditions on such an imaginary surface

is clearly unphysical. In § 3.2 we studied how different choices of boundary conditions at the entangling surface affect the vacuum entanglement entropy in the simple setting of a free boson in 1+1D, elucidating boundary condition-dependent finite-size corrections that are missed in the naive calculation.

Another open question is the definition of entanglement between spatial regions in a theory with nonlocal degrees of freedom, such a gauge theory, or gravity. This issue is further confounded by the existence of symmetries that do not respect the local organization of degrees of freedom. A canonical example is electromagnetic duality, where the dual potential is determined from the original one by the nonlocal inversion of a differential operator. This suggests that the entanglement entropy might be different in the two duality frames. In § 3.3 we showed that this is indeed the case, quantifying the difference in entanglement between an abelian p-form theory and its electromagnetic dual. This discrepancy agrees with the electromagnetic anomaly of a $p-1$-form ghost theory living on the entangling surface, lending credence to the conjecture (known to be true in 1-form gauge theories) that correct calculation of entanglement requires treatment of an edge mode theory on the entangling surface. These results also resolved several discrepancies in the literature on electromagnetic duality.

One might wonder whether the proper treatment of edge modes or boundary conditions is relevant for the calculation of more traditional field theory quantities, such as correlation functions. In § 3.4 I studied correlation functions in Rindler quantization, where the Minkowski vacuum corresponds to a state with thermal entanglement between the two halves of space. A rigorous definition of the Rindler states requires the imposition of boundary conditions where the two halves meet,

but I found that the Rindler description reproduces Minkowski correlation functions without attention to this issue, even for trans-horizon correlators, in scalar and abelian gauge theories.

Finally, in § 3.5 we studied edge modes in the context of holographic codes: tensor network representations of holography that capture the quantum-error-correcting nature of bulk reconstruction. Most examples of holography involve bulk gauge fields, however, which cannot be described by a local bulk model. We extended the construction of holographic codes to include gauge degrees of freedom and studied how gauge degrees of freedom straddling the entangling surface are encoded on the corresponding boundary region. We found that the flux degrees of freedom on the entangling surface can be encoded in both the boundary region and its complement, but that the conjugate degrees of freedom are then encoded in neither side. We further suggest that this models the Ryu-Takayanagi proposal, in which the flux of the bulk graviton field – the area operator – is encoded in both of two complementary boundary regions.

3.2 Entanglement Entropy and Boundary Conditions in 1+1 Dimensions

This section aims to address the question of how to define entanglement entropy. When one attempts to calculate the entanglement entropy

$$S_A = -\mathrm{Tr}_A\ \rho_A \log\ \rho_A \tag{3.1}$$

of a reduced density matrix

$$\rho_A = \mathrm{Tr}_{\overline{A}}\ \rho, \tag{3.2}$$

the traces over $\mathcal{H}_A$ and $\mathcal{H}_{\bar{A}}$ are both taken on Hilbert spaces defined on regions with boundaries. These Hilbert spaces are only well-defined once boundary conditions have been specified at ∂A and $\partial \bar{A}$. However, this technical need to impose boundary conditions is at odds with the physics of the problem: the entangling surface is entirely fiducial, and the imposition of any condition on the fields on such a surface is unphysical.

We do not attempt to prescriptively resolve this tension, instead focusing on the simpler problem of characterizing how the choice of boundary conditions affects the entanglement calculation. Working in the highly simplified setting of a free boson in 1+1D, with A an interval of length ℓ, we find new contributions to the entanglement entropy from the choice of boundary conditions, independent of ℓ, which can in turn be identified with the boundary entropy of the associated state in boundary conformal field theory. We showed that these contributions can be derived either from the canonical trace or from the heat kernel calculation of

the oscillator partition function.

3.2.1 Introduction

Entanglement entropy in quantum field theory is now a well studied subject with a number of different applications [208, 209, 210]. In the case of conformal field theories, conformal symmetry provides powerful analytical tools for the computation of entanglement entropy (and, more generally, Renyi entropies) [211, 212, 213, 214], particularly in two spacetime dimensions [215]. The abstract nature of the these methodologies, however, can make the underlying physics obscure. In this paper, we do some basic computations with more pedestrian methods, in an attempt to elucidate some of the underlying issues.

In general, we consider a quantum field theory in d-dimensional Minkowski space $R^{1,d-1}$ with a unique ground state $|0\rangle$. We choose a finite contiguous spatial region A, and construct a density matrix ρ by tracing out the fields that live in the complementary region $\overline{A}$,

$$\rho = \mathrm{Tr}_{\overline{A}} |0\rangle\langle 0|. \tag{3.3}$$

The Renyi entropy with index β is then defined as

$$S_\beta = \frac{1}{1-\beta} \log \mathrm{Tr}\, \rho^\beta, \tag{3.4}$$

and the entanglement entropy is

$$S_{\rm EE} = -\operatorname{Tr}\rho\log\rho = \lim_{\beta\to 1} S_\beta. \tag{3.5}$$

Even in the simplest case of scalar fields, this procedure is ambiguous in the continuum, because in general specifying "the fields in region A" requires boundary conditions at the surface of A. In the case of gauge fields, much more analysis is needed [216, 217], since there is in general no gauge-invariant prescription for assigning the gauge potentials to regions [218]. For scalar fields in more than two spacetime dimensions, the issue of boundary conditions has been raised in the context of the influence of the conformal coupling of the scalar fields to the background spacetime curvature [219]. In two spacetime dimensions, the conformal coupling vanishes and so does not affect the calculation, but the question of appropriate boundary conditions remains [220, 221]. If the conformal field theory in question arises as the continuum limit of an underlying lattice theory, or we have other information about the ultraviolet regulator, it can be possible to deduce the appropriate boundary condition.

Our goal here is to explore the role of boundary conditions at the surface A in a simple example of a continuum conformal field theory in 1+1 dimensions, the $c = 1$ compact-boson. We will find that zero modes play the key role, as they do in gauge theories [217, 222, 223, 224]. Other related work includes [225, 226, 227, 228, 229, 230].

3.2.2 Review of results from conformal field theory

The density matrix ρ of Eq. (3.3) is a hermitian operator with nonnegative eigenvalues that obeys

$$\mathrm{Tr}\,\rho = 1. \tag{3.6}$$

Therefore we can write it in the form

$$\rho = \exp(-K), \tag{3.7}$$

where K is a hermitian operator known as the modular hamiltonian [231]. In the special case that the quantum field theory is a conformal field theory, and the spatial region A is a $(d-1)$-dimensional ball of radius R, conformal symmetry can be used to show that, up to a possible boundary term that will be discussed further below, K is the generator of a conformal transformation that preserves the boundary of the ball; this leads to [211, 213]

$$K = \int_A d^{d-1}x\, f(x) T^{00}(x) + c, \tag{3.8}$$

where $T^{\mu\nu}$ is the conformal traceless stress-energy tensor, obtained by varying the background metric, and including contributions from a coupling $\xi R\varphi^2$ to the background curvature R, with $\xi = \frac{d-2}{4(d-1)}$; the function $f(x)$ is

$$f(x) = \frac{\pi}{R}(R^2 - |x|^2), \tag{3.9}$$

and c is a constant that is fixed by the requirement $\mathrm{Tr}\,\rho = 1$.

The simplest case to consider is a single real compact scalar field φ,

$$\varphi(x) \sim \varphi(x) + V, \tag{3.10}$$

where V is the circumference of the target-space circle for φ. The conformal energy density is

$$T^{00} = \tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\partial_x\varphi)^2 + \xi\nabla^2\varphi^2, \tag{3.11}$$

where π is canonically conjugate to φ,

$$[\varphi(x,t), \pi(x',t)] = i\delta(x-x'). \tag{3.12}$$

There has been extensive discussion of whether an extra boundary term is also needed; see [219] and references therein. This issue arises from the coupling to the background curvature, and is absent in $d = 2$, where $\xi = 0$. We therefore specialize to this case, since the issues we wish to explore arise even in this simplest situation.

Having specialized to $d = 2$ and $\xi = 0$, we can make a change of the spatial coordinate,

$$dy = \frac{dx}{f(x)}, \tag{3.13}$$

$$y = \frac{1}{\pi}\tanh^{-1}(x/R). \tag{3.14}$$

The range $x \in (-R, R)$ corresponds to $y \in (-\infty, \infty)$. We then have

$$\partial_y\varphi(y) = f(x)\partial_x\varphi(x). \tag{3.15}$$

If we now define a rescaled conjugate momentum

$$\tilde{\pi}(y) = f(x)\pi(x), \tag{3.16}$$

then we have

$$[\varphi(y), \tilde{\pi}(y')] = i\delta(y - y') \tag{3.17}$$

and

$$K = \int_{-\infty}^{+\infty} dy \, \left[\tfrac{1}{2}\tilde{\pi}^2 + \tfrac{1}{2}(\partial_y\varphi)^2\right] + c. \tag{3.18}$$

Thus the modular hamiltonian corresponds to the hamiltonian of a free-field theory on an infinite line.

The entanglement entropy computed from Eqs. (3.5,3.7,3.18) is infinite. To regulate it, we give the y coordinate a finite range, $y \in (-L/2, L/2)$, with $L \gg 1$. This corresponds to a cutoff on the x coordinate at $|x| = R - \varepsilon$ with $\varepsilon \ll R$, and

$$L = \frac{1}{\pi} \log\left(\frac{2R}{\varepsilon}\right) + O(\varepsilon). \tag{3.19}$$

Thus the $\varepsilon \to 0$ limit corresponds to $L \to \infty$. It will also be convenient to shift the origin of the y coordinate, so that $y \in (0, L)$. We now have

$$K = \int_0^L dy \, \left[\tfrac{1}{2}\tilde{\pi}^2 + \tfrac{1}{2}(\partial_y\varphi)^2\right] + c, \tag{3.20}$$

which is the hamiltonian of a free-field theory on a finite interval of length L. To fully specify K, we will need to choose boundary conditions at the ends of the interval.

We begin by writing general mode expansions for φ, $\partial_y\varphi$, and $\tilde{\pi} = \partial_t\varphi$,

$$\varphi(y,t) = \varphi_0 + \varphi_1 y + \pi_0 t + \pi_1 y t + \sum_{k=1}^{\infty} \chi_k(y)\big[a_k e^{-i\omega_k t} + a_k^\dagger e^{i\omega_k t}\big], \tag{3.21}$$

$$\partial_y\varphi(y,t) = \varphi_1 + \pi_1 t + \sum_{k=1}^{\infty} \chi_k'(y)\big[a_k e^{-i\omega_n t} + a_k^\dagger e^{i\omega_k t}\big], \tag{3.22}$$

$$\tilde{\pi}(y,t) = \pi_0 + \pi_1 y - i\sum_{k=1}^{\infty} \omega_n \chi_k(y)\big[a_k e^{-i\omega_n t} - a_k^\dagger e^{i\omega_k t}\big], \tag{3.23}$$

where

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \tag{3.24}$$

and the mode function $\chi_k(y)$ satisfies $\chi_k'' + \omega_k^2\chi_k = 0$. We have assumed that the boundary conditions will render the allowed frequencies discrete and the mode functions real.

Since $\tilde{\pi}(y) = f(x)\pi(x)$ and $\partial_y\varphi(y) = f(x)\partial_x\varphi(x)$, and since $f(x)$ vanishes at the boundary $|x| = R$, a natural choice of boundary condition is to require both $\tilde{\pi}(y)$ and $\partial_y\varphi(y)$ to vanish at the regularized boundary points $y = 0$ and $y = L$. However, requiring both is incompatible with the commutation relation, Eq. (3.12). So we must make choice. We discuss three possible choices in the next three sections.

3.2.3 Spatial Neumann boundary conditions

We first consider what we will call spatial Neumann boundary conditions,

$$\partial_y\varphi(y,t)\big|_{y=0} = 0, \quad \partial_y\varphi(y,t)\big|_{y=L} = 0. \tag{3.25}$$

From Eq. (3.22), we see that spatial Neumann boundary conditions require $\chi_k'(0) = \chi_k'(L) = 0$, which fixes $\chi_k(y) \sim \cos(\omega_k y)$ and

$$\omega_k = \pi k/L. \tag{3.26}$$

The correctly normalized mode functions are then

$$\chi_k(y) = \frac{1}{\sqrt{\pi k}} \cos(\pi k y/L). \tag{3.27}$$

We also have the zero-mode conditions

$$\varphi_1 = 0, \quad \pi_1 = 0. \tag{3.28}$$

The remaining zero modes are then given by

$$\varphi_0 = \int_0^L dy\ \varphi(y,t), \quad \pi_0 = \int_0^L dy\ \tilde{\pi}(y,t), \tag{3.29}$$

and from the commutation relation, Eq. (3.12), we find

$$[\varphi_0, \pi_0] = \frac{i}{L}. \tag{3.30}$$

Since φ is compact, we have from Eq. (3.10) that $\varphi_0 \sim \varphi_0 + V$, and hence π_0 is quantized,

$$\pi_0 = \frac{2\pi}{LV} m, \quad m = 0, \pm 1, \ldots. \tag{3.31}$$

The modular hamiltonian is now

$$K = \frac{2\pi^2}{LV^2}m^2 + \sum_{k=1}^{\infty} \omega_k a_k^\dagger a_k + c. \tag{3.32}$$

We can now compute

$$\operatorname{Tr} e^{-\beta K} = e^{-\beta c} Z_{\text{osc}} Z_{0,\text{SN}}, \tag{3.33}$$

where the oscillator partition function is

$$Z_{\text{osc}} = \prod_{k=1}^{\infty} \sum_{n=0}^{\infty} e^{-\beta \omega_k n} \tag{3.34}$$

$$= \prod_{k=1}^{\infty} \frac{1}{1 - e^{-\beta\omega_k}} \tag{3.35}$$

$$= \left[e^{\pi\beta/24L}\, \eta(i\beta/2L) \right]^{-1}, \tag{3.36}$$

where $\eta(\tau)$ is the Dedekind eta function. For large L (more specifically, $L \gg \beta$), we have

$$\log Z_{\text{osc}} = \frac{\pi L}{6\beta} - \frac{1}{2}\log\left(\frac{2L}{\beta}\right) + O(\beta/L). \tag{3.37}$$

If we ignore the zero mode completely, and compute the Renyi entropy from $\operatorname{Tr}\rho^\beta = e^{-\beta c} Z_{\text{osc}}$ with c adjusted to make $\operatorname{Tr}\rho = 1$, we find

$$S_\beta^{\text{osc}} = \frac{(1+\beta)}{6\beta}\pi L - \frac{1}{2}\log 2L + \frac{\log\beta}{2(1-\beta)}. \tag{3.38}$$

Recalling that $\pi L = \log(2R/\varepsilon)$, we see that the first term is the usual result. However there is an additional $\log L \sim \log\log R/\varepsilon$ term, which is anomalous and not expected to appear in the final answer. The possibility of such a term was

noted in [217]. This result for the entanglement entropy was first found by [232], where the subleading term was not retained.

This anomalous term is canceled when we include the contribution of the zero modes. The zero-mode partition function with spatial Neumann boundary conditions is

$$Z_{0,\mathrm{SN}} = \sum_{m=-\infty}^{+\infty} e^{-2\pi^2\beta m^2/V^2 L} \tag{3.39}$$

$$= \vartheta(2\pi i\beta/LV^2). \tag{3.40}$$

where $\vartheta(\tau)$ is a Jacobi theta function (with the other argument $z = 0$). For large L, we have

$$\log Z_{0,\mathrm{SN}} = \frac{1}{2}\log\left(\frac{2L}{\beta}\right) + \frac{1}{2}\log\left(\frac{V^2}{4\pi}\right) + O(e^{-LV^2/2\pi\beta}). \tag{3.41}$$

Adding this to Eq. (3.37), we have

$$\log Z_{\mathrm{osc}} + \log Z_{0,\mathrm{SN}} = \frac{\pi}{6\beta}L + \frac{1}{2}\log\left(\frac{V^2}{4\pi}\right) + O(\beta/L). \tag{3.42}$$

The anomalous $\log L$ has now been canceled. The resulting Renyi entropy is

$$S_\beta = \frac{\pi(1+\beta)L}{6\beta} + \frac{1}{2}\log\left(\frac{V^2}{4\pi}\right) + O(\beta/L). \tag{3.43}$$

This is the usual result. The constant term, independent of both L and β but depending on V, can be understood as a contribution from the Affleck-Ludwig boundary entropy [233] with these boundary conditions [221, 220] (see [234] for a

review).

3.2.4 Temporal Neumann boundary conditions

We next consider what we will call temporal Neumann boundary conditions,

$$\tilde{\pi}(0,t) = 0, \quad \tilde{\pi}(L,t) = 0. \tag{3.44}$$

Again, these are motivated by $\tilde{\pi}(y) = f(x)\pi(x)$, and the vanishing of $f(x)$ at the ends of the interval in the original x coordinate. From Eq. (3.22), we see these boundary conditions require $\chi_k(0) = \chi_k(L) = 0$, which fixes $\chi_k(y) \sim \sin(\omega_k y)$ and Eq. (3.26). The correctly normalized mode functions are then

$$\chi_k(y) = \frac{1}{\sqrt{\pi k}} \sin(\pi k y/L). \tag{3.45}$$

We also have the zero-mode conditions

$$\pi_0 = 0, \quad \pi_1 = 0. \tag{3.46}$$

The remaining zero modes are then φ_0 and φ_1, with $\varphi_0 \sim \varphi_0 + V$; φ_1 has an infinite range. The modular hamiltonian is now

$$K = \frac{L}{2}\varphi_1^2 + \sum_{k=1}^{\infty} \omega_k a_k^\dagger a_k + c. \tag{3.47}$$

The oscillator contribution is therefore the same as it is for spatial Neumann boundary conditions, Eq. (3.36). The zero-mode contribution is now

$$Z_{0,\mathrm{TN}} = L \int_0^V d\varphi_0 \int_{-\infty}^{+\infty} d\varphi_1 \, e^{-\beta L \varphi_1^2/2} \tag{3.48}$$

The prefactor of L in Eq. (3.48) arises from the measure for a trace of a functional of $\varphi(x)$,

$$\prod_x d\varphi(x) = L \, d\varphi_0 \, d\varphi_1 \prod_{n=1}^{\infty} c_n, \tag{3.49}$$

where c_n is the coefficient of $\chi_n(x)$ in the mode expansion of $\varphi(x)$. The factor of L comes from the jacobian for this change of integration variables; its necessity can be seen from dimensional analysis. Evaluating the integral in Eq. (3.48), we have

$$Z_{0,\mathrm{TN}} = V \left(\frac{2\pi L}{\beta} \right)^{1/2}. \tag{3.50}$$

This yields

$$\log Z_{0,\mathrm{TN}} = \frac{1}{2} \log\left(\frac{2L}{\beta} \right) + \frac{1}{2} \log\left(\frac{V^2}{4\pi} \right) + O(e^{-LV^2/2\pi\beta}). \tag{3.51}$$

For large L, this is the same as $\log Z_{0,\mathrm{SN}}$, Eq. (3.41), up to exponentially small corrections, and therefore the result for the Renyi entropy is also the same, Eq. (3.43).

3.2.5 Dirichlet boundary conditions

Although not motivated by the vanishing of $f(x)$ at the endpoints, in this section we consider Dirichlet boundary conditions

$$\varphi(0,t) = 0, \quad \varphi(L,t) = 0 \bmod V. \tag{3.52}$$

From Eq. (3.22), we see these conditions require $\chi_k(0) = \chi_k(L) = 0$, which results in Eqs. (3.26) and (3.45), the same as for temporal Neumann boundary conditions. The zero-mode conditions are now

$$\varphi_0 = 0, \quad \pi_0 = 0, \quad \pi_1 = 0. \tag{3.53}$$

The remaining zero mode is then

$$\varphi_1 = \frac{wV}{L}, \quad w = 0, \pm 1, \ldots \tag{3.54}$$

The modular hamiltonian is now

$$K = \frac{V^2}{2L} w^2 + \sum_{k=1}^{\infty} \omega_k a_k^\dagger a_k + c. \tag{3.55}$$

The oscillator contribution is therefore the same as it is for spatial or temporal Neumann boundary conditions, Eq. (3.36). The zero-mode contribution is now

$$Z_{0,\mathrm{D}} = \sum_{w=-\infty}^{+\infty} e^{-\beta V^2 m^2/2L} \tag{3.56}$$

$$= \vartheta(i\beta V^2/2\pi L). \tag{3.57}$$

For large L, we have

$$\log Z_{0,\mathrm{D}} = \frac{1}{2}\log\left(\frac{2L}{\beta}\right) - \frac{1}{2}\log\left(\frac{V^2}{\pi}\right) + O(e^{-2\pi L/\beta V^2}). \tag{3.58}$$

Combining Eqs. (3.37) and (3.58) for large L, we have

$$\log Z_{\mathrm{osc}} + \log Z_{0,\mathrm{D}} = \frac{\pi L}{6\beta} - \frac{1}{2}\log\left(\frac{V^2}{\pi}\right) + O(\beta/L). \tag{3.59}$$

The anomalous $\log L$ has now again been canceled. The resulting Renyi entropy is

$$S_\beta = \frac{(1+\beta)}{6\beta}\pi L - \frac{1}{2}\log\left(\frac{V^2}{\pi}\right) + O(\beta/L). \tag{3.60}$$

The constant term is different than it is in the case of spatial or temporal Neumann boundary conditions, and again can be understood as a contribution from the Affleck-Ludwig boundary entropy [233].

3.2.6 Duality

Comparing Eq. (3.43) for the Renyi entropy with Neumann boundary conditions (spatial or temporal) with Eq. (3.60) for the Renyi entropy with Dirichlet boundary conditions, we see that they are related by

$$V \leftrightarrow \frac{2\pi}{V}. \tag{3.61}$$

For the case of spatial Neumann boundary conditions, this follows from the same relation for the zero-mode partition functions, Eqs. (3.40) and (3.57). This is

related to the T-duality transformation for the compact boson on a spatial circle with circumference $2L$ with periodic boundary conditions. In this case, the mode expansion is

$$\varphi(y,t) = \varphi_0 + \varphi_1 y + \pi_0 t + \frac{1}{\sqrt{2\pi}} \sum_{k=1}^{\infty} \frac{1}{k} \big[a_k e^{-i\omega_k (y-t)} + \tilde{a}_k e^{i\omega_k (y+t)} + \text{h.c.} \big], \quad (3.62)$$

with

$$\varphi_1 = \frac{V}{2L} w, \quad \pi_0 = \frac{\pi}{LV} m, \quad (3.63)$$

where again w and m are integers representing winding and momentum modes, and now there are two types of oscillators (left moving and right moving modes) with $\omega_k = \pi k/L$; matching Eq. (3.26) is the reason for having the circle be twice as long as the interval. The hamiltonian is

$$H = \frac{\pi^2}{LV^2} m^2 + \frac{V^2}{4L} w^2 + \sum_{k=1}^{\infty} \omega_k \big(a_k^\dagger a_k + \tilde{a}_k^\dagger \tilde{a}_k \big) + c. \quad (3.64)$$

With a judicious choice of c, and introducing the modular parameter

$$\tau = i\beta/2L, \quad (3.65)$$

the complete partition function is

$$Z_{\text{circle}}(\tau, V) = \frac{\vartheta(2\pi\tau/V^2)\vartheta(\tau V^2/2\pi)}{\eta(\tau)^2} \quad (3.66)$$

which is manifestly invariant under Eq. (3.61). We also have the relation [235]

$$Z_{\text{circle}}(\tau, V) = Z_{\text{SN}}(\tau, \sqrt{2}V) Z_{\text{D}}(\tau, V/\sqrt{2}) \tag{3.67}$$

for particular choices of c in Eqs. (3.32) and (3.55). Since $Z_{\text{circle}}(\tau, V)$ can be written as a euclidean path integral over the field on a 2-torus with cycle lengths β and $2L$, we also have invariance under $\beta \leftrightarrow 2L$ or equivalently $\tau \leftrightarrow -1/\tau$, which follows from

$$\eta(-1/\tau) = (-i\tau)^{1/2} \eta(\tau), \tag{3.68}$$

$$\vartheta(-1/\tau) = (-i\tau)^{1/2} \vartheta(\tau). \tag{3.69}$$

The boundary conditions on the interval of length L break this symmetry, but as a remnant of it we have the relations

$$Z_{\text{SN}}(-1/\tau, V) = \sqrt{V^2/4\pi}\, Z_{\text{D}}(\tau, V/2), \tag{3.70}$$

$$Z_{\text{D}}(-1/\tau, V) = \sqrt{\pi/V^2}\, Z_{\text{SN}}(\tau, 2V). \tag{3.71}$$

3.2.7 Comparison with heat kernel methods

In this section we evaluate the entanglement entropy using heat-kernel methods to facilitate comparison with the computation of Casini and Huerta [212]. For a review of these techniques, see [236].

The key formula is

$$\log Z_{\rm osc} = \frac{1}{2}\lim_{s\to 1}\int_0^\infty \frac{dt}{t^s}K(t), \tag{3.72}$$

where $K(t) = \mathrm{Tr}'\, e^{-t\Delta}$ and Δ is minus the Laplacian on the target manifold (in our case, an interval I of length L with one of our sets of boundary conditons) times a circle S^1 of circumference β, and the prime on the trace indicates that we omit the zero modes on I. Thus we have $\Delta = -\partial_\tau^2 - \partial_y^2$, where τ is periodic with period β. Since this is a product space, the trace factorizes,

$$K(t) = K_{S^1}(t)K_I(t) \tag{3.73}$$

where $K_{S^1}(t)$ is the regulated form (defined shortly) of the unregulated heat kernel $\tilde{K}_{S^1}(t)$ on a circle of circumference β,

$$\tilde{K}_{S^1}(t) = \sum_{n=-\infty}^{\infty} e^{t(2\pi i n/\beta)^2} \tag{3.74}$$

$$= \vartheta(4\pi i t/\beta^2) \tag{3.75}$$

$$= \sqrt{\beta^2/4\pi t}\,\vartheta(i\beta^2/4\pi t) \tag{3.76}$$

$$= \frac{\beta}{\sqrt{4\pi t}}\sum_{n=-\infty}^{\infty} e^{-n^2\beta^2/4t}. \tag{3.77}$$

The regulated heat kernel is obtained by dropping the $n = 0$ term, or equivalently, subtracting the $\beta \to \infty$ limit:

$$K_{S^1}(t) = \frac{2\beta}{\sqrt{4\pi t}}\sum_{n=1}^{\infty} e^{-n^2\beta^2/4t}. \tag{3.78}$$

We can check that this regulated heat kernel on the circle gives the correct answer for the partition function by noting that, for a partition function with the general form of Eq. (3.35),

$$Z = \prod_{k=1}^{\infty} \frac{1}{1 - e^{-\beta\omega_k}}, \tag{3.79}$$

we have

$$-\log(1 - e^{-\beta\omega_k}) = \frac{1}{2} \lim_{s\to 1} \int_0^\infty \frac{dt}{t^s} K_{S^1}(t) e^{-t\omega_k^2}. \tag{3.80}$$

Summing over k then gives $\log Z$ on the left, and we identify

$$K_I(t) = \sum_{k=1}^{\infty} e^{-t\omega_k^2} \tag{3.81}$$

on the right.

For all our choices of boundary conditions, we have $\omega_k = \pi k/L$, and hence

$$\begin{aligned} K_I(t) &= \sum_{k>0} e^{-\pi^2 k^2 t/L^2} && (3.82)\\ &= \frac{1}{2}\left[\vartheta(i\pi t/L^2) - 1\right] && (3.83)\\ &= \frac{1}{2}\left[\sqrt{L^2/\pi t}\,\vartheta(iL^2/\pi t) - 1\right] && (3.84)\\ &= \frac{1}{2}\left[\sqrt{L^2/\pi t}\left(2\sum\nolimits_{m>0} e^{-m^2L^2/t} + 1\right)\right] - 1 && (3.85)\\ &= \frac{L}{\sqrt{4\pi t}}\left[1 + 2\sum_{m>0} e^{-m^2L^2/t} - \frac{\sqrt{\pi t}}{L}\right]. && (3.86) \end{aligned}$$

Plugging back into Eq. (3.72), we find

$$\log Z_{\text{osc}} = \frac{\beta L}{4\pi} \lim_{s\to 1} \int_0^\infty \frac{dt}{t^{s+1}} \sum_{n>0} e^{-n^2\beta^2/4t} \left[1 + 2\sum_{m>0} e^{-m^2L^2/t} - \frac{\sqrt{\pi t}}{L}\right]$$

$$= I_1 + I_2 + I_3. \tag{3.87}$$

In the analysis of [212], only I_1 is kept; the second and third terms are subleading in L, and are dropped. However, I_2 and I_3 must also be included in order to reproduce the canonical oscillator partition function, Eq. (3.36), and supplemented by the appropriate zero-mode contribution to obtain the entropy. We compute the integrals in Appendix 3.2.9, with the result

$$\begin{aligned} I_1 &= \frac{\pi L}{6\beta}, \\ I_2 &= -\frac{\pi L}{6\beta} - \frac{\pi\beta}{24L} - \log\eta(i\beta/2L) + \frac{1}{4}\lim_{s\to 1}\left(\frac{1}{s-1} + \gamma_E - 2\log\beta\right), \\ I_3 &= -\frac{1}{4}\lim_{s\to 1}\left(\frac{1}{s-1} + \gamma_E - 2\log\beta\right) \end{aligned} \tag{3.88}$$

where γ_E is the Euler-Mascheroni constant. Thus the heat kernel gives

$$\log Z_{\rm osc} = I_1 + I_2 + I_3 = -\frac{\pi\beta}{24L} - \log\eta(i\beta/2L) \tag{3.89}$$

in exact agreement with Eq. (3.36).

3.2.8 Discussion

We have undertaken a detailed analysis of the computation of the entanglement entropy (and the Renyi entropies) of an interval for the compact-boson $c = 1$ conformal field theory in 1+1 dimensions, paying particular attention to the role of boundary conditions at the endpoints and the contributions of zero modes, and using operator methods rather than euclidean path integral manipulations.

The possibility of temporal Neumann boundary conditions, and the distinction of them from Dirichlet boundary conditions, does not seem to have been previously considered. Our results emphasize the necessity of paying careful attention to these issues.

Acknowledgements

We are happy to thank Vlad Rosenhaus, Dan Harlow, and especially Will Donnelly and Aron Wall for extensive discussions. This work was supported in part by NSF Grant PHY13-16748.

3.2.9 Appendix: Integrals from §3.2.7

In this appendix we evaluate the integrals in (3.87).

$$\begin{aligned} I_1 &= \frac{\beta L}{4\pi} \lim_{s\to 1} \sum_{n>0} \int_0^\infty \frac{dt}{t^{s+1}} e^{-n^2\beta^2/4t} \\ &= \frac{\beta L}{4\pi} \lim_{s\to 1} 4^s \beta^{-2s} \Gamma(s)\zeta(2s) \\ &= \frac{\pi L}{6\beta}. \end{aligned} \tag{3.90}$$

Note that we could have set $s = 1$ at the start for I_1, but the same is not true for I_2 and I_3. We have

$$\begin{aligned} I_3 &= -\frac{\beta}{4\sqrt{\pi}} \lim_{s\to 1} \int_0^\infty \frac{dt}{t^{s+1/2}} \sum_{n>0} e^{-n^2\beta^2/4t} \\ &= -\frac{\beta}{4\sqrt{\pi}} \lim_{s\to 1} (\beta/2)^{1-2s} \Gamma(s-1/2)\zeta(2s-1) \end{aligned}$$

$$= -\frac{1}{4}\lim_{s\to 1}\left(\frac{1}{s-1} + \gamma_E - 2\log\beta\right). \tag{3.91}$$

Finally, we evaluate I_2:

$$\begin{aligned} I_2 &= \frac{\beta L}{2\pi}\lim_{s\to 1}\int_0^\infty \frac{dt}{t^{s+1}} \sum_{m,n>0} e^{-n^2\beta^2/4t - m^2L^2/t} \\ &= \frac{\beta L}{2\pi}\lim_{s\to 1}\Gamma(s)L^{-2s}\sum_{m,n>0}\frac{1}{[(n\beta/2L)^2+m^2]^s}. \end{aligned} \tag{3.92}$$

The double sum can be expressed in terms of an Eisenstein series

$$E(\tau,s) = \sum_{(m,n)\neq(0,0)} \frac{(\operatorname{Im}\tau)^s}{|n\tau+m|^{2s}} \tag{3.93}$$

with $\tau = i\beta/2L$. We have

$$\begin{aligned} \sum_{m,n>0}\frac{1}{[(n\beta/2L)^2+m^2]^s} &= \frac{1}{4}(\beta/2L)^{-s}E(i\beta/2L,s) - \frac{1}{2}(\beta/2L)^{-2s}\sum_{n>0}\frac{1}{n^{2s}} - \frac{1}{2}\sum_{m>0}\frac{1}{m^{2s}} \\ &= \frac{1}{4}(\beta/2L)^{-s}E(i\beta/2L,s) - \frac{1}{2}(\beta/2L)^{-2s}\zeta(2s) - \frac{1}{2}\zeta(2s). \end{aligned} \tag{3.94}$$

Here the sum over the positive quadrant of $\mathbb{Z}^2$ was rewritten as the sum over $(m,n)\neq(0,0)$, minus the $(m,0)$ and $(0,n)$ lines, all divided by 4. We now have

$$I_2 = \frac{\beta L}{8\pi}\lim_{s\to 1}\Gamma(s)\left[(\beta L/2)^{-s}E(i\beta/2L,s) - 2(\beta/2)^{-2s}\zeta(2s) - 2L^{-2s}\zeta(2s)\right]. \tag{3.95}$$

The limits of the last two terms are simple to evaluate, and yield $-\pi L/6\beta$ and $-\pi\beta/24L$, respectively. To evaluate the first term, we need the Kronecker limit

formula for Eisenstein series near $s = 1$,

$$E(\tau, s) = \frac{\pi}{s-1} + 2\pi \left[\gamma_E - \log 2 - \log\left(\sqrt{\operatorname{Im}\tau}\,|\eta(\tau)|^2\right)\right] + O(s-1). \tag{3.96}$$

With this we find

$$I_2 = -\log \eta(i\beta/2L) + \frac{1}{4}\lim_{s\to 1}\left(\frac{1}{s-1} + \gamma_E - 2\log\beta\right) - \frac{\pi L}{6\beta} - \frac{\pi\beta}{24L}. \tag{3.97}$$

3.3 Electromagnetic duality and entanglement anomalies

Electromagnetic duality exchanges electric and magnetic fields via

$$F \leftrightarrow \star F. \tag{3.98}$$

It is a symmetry of pure abelian gauge theories, as noted by Maxwell, and also generalizes to some theories with matter [362, 363]. However, under this symmetry the gauge potential A (with $dA = F$) is exchanged for the dual potential $\tilde{A}$ (with $d\tilde{A} = \star F$). From (3.98) it follows that

$$\tilde{A} = d^{-1} \star dA \tag{3.99}$$

and so the relation between the two potentials is nonlocal. This raises the question of how quantities that probe of the localization of degrees of freedom – such as entanglement entropy – behave under duality.

In the following work we studied this question in the context of abelian p-form electromagnetism, which describes the dynamics of fluxes sourced by charges extended in $p-1$ spatial dimensions. Building on [257, 217], which carefully analyzed the partition function in 1-form electromagnetism, we generalized to the partition function of the p-form theory and then considered the question of duality. We found that erroneous neglect of certain ingredients in the partition function calculation – instantons and zeromodes – had led to conflicting results [243, 244] in the literature on duality. We showed that the inclusion of these ingredients

resolves the discrepancy, with the result that the duality symmetry is anomalous when the spacetime dimension is even but exact when the spacetime dimension is odd.

We used these results to compute the change in entanglement entropy under a duality transformation, relating the change in the partition function to a change in entanglement entropy via the replica trick. We found that in odd dimensions there is no change in entanglement entropy under duality, while in even dimensions the change in entanglement entropy agrees with the duality anomaly of a $p-1$-form ghost theory confined to the entangling surface. It has previously been suggested that such "edge mode" theories naturally arise in computations of entanglement, and that their effects resolve a discrepancy between the gauge theory entropy and the conformal anomaly of the theory [217, 275, 364]. Our work lends credence to this claim in the more general context of p-form theories, as the entanglement anomaly we discovered would arise naturally from the presence of an edge mode theory in the replica partition function.

3.3.1 Introduction

Electromagnetic duality is a symmetry of many gauge theories, but the dual degrees of freedom are not local in the original variables. Thus one might expect probes of the localization of correlations – such as entanglement entropy – to depend on the duality frame.

One particularly tractable example is the electromagnetic duality of Maxwell theory, where the the field strength $F_{\mu\nu}$ is interchanged with its Hodge dual

$(\star F)_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$.[1] Depending on the context, this is also called Poincaré or S-duality, and generalizes to Maxwell theories with p-form potentials in D spacetime dimensions.[2] The dual $\tilde{p} = D - p - 2$-form potentials satisfy $dA = F$ and $d\tilde{A} = \star F$ and so determining one in terms of the other requires the nonlocal inversion of a differential operator.

For many purposes the dual theories are identical, *e.g.* the space of classical solutions is the same. But on the quantum level there are conflicting opinions [239, 240, 241, 242, 243, 244, 245] about the extent to which the dual theories may be regarded as equivalent, and even whether the trace anomaly agrees when $p \neq \tilde{p}$. Quantum equivalence requires the dual partition functions Z and $\tilde{Z}$ to be exactly equal, which is not always the case. For example the partition function of Maxwell theory in $D = 4$ transforms under electromagnetic duality as a modular form [244], which characterizes the anomaly in the duality symmetry.

The results [243, 244] are in conflict. In [243] Schwarz and Tyupkin compute the ratio of partition functions of dual p-form theories by computing a ratio of functional determinants, which come from Gaussian integrals over the non-zero modes of the Laplacian. This calculation yields a vanishing anomaly in even dimensions, but Witten later computed a nontrivial anomaly in ordinary 1-form Maxwell theory in four dimensions [244], which was confirmed by subsequent calculations [246, 237]. To our knowledge this conflict has not been resolved in the literature. However, our interest in the duality properties of entropy led to a more thorough calculation of the duality anomaly of p-form gauge theories, which

[1] Note that we assume oriented manifolds throughout, since Hodge and Poincaré duality hold only in that case. Both of these dualities can be generalized to the non-orientable case [237], but we leave such a generalization to future work.

[2] These p-forms appear in string theory as the gauge fields coupled to D-branes [238].

enables us to reconcile these results. We find that previously-neglected zero modes and instantons contribute factors that 1) give rise to the even-dimensional anomaly and 2) trivialize the duality in odd dimensions. This proves, for example, quantum equivalence of the scalar and photon in $D = 3$, and reproduces the known anomaly in $D = 4$. Our method extends to arbitrary D and p.[3]

In general, we find the following for the duality anomaly of an abelian p-form gauge theory on M:

$$\log \frac{Z_p}{\tilde{Z}_{\tilde{p}}} = \begin{cases} (-1)^{p+1}\ \chi(M) \log \sqrt{\frac{q}{\tilde{q}}} & D \text{ even} \\ 0 & D \text{ odd}, \end{cases} \tag{3.100}$$

where Z_p is the partition function of the gauge theory and $\tilde{Z}_{\tilde{p}}$ the partition function of its electromagnetic dual. q and $\tilde{q} = 2\pi/q$ are the couplings of the dual theories, which enter into the path integral of a U(1) gauge field via the flux quantization condition. The argument of the log is in units of a mass scale μ that must be introduced to define the quantum theory. The scale μ does not enter into the classical theory, the classical duality, or the quantum correlation functions, but it does enter into the anomalous quantum duality, where its role is to fix the units.

The vanishing of the odd-dimensional anomaly follows from a theorem of Cheeger [248] that equates the ratio of analytic torsion [249] (a product of functional determinants related to the partition function of Chern-Simons theory) and Reidemeister torsion [250] (a combinatorial quantity invented in the 1930s to distinguish lens spaces) with a ratio of the sizes of torsion subgroups. Although this

[3]We do not consider the case of massive p-form theories, for which the duality relation is instead $\tilde{p} = D - p - 1$, but [247] argues that there is no duality anomaly in this case.

combination of quantities from entirely different branches of mathematics may seem obscure, each quantity appears naturally in the ratio of partition functions. This physical application of the Cheeger-Müller theorem may be of interest apart from our study of the duality properties of entanglement entropy.

The even-dimensional duality anomaly is purely topological, and may also be absorbed into a local counterterm. In computations involving renormalization, often one already needs the leeway to shift the action by a local counterterm, and the duality anomaly may simply be absorbed. But it other contexts knowing the exact form of the anomaly is important.[4]

Our calculation of entanglement entropy relies on the replica trick [208, 215], which expresses the entropy in terms of a partition function. Thus one might expect that entanglement inherits some of the duality structure. In detail, we calculate the entanglement entropy $S_A = -\operatorname{Tr}\rho_A \log \rho_A$ of a region A by first computing $\operatorname{Tr}\rho_A^n$, analytically continuing to non-integer n, then using the identity $S_A = \lim_{n\to 1}(1- n\partial_n)\ \log \operatorname{Tr}\rho_A^n$ to obtain the entropy. Specializing to the vacuum and constructing powers of the vacuum reduced density matrix as euclidean path integrals, one finds $\operatorname{Tr}\rho_A^n = Z(M_A^{(n)})$, the partition function of the theory on the "replica manifold" $M_A^{(n)}$ of index n. Calculation of the entropy is reduced to the calculation of a replica partition function, and if one identifies the replica index with an inverse temperature the entanglement entropy is its thermodynamic entropy at $n = 1$.

However, eq. (3.100) implies that $Z(M_A^{(n)})$ can transform anomalously under

[4]Examples include: (i) nonrenormalization theorems, where a finite shift in a quantity might violate the theorem, or (ii) p-form fields that arise from Kaluza-Klein reduction, where it is important to preserve the local covariance of the higher dimensional theory [217].

electromagnetic duality, and so vacuum entanglement may also depend on the duality frame. We call this phenomenon an "entanglement anomaly". [5] It is given by

$$\Delta S_A = (1 - n\partial_n) \log \frac{Z(M_A^{(n)})}{\tilde{Z}(M_A^{(n)})} \tag{3.101}$$

evaluated at $n = 1$. The ratio can be computed using (3.100). For a p-form theory at coupling q, the change in entanglement entropy A is

$$\Delta S_A = \begin{cases} (-1)^{p-1}\ \chi(\partial A) \log \sqrt{\frac{q}{\tilde{q}}} & D \text{ even} \\ 0 & D \text{ odd,} \end{cases} \tag{3.102}$$

where $\chi(\partial A)$ is the Euler characteristic of the entangling surface. This ratio is the duality anomaly of a $(p-1)$-form edge mode theory on the entangling surface. Since the partition function of an abelian gauge theory on a replica manifold contains replica index-independent pieces that correspond to edge modes living on the entangling surface [257, 217], the entanglement anomaly arises naturally from this effect. This is discussed in § 3.3.4. We also consider theories with a θ-term, see § 3.3.4.

While the constant term in the even-D entanglement entropy changes under rescalings of the cutoff [208], we show in § 3.3.4 that the constant term in the even-dimensional entanglement anomaly is actually unchanged under simultaneous transformation of the two theories. This is consistent with recent results in the condensed matter literature [258, 259] and we give a general derivation.

[5] Previously entanglement has been shown to transform anomalously under other symmetries, e.g. under a Lorentz boost in theories with chiral anomalies [251, 252, 253, 254, 255, 256]. In this work we find an analogous effect when the duality symmetry is anomalous.

Now we outline the body of the paper. In § 3.3.2 we describe our calculation of the partition function of p-form gauge theory on an arbitrary manifold and outline our calculation of the ratio of electromagnetic dual partition functions $Z_p/\tilde{Z}_{\tilde{p}}$. We explain how to reconcile the conflicting results of [243, 244] and why the anomaly vanishes in odd D. In § 3.3.3, we explain why the stress tensor is the same for the dual theories in even D (including the trace), in agreement with the arguments of [242]. Finally, in § 3.3.4 we use the partition function to compute the entanglement anomaly. We show that thermal entropy is duality-invariant and address the question of universality under a change of regulator, and conclude by interpreting the entanglement anomaly physically as the duality anomaly of an edge mode theory living on the entangling surface.

Details of the duality calculations are left to appendices 3.3.6 and 3.3.7. In appendix 3.3.8 we work out a simple example that illustrates the importance of zero modes: Maxwell theory in one spacetime dimension, which has no states besides the vacuum. The oscillator partition function fails to reproduce the trivial canonical sum over states, unless accompanied by the zero mode contribution.

Recent related work includes [229], where the author considers the interplay between entanglement and duality in discrete spin systems; [260], which discusses the conformal p-form theories; [261], which develops the extended Hilbert space in the magnetic representation. [262] carried out some explicit calculations of p-form partition functions on the sphere and confirms the existence of the anomaly in even but not odd dimensions. [223] made use of duality to relate the entanglement entropy of a Maxwell theory to a compact scalar in $2+1$ dimensions; our results justify their use of this duality, since we show that it is exact in odd dimensions.

3.3.2 Partition functions and duality

In this section we calculate the partition function of p-form Maxwell theory and the ratio between the partition function and its electromagnetic dual.

The partition function of p-form gauge theory

Consider p-form electrodynamics on a compact manifold M, with gauge potential A and field strength $F = dA$. The Euclidean action is[6]

$$I = \frac{1}{2q^2}\int_M (\star F)\wedge F = \frac{1}{2(p+1)!q^2}\int_M \sqrt{g}F^{\mu\nu\cdots}F_{\mu\nu\cdots}. \tag{3.103}$$

When $p = 1$ this reduces to the familiar Maxwell action $\frac{1}{4q^2}\int\sqrt{g}F_{\mu\nu}F^{\mu\nu}$. The constant q is the coupling constant, the fundamental unit of charge.

We will compute the partition function by generalizing the approach developed in Ref. [263] to p-form theories. The partition function on M is given by the Euclidean path integral

$$Z = \sum_{\text{bundles}}\int \mathcal{D}[A/\mathcal{G}]\, e^{-I[A/\mathcal{G}]}. \tag{3.104}$$

The path integral is over all equivalence classes of connections, which we denote $A/\mathcal{G}$. This includes a sum over all gauge bundles (allowing for field strengths F that cannot be globally expressed as $F = dA$) and over connections with vanishing curvature $F = 0$ (zero modes).

We first decompose the field strength F into a piece that comes from a globally-

[6]This defines our convention for the Hodge dual $\star$.

defined p-form potential and a piece that does not:

$$F = \mathscr{F} + dA. \tag{3.105}$$

where A is the p-form potential and $\mathscr{F}$ is the part of F that cannot be written as dA. The Bianchi identity implies $dF = 0$, so $\mathscr{F}$ is an element of the $(p+1)^{\text{st}}$ cohomology group of M. The Dirac quantization condition further restricts $\mathscr{F}$ to be an element of the integer cohomology,

$$\mathscr{F} \in 2\pi H^{p+1}(M, \mathbb{Z}). \tag{3.106}$$

The decomposition (3.105) only fixes $\mathscr{F}$ up to the addition of an exact form. We can fix this remaining freedom by choosing $\mathscr{F}$ to be harmonic, $\Delta_{p+1}\mathscr{F} = 0$. This choice makes the decomposition (3.105) orthogonal, and as a result the action I splits as a sum over $\mathscr{F}$ and an integral over A:

$$I = \frac{1}{2q^2}\int_M \left[(\star\mathscr{F}) \wedge \mathscr{F} + (\star dA) \wedge dA\right]. \tag{3.107}$$

The sum over instantons therefore decouples from the remainder of the partition function; we will return to it after first considering the functional integral over the potential A.

To carry out the integration over the p-form potential A, we write the mode expansion

$$A = A_{\text{zero}} + \sum_n \alpha_n A_n, \tag{3.108}$$

where A_n are the nonzero modes of the p-form Laplacian, $\Delta_p A_n = \lambda_n A_n$, which

are chosen to be orthonormal. A_{zero} is the zero mode satisfying $\Delta_p A_{\text{zero}} = 0$. We will first deal with the nonzero modes, then treat A_{zero} separately.

We must introduce Faddeev-Popov ghosts that cancel out the unphysical polarizations in order to carry out the gauge-invariant path integral (3.103).[7] In p-form gauge theory, these ghosts are $(p-1)$-forms with fermionic statistics. However, these $(p-1)$-forms have their own $(p-2)$-form gauge symmetry: some of the gauge transformations are redundant, and the ghosts subtract too many degrees of freedom. It is then necessary to add in further positive degrees of freedom via ghosts-for-ghosts [264, 265, 243]. One must continue in this way, introducing k-form fields for all $k = 0, \ldots, p$ with alternatingly bosonic and fermionic statistics. The number of ghosts increases as the form degree decreases, so we have one p-form gauge field, two $(p-1)$-form ghosts, three $(p-2)$-form ghosts-for-ghosts, etc.

Having introduced the ghosts, the action for the field A is non-degenerate, $I = \frac{1}{2q^2}\langle A, \Delta_p A\rangle$, so we can carry out the path integral as usual. The space of p-forms comes with a natural measure induced by the inner product on p-forms. Because the modes A_n are orthonormal, this measure can be expressed in terms of the coefficients α_n of the mode expansion as

$$\mathcal{D}A = \prod_n \frac{\mu\, d\alpha_n}{\sqrt{2\pi} q}. \tag{3.109}$$

The reason for our choice of overall multiplicative constant $1/\sqrt{2\pi}q$ will become

[7] One does not usually consider ghosts in abelian theories, since they decouple from the physical polarizations and hence do not contribute to correlation functions. However, the ghosts still contribute to the partition function and therefore to the entropy.

clear in the course of the calculation. We also had to introduce a parameter μ, with dimensions of mass, since the measure must be dimensionless. Most quantities are independent of μ but it is part of the definition of the theory; it will set the units in the duality.

The path integral over nonzero modes of A is a Gaussian integral, and hence reduces to a functional determinant

$$\int \mathcal{D}A e^{-\frac{1}{2q^2}\langle A,\Delta_p A\rangle} = \prod_n \int \frac{\mu\, d\alpha_n}{\sqrt{2\pi}q} e^{-\frac{1}{2q^2}\lambda_n \alpha_n^2} = \prod_n \left(\frac{\lambda_n}{\mu^2}\right)^{-1/2} = \det\left(\frac{\Delta_p}{\mu^2}\right)^{-1/2}. \tag{3.110}$$

Carrying out the analogous integrals for the various ghost fields then leads to a string of determinants:

$$\begin{aligned} Z_{\text{osc}} &= \det(\tilde{\Delta}_p)^{-1/2} \det(\tilde{\Delta}_{p-1})^{+1} \cdots \det(\tilde{\Delta}_0)^{(-1)^{p+1}\frac{p+1}{2}} \\ &= \prod_{k=0}^{p} \det(\tilde{\Delta}_k)^{(-1)^{p+1-k}(p+1-k)/2}. \end{aligned} \tag{3.111}$$

where $\tilde{\Delta}_k = \Delta_k/\mu^2$. The exponent reflects the proliferation of ghosts, and the sign reflects their alternatingly fermionic and bosonic character.

Next, there is the integral over flat connections A_{zero}. Let $H^p(M)$ denote the space of harmonic p-forms, a real vector space of dimension b_p (the p^{th} Betti number). For flat connections the action vanishes, so the path integral for these modes simply computes the volume of the space of flat connections in the measure (3.109). This volume is finite because we identify p-form potentials under large gauge transformations, which are elements of the integer cohomology group $H^p(M,\mathbb{Z})$. As a discrete abelian group, it splits into a free part (given by b_p copies

of $\mathbb{Z}$) and a torsion part (a finite abelian group):

$$H^p(M,\mathbb{Z}) = \text{Free}\, H^p(M,\mathbb{Z}) \oplus \text{Tor}\, H^p(M,\mathbb{Z}) = \mathbb{Z}^{b_p} \oplus T^p. \tag{3.112}$$

First we deal with the free part. The subspace Free $H^p(M,\mathbb{Z})$ is simply the space of harmonic p-forms whose integrals around all noncontractible p-dimensional surfaces are integers. As a group it is equal to b_p copies of the additive group of the integers: Free $H^p(M,\mathbb{Z}) = \mathbb{Z}^{b_p}$. Thus two p-form potentials A and A' are equivalent if the integral of $A - A'$ over any p-dimensional surface is a multiple of 2π. This is just a generalization of the Aharonov-Bohm phase to higher-form gauge theories.

We define for each k a topological basis $\{w_i\}_{i=1}^{b_k}$ of Free $H^k(M,\mathbb{Z})$. The space of flat connections can then be parametrized as

$$A_{\text{zero}} = \sum_{i=1}^{b_p} \beta_i w_i, \tag{3.113}$$

where $\beta_i = [0, 2\pi)$. Written in the topological basis, the inner product on p-forms Γ_p has components

$$[\Gamma_p]_{ij} = \int_M (\star w_i) \wedge w_j. \tag{3.114}$$

Integrating over the space of flat connections modulo large gauge transformations with the measure (3.109) then gives a factor of

$$\det\left(\frac{2\pi\mu^2}{q^2}\Gamma_p\right)^{1/2} \tag{3.115}$$

There is a similar term that appears for the $(p-1)$ form gauge transformations. In

dividing by the volume of the gauge group, we must also divide by zero modes of the gauge transformations, which are flat $(p-1)$-forms. These gauge symmetries have their own gauge redundancy given by $(p-2)$-forms, we must *multiply* by the volume of the space of flat $(p-2)$ forms. Continuing in this way we obtain the complete zero mode contribution

$$Z_{\text{zero}} = \det\left(\frac{2\pi\mu^2}{q^2}\Gamma_p\right)^{1/2} \det\left(\frac{2\pi\mu^4}{q^2}\Gamma_{p-1}\right)^{-1/2} \cdots \det\left(\frac{2\pi\mu^{2(p+1)}}{q^2}\Gamma_0\right)^{(-1)^p/2}. \tag{3.116}$$

Note that for a manifold with a single connected component of volume V, Γ_0 is the 1×1 matrix V. This generalizes the volume correction that appears in the path integral for Maxwell theory in refs. [263, 237] and is essential for unitarity.

The torsion part $T^k := \text{Tor}\, H^k(M,\mathbb{Z})$ of the cohomology groups is perhaps less familiar, but will play an important role in the duality in the most general case. When we integrate over the space of flat p-forms we must divide by the large gauge transformations. Taking the quotient by the discrete subgroup T^p simply amounts to dividing the partition function by the number of elements $|T^p|$. To account for torsion-valued $p-1$-form large gauge transformations we must multiply by $|T^{p-1}|$ and so on, resulting in the contribution

$$Z_{\text{tors}} = |T^p|^{-1}|T^{p-1}|\cdots|T^0|^{(-1)^{p+1}} \tag{3.117}$$

to the partition function.

Finally we return to the instanton contribution, i.e. gauge connections that cannot be expressed as $F = dA$. The Bianchi identity identifies them as elements

of the cohomology group

$$\mathscr{F} \in 2\pi\, H^{p+1}(M, \mathbb{Z}). \tag{3.118}$$

Again, we can split the cohomology group into its free and torsion parts:

$$H^{p+1}(M, \mathbb{Z}) = \text{Free}\, H^{p+1}(M, \mathbb{Z}) \oplus \text{Tor}\, H^{p+1}(M, \mathbb{Z}) = \mathbb{Z}^{b_{p+1}} \oplus T^{p+1}. \tag{3.119}$$

Elements of Free $H^{p+1}(M, \mathbb{Z})$ are equivalence classes of $(p+1)$-forms, from which we choose $\mathscr{F}$ to be the unique harmonic representative. Elements of the torsion subgroup are associated with vanishing field strengths, and hence these instantons simply lead to an overall factor of $|T_{p+1}|$. (These correspond to flat connections that have nontrivial holonomy around certain noncontractible p-surfaces, and yet do not come from harmonic p-forms.) Thus the full sum over instantons is given by

$$Z_{\text{inst}} = |T^{p+1}| \sum_{\mathscr{F} \in \mathbb{Z}^{b_{p+1}}} e^{-I[\mathscr{F}]}. \tag{3.120}$$

Summarizing, we find that the partition function of p-form gauge theory is

$$Z_p = \prod_{k=0}^{p} \left(\frac{\det\left(\frac{2\pi\mu^{2(p-k+1)}}{q^2}\Gamma_k\right)^{1/2}}{|T^k| \det(\tilde{\Delta}_k)^{(p-k+1)/2}} \right)^{(-1)^{p-k}} |T^{p+1}| \sum_{\mathscr{F} \in \mathbb{Z}^{b_{p+1}}} e^{-I[\mathscr{F}]}. \tag{3.121}$$

Except for an anomaly in even dimensions, the factors of μ cancel between numerator and denominator. This anomaly fixes the units in (3.100) but otherwise plays no role. We will not keep the factors of μ explicit in the remainder of this section, deferring the details to the end of appendix 3.3.6.

Next we compute the change in the effective action under electromagnetic

duality.

Electromagnetic duality

With the expression (3.121) it is a straightforward exercise to compute the relation between the dual partition functions. We will use zeta function regularization. The partition function of the electromagnetic dual is (3.103) with $p \to \tilde{p} = D - p - 2$ and $q \to \tilde{q} = 2\pi/q$, which is equivalent to the replacement

$$\tilde{F} = \left(\frac{\tilde{q}}{q}\right) \star F. \tag{3.122}$$

Poincaré duality and Poisson summation are the only tools needed to compute the ratio $Z_p/\tilde{Z}_{\tilde{p}}$, a task we defer to appendix 3.3.6.

First we isolate the oscillator contribution to the ratio. The ratio of oscillator partition functions is

$$\frac{Z_{\text{osc}}}{\tilde{Z}_{\text{osc}}} = \left[\prod_{k=0}^{D} (\det \Delta_k)^{k(-1)^{k+1}/2}\right]^{(-1)^{p+1}} . \tag{3.123}$$

This expression is related to a quantity known as the Ray-Singer analytic torsion [266, 249],

$$\tau_{\text{RS}} = \prod_{k=0}^{D} (\det \Delta_k)^{k(-1)^{k+1}/2} . \tag{3.124}$$

It plays a role in abelian Chern-Simons theory as the magnitude squared of the partition function [246] but was originally defined as an analytic analog to a combinatorial invariant called Reidemeister torsion, which we will encounter soon.

When D is even, $\tau_{\text{RS}} = 1$ by Poincaré duality and so the ratio of oscillator

partition functions is

$$\log \frac{Z_{\rm osc}}{\tilde{Z}_{\rm osc}} = \begin{cases} 0 & D = 2n \\ (-1)^{p+1} \ \log \tau_{\rm RS} & D = 2n+1. \end{cases} \tag{3.125}$$

This is the result obtained by Schwarz and Tyupkin [243], who considered only the oscillator modes. Note in particular that the contribution to the anomaly vanishes in even D.

The rest of $Z_p/\tilde{Z}_{\tilde{p}}$ comes from the zero modes and instantons. Making the simplifying assumption that the torsion subgroups of $H^k(M,\mathbb{Z})$ are trivial, they contribute

$$\frac{Z_{\rm zero}Z_{\rm inst}}{\tilde{Z}_{\rm zero}\tilde{Z}_{\rm inst}} = \left[\left(\frac{2\pi}{q^2} \right)^{\chi} \prod_{k=0}^{D} \det \left(\Gamma_k \right)^{(-1)^k} \right]^{(-1)^p/2} \tag{3.126}$$

which follows from Poincaré duality and the relation $\chi = \sum (-1)^k b_k$ between the Euler characteristic and the Betti numbers. This ratio is related to the Reidemeister torsion:

$$\tau_{\rm Reid} = \prod_{k=0}^{D} \det \Gamma_k^{(-1)^k/2}. \tag{3.127}$$

This flavor of torsion was invented in 1935 to classify lens spaces [250], which have the same homotopy groups but are not homeomorphic. It is actually the oldest non-homotopy invariant [267], and is defined in terms of chain complexes on M. Ray and Singer defined their torsion (3.124) as an analytic analog in the early 1970s. Cheeger and Müller independently proved a few years later that the two are actually equal (up to the torsion subgroups T^k to be discussed), culminating

in the Cheeger-Müller theorem [268, 248, 269]

$$\tau_{\rm RS} = \tau_{\rm Reid}. \tag{3.128}$$

This equation is the key to the triviality of the odd-dimensional anomaly. Thus

$$\frac{Z_{\rm zero}Z_{\rm inst}}{\tilde{Z}_{\rm zero}\tilde{Z}_{\rm inst}} = \begin{cases} (-1)^{p+1}\ \chi(M)\log\sqrt{\frac{q}{\tilde{q}}} & D = 2n \\ (-1)^{p}\ \log\tau_{\rm Reid} & D = 2n+1, \end{cases} \tag{3.129}$$

and the duality anomaly of a p-form theory on a D-manifold M is

$$\log\frac{Z_p}{\tilde{Z}_{\tilde{p}}} = \begin{cases} (-1)^{p+1}\ \chi(M)\log\sqrt{\frac{q}{\tilde{q}}} & D = 2n \\ (-1)^{p+1}\log\frac{\tau_{\rm RS}}{\tau_{\rm Reid}} = 0, & D = 2n+1. \end{cases} \tag{3.130}$$

In odd D the zero mode and instantons cancel the oscillator contribution to the ratio and so the duality is exact. Our even-dimensional result agrees with Witten's computation [244] of the duality anomaly of $D = 4$ Maxwell theory; in appendix 3.3.7 we reproduce the θ-dependence and discuss a phase. Only the zero modes and instantons contribute to the anomaly, which is why Schwarz and Tyupkin found an exact duality. It is somewhat ironic that the zero modes and instantons trivialize the odd-dimensional duality instead.

The quantity $\frac{q}{\tilde{q}}$ in (3.130) has mass dimension $2(p+1) - D$ and so must be accompanied by a dimensionful factor; this is furnished by the parameter μ that we had to introduce in order to make the measure (3.109) dimensionless. We will see in appendix 3.3.6 that an anomaly in rescaling μ out of the functional

determinants multiplies (3.130) by an extra term

$$(-1)^p \chi(M) \log \mu^{p+1-D/2}, \tag{3.131}$$

after application of the McKean-Singer formula [270]. This precisely fixes the units.

Last we consider the contribution from any torsion subgroups $T^k \subset H^k(M, \mathbb{Z})$. In the presence of nontrivial T^k, Cheeger [248] found that the relation (3.128) is modified to

$$\frac{\tau_{\rm RS}}{\tau_{\rm Reid}} = \prod_{k=0}^{D} |T^k|^{(-1)^{k+1}} \tag{3.132}$$

and so to maintain duality invariance we must show that their effect on the ratio of partition functions is to divide by this factor. In appendix 3.3.6 we show that their contribution to the ratio of partition functions is

$$\frac{Z_{p,\rm tors}}{\tilde{Z}_{\tilde{p},\rm tors}} = \left[\prod_{k=0}^{D} |T^k|^{(-1)^{k+1}} \right]^{(-1)^p} \tag{3.133}$$

which is trivial in even dimensions and cancels the non-torsion contribution in odd dimensions. In fact, the vanishing of the odd-dimensional anomaly is equivalent to Cheeger's refinement of the Cheeger-Müller theorem.

We find it rather surprising to find any application to physics in this somewhat obscure relation between quantities from different branches of mathematics.

3.3.3 Trace of the stress tensor

Since the effective action in even dimensions is not invariant under a duality transformation, it is natural to wonder whether the duality shifts the value of any other observables. One natural choice of observable is the stress tensor, obtained by varying the effective action with respect to the metric:

$$T_{ab} = -2\frac{\delta}{\delta g^{ab}} \log Z \tag{3.134}$$

In our regulator scheme, it is easy to see that T_{ab} will *not* be affected by a duality transformation: the duality anomaly is simply a finite number times a topological invariant χ, which is independent of the metric. Therefore, T_{ab} is the same in both theories.

Nevertheless, there has been a considerable amount of confusion about this topic in the literature, due to an apparent discrepancy in the trace anomaly between the dual theories. In order to clarify this issue, we first remind the reader that we needed three dimensionful parameters in order to define the quantum partition function (3.104) (apart from any geometric parameters of M):

$\boldsymbol{\mu}$: the dimensional measure factor appearing in each mode of the path integral;

$\boldsymbol{\Lambda}$: an additional parameter present in some regularization schemes (such as an ultraviolet momentum cutoff or lattice scale);

$\boldsymbol{q}$: the fundamental charge, which is dimensionful when $D \neq 2p + 2$.

The first two quantities, Λ and μ, do not appear in the classical theory but are needed to make the quantum theory well-defined.

Note that some regularization schemes, such as the zeta-function regularization we have been using, introduce only a single dimensionful scale and are effectively identifying $\Lambda = \mu$. In this case even a single mode in the partition function gives rise to a logarithmic divergence as $\Lambda \to \infty$. While consistent, this can lead to confusing outcomes like apparent UV divergences in theories with no local degrees of freedom, such as $D = 2$ Maxwell theory [271]. In such a scheme, our duality anomaly will *appear* to take the form of a logarithmic divergence proportional to χ, although it really comes from the zero mode integrals.

This point is important because the trace anomaly is frequently calculated using the log divergences of the theory. However, it is necessary to consider the dependence on the logs of all three kinds of dimensionful parameters to get the correct result. Thus, suppose we have a compact manifold $M(R)$ with "radius" R, where $M(R)$ is given by acting on $M(1)$ with a uniform scaling factor $\Omega = R$. Then the dependence of the effective action on $\log R$ is given by

$$\frac{\partial}{\partial \log R} \log Z_p = -\int_M T. \tag{3.135}$$

Dimensional analysis now says that this term can be calculated if you know how $\log Z$ scales when you simultaneously adjust the mass scales of Λ, μ, and q:

$$\frac{\partial}{\partial \log R} \log Z_p = \left[\frac{\partial}{\partial \log \Lambda} + \frac{\partial}{\partial \log \mu} + \left(p + 1 - \frac{D}{2}\right)\frac{\partial}{\partial \log q}\right] \log Z_p \tag{3.136}$$

The simplest case is that of a conformal p-form field in $D = 2p+2$ dimensions, where the theory is dual to another p-form of the same rank. In this case, q is dimensionless. Since the theory is conformal, $T = 0$ classically, q is dimensionless,

and any nonzero value of T must come from quantum anomalies. After setting $\mu = \Lambda$, balancing of logarithms (i.e. demanding that their arguments be dimensionless) requires that any dependence on $\log \Lambda$ must match with the dependence of $\log Z$ on a local conformal rescaling $g'_{ab} = \Omega^2 g_{ab}$ of the metric. Hence the trace $T = T_{ab} g^{ab}$ may be calculated from the log divergences of the theory. However, $p = \tilde{p}$, so the log divergences of the two theories are identical, and the duality anomaly (3.130) is merely a finite function of q. Note that T is determined locally in this conformal case.

Things are more subtle when $D \neq 2p+2$. In this case, the action (3.103) is not conformal, even classically. Thus, in general T may depend in a complicated and nonlocal way on the state of the fields. The classical theory is *almost* invariant under the global scale-invariance (Ω = constant) generated by (3.135), but even this symmetry is partially broken by flux quantization effects that depend on q.

Now we discuss electromagnetic duality. In even dimensions, (3.130) and (3.131) tell us that the duality anomaly coming from zero modes is proportional to

$$\Delta \log Z \propto \log \left(\frac{q}{\mu^{p+1-D/2}} \right) \tag{3.137}$$

This logarithm is already balanced, and hence it does not produce any $\log R$ dependence. Therefore, the integrated trace is unaffected by the duality, consistent with our claim above.

This conclusion is essentially the same as that of [242], who argued that the *total* trace of the stress tensor is duality invariant, even though the amount attributable to varying $\log \Lambda$[8] depends on the duality frame. However, they regu-

[8] which they confusingly refer to as the "trace anomaly"

lated their zero modes by inserting a small mass, while we consider a U(1) gauge field whose IR divergences are regulated by finite-q effects.

In the next section we will consider the effects of the duality anomaly on the entanglement entropy S. Unlike T_{ab}, S is sensitive to topological terms, and thus we will find a nonzero shift under duality.

3.3.4 Entanglement anomalies

We now derive the entanglement anomaly, i.e. the difference in entanglement entropies between the p-form theory and its dual. We use the replica trick, which enables us to compute the difference of entropies using the results of § 3.3.2 in conjunction with the definition of the entanglement anomaly (3.101). The procedure is straightforward: we substitute the replica manifold $M = M_A^{(n)}$ into (3.130), then determine the n-dependence in order to compute the entanglement anomaly (3.102). As described in the previous section the anomaly vanishes in odd spacetime dimension.

The anomaly

The change in vacuum entanglement of a region A under duality (3.101) depends on the ratio of replica partition functions, i.e. the duality anomaly (3.130) of the theory on the replica manifold $M_A^{(n)}$. This is determined by the Euler characteristic of $M_A^{(n)}$, which follows immediately from its cut-and-paste construction: $M_A^{(n)}$ consists of n copies of $M \setminus A$, glued together along n copies of $A \setminus \partial A$, all glued to a single copy of the entangling surface ∂A. Since each piece is disjoint

their Euler characters just add:

$$\begin{aligned}\chi(M_A^{(n)}) &= n\chi(M \setminus A) + n\chi(A \setminus \partial A) + \chi(\partial A) \\ &= n\chi(M \setminus \partial A) + (1-n)\chi(\partial A).\end{aligned} \tag{3.138}$$

Using

$$\log \frac{Z(M_A^{(n)})}{\tilde{Z}(M_A^{(n)})} = (-1)^{p+1}\ \chi\left(M_A^{(n)}\right) \log \sqrt{\frac{q}{\tilde{q}}} \tag{3.139}$$

together with the thermodynamic expression for the anomaly (3.101), we find that the change in the entanglement entropy of a region A in p-form Maxwell theory under electromagnetic duality is

$$\Delta S_A = (-1)^{p+1}\ \chi(\partial A) \log \sqrt{\frac{q}{\tilde{q}}}. \tag{3.140}$$

Here $\tilde{q} = 2\pi/q$ is the coupling of the dual gauge theory. We will argue that the entanglement anomaly arises physically from the global anomaly of a $p-1$-form edge mode theory living on the entangling surface; see § 3.3.4.

Now we discuss the universality of our results. A constant term in the entanglement entropy in even dimensions can usually be absorbed into a shift of the cutoff, and so our results may appear to depend on the choice of renormalization scheme. For simplicity consider the entanglement entropy of a ball-shaped region in $D = 4$ flat vacuum, whose general form is [208]

$$S = c_1 R^2 \Lambda^2 + c_2 \log(R\Lambda) + c_3 \tag{3.141}$$

where R is the radius of the sphere, Λ is the cutoff, and c_i are various constants.

In the absence of duality only c_2 is universal, as changes in $c_{1,3}$ can be absorbed into shifts of the cutoff. The entanglement entropy of the same region in the dual theory is

$$\tilde{S} = \tilde{c}_1 R^2 \tilde{\Lambda}^2 + \tilde{c}_2 \log(R\tilde{\Lambda}) + \tilde{c}_3. \tag{3.142}$$

Universality of the coefficient of the log guarantees $c_2 = \tilde{c}_2$.

In these calculations of entanglement, Λ is a physical inverse distance to the entangling surface and should be matched between the two theories in order to compare like quantities. If we rescale $\Lambda \to \Lambda' = \alpha\Lambda$ in the original theory, c_3 picks up a shift proportionate to c_2. In the dual theory we must do the same rescaling and so $\tilde{c}_3$ picks up a shift, but universality of the log coefficient guarantees that it matches the shift in c_3. Thus, the entanglement anomaly $\Delta c_3 = c_3 - \tilde{c}_3$ is independent of Λ.[9]

More generally, we expect that Δc_3 does not depend on the choice of (reasonable) regulator scheme for the theory. Recall that the entanglement anomaly arises purely from the zero modes and instantons of the theory, since (in even dimensions, where the anomaly can exist) the nonzero modes of the dual theories are in correspondence. Therefore, so long as (a) we use the same regulator for modes of the same wavelength on both sides of the duality, and (b) the regulator only affects UV divergent quantities, not zero modes or instantons, it follows that Δc_3 is a universal quantity.

[9] We are grateful to Mark Srednicki for pointing out the preceding argument.

With a θ term

We can also calculate the entanglement anomaly in the presence of a topological term in the action. For concreteness consider the theory with $p = 1$ and $D = 4$, with action

$$I = \frac{2\pi^2}{q^2}\left(\frac{1}{8\pi^2}\int_M F_{\mu\nu}F^{\mu\nu}\right) + \frac{i\theta}{2}\left(\frac{1}{8\pi^2}\int_M \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}\right). \tag{3.143}$$

We derive the duality relation in appendix 3.3.7. Here we just quote the result:

$$\tilde{Z} = Z\left(\frac{-1}{\tau}\right) = e^{i\pi\sigma/4}\bar{\tau}^{(\chi+\sigma)/4}\tau^{(\chi-\sigma)/4}Z(\tau). \tag{3.144}$$

where $\tau = \frac{\theta}{2\pi} + \frac{2\pi i}{q^2}$ and $\sigma = b_2^+ - b_2^-$ is the topological (as opposed to metric) signature of the manifold. The partition function transforms as a modular form up to a phase.

Before we can apply this result to the entanglement anomaly, we need to determine the signature of the replica manifold. The contribution of the phase in (3.144) to the entanglement anomaly is

$$\Delta S_{\text{phase}} = \left(-\frac{i\pi}{4}\right)(1 - n\partial_n)\sigma(M_A^{(n)})|_{n=1} \tag{3.145}$$

If this quantity is nonzero the phase would contribute an imaginary piece to the entanglement anomaly. Since entropy is a real quantity, this is only consistent if $\sigma(M_A^{(n)})$ is linear in n so that it does not contribute to the anomaly. It was shown in [272] that this is indeed the case, at least for $D = 4$: the signature is well-defined for manifolds with conical defects, and is linear in the replica number

n. We conjecture that this will also be the case for general D. While we do not have a complete proof, we note that the non-additivity of the signature is given by a result of [273], and appears to vanish by symmetry considerations.

Assuming that the signature contribution vanishes, we can now substitute $\chi(M_A^{(n)})$ and $\sigma(M_A^{(n)})$ into (3.101) to find the generalization of (3.140):

$$\Delta S_A = -\frac{1}{4}\log\left[\left(\frac{\theta}{2\pi}\right)^2 + \left(\frac{2\pi}{q^2}\right)^2\right]\chi(\partial A). \tag{3.146}$$

This analysis extends to the case where D is a multiple of 4 and $p+1 = D/2$. A case not covered by this analysis is $D = 2$ Maxwell theory with a $\theta\int F$ term; in appendix 3.3.7 we show that the entanglement anomaly is independent of such a θ term.

Invariance of the thermal entropy

The results of the previous sections imply that thermal entropy does not change under electromagnetic duality. This is reassuring, since the total number of degrees of freedom should be duality invariant.

We compute thermal entropies using ordinary thermodynamics. The first law of thermodynamics $S_{\text{therm}} = \beta\langle E\rangle + \log Z_{\text{therm}}$ relates thermal entropy to the thermal partition function $Z_{\text{therm}} = \operatorname{Tr} e^{-\beta H}$. When the theory lives on a spatial manifold Σ, the thermal partition function is equal to the path integral on $M^{(\beta)} = \Sigma \times S^1_\beta$. We rewrite the first law as

$$S_{\text{therm}}(\beta) = (1 - \beta\partial_\beta)\log Z(M^{(\beta)}) \tag{3.147}$$

and specialize again to the case of p-form theories. Under electromagnetic duality, the change in thermal entropy is

$$\Delta S_{\rm therm}(\beta) = (1 - \beta\partial_\beta) \log \frac{Z_p(M^{(\beta)})}{\tilde{Z}_{\tilde{p}}(M^{(\beta)})}. \tag{3.148}$$

We can compute the right hand side using (3.130) with $M = M^{(\beta)}$. This gives

$$\log \frac{Z_{\rm therm}}{\tilde{Z}_{\rm therm}} = \begin{cases} (-1)^{p+1}\ \chi(M^{(\beta)}) \log \sqrt{\frac{q}{\tilde{q}}} & D = 2n \\ 0 & D = 2n+1. \end{cases} \tag{3.149}$$

The Euler character of a product manifold $A \times B$ satisfies $\chi(A \times B) = \chi(A) \cdot \chi(B)$, so the Euler character of $M^{(\beta)} = \Sigma \times S^1_\beta$ vanishes and hence $\Delta S_{\rm therm}(\beta) = 0$ in even dimensions.

Thus for any p and D

$$S_{\rm therm} = \tilde{S}_{\rm therm}. \tag{3.150}$$

Edge modes

In theories with gauge symmetry there is a question of how to define the entanglement entropy. For example, [216] proposed a number of inequivalent definitions in terms of the algebra of observables inside the entangling surface. Here we have adopted the definition of entanglement entropy via the replica trick. The replica trick requires no additional input, such as a choice of algebra, and so must single out a particular definition. In Refs. [257, 217] it was shown that the replica trick coincides with the "extended Hilbert space" definition of entanglement entropy for 1-form gauge fields. In the case of abelian gauge fields, this coincides with what

Ref. [216] calls the "electric" definition of entanglement entropy and in condensed matter is sometimes called the "rough edge" [259].

Analysis of these replica partition functions $Z(M_A^{(n)})$ [217] yields a decomposition into a bulk and an edge piece:

$$Z(M_A^{(n)}) = Z_{\text{bulk}} Z_{\text{edge}} \tag{3.151}$$

where Z_{bulk} describes degrees of freedom on $M \setminus \partial A$ and Z_{edge} describes degrees of freedom on the entangling surface ∂A.

The edge mode partition function is

$$Z_{\text{edge}} = \int \mathcal{D}E_\perp e^{-I_{\text{cl}}(E_\perp)}; \tag{3.152}$$

the sum is over configurations of the normal electric field at the entangling surface and the exponent is the action of a classical solution with corresponding boundary configuration — the action is itself a boundary term. The reduced density matrix thus splits into superselection sectors labelled by $E_\perp$: $\rho = \oplus p(E_\perp)\rho_{E_\perp}$. This gives rise to a Shannon term $-\sum p(E_\perp)\log p(E_\perp)$ in the entanglement entropy [274].

In this case the entropy of the edge modes is given by the log of the partition function of a ghost scalar confined to the entangling surface:

$$S_{\text{edge}} = -\log Z_{p=0}(\partial A). \tag{3.153}$$

This suggests that the generalization of the edge mode entropy for a p-form theory is the partition function of a ghost $(p-1)$-form theory on the codimension-2

entangling surface.

Note that when the p-form theory in D dimensions is conformal, the $(p-1)$-form theory in $D-2$ dimensions is also conformal. In this case there is a universal logarithmic divergence in the entanglement entropy of a sphere related to the conformal anomaly of the theory [275, 213]. For example, the sphere entropy of Maxwell theory in $D = 4$, contains a contribution from a ghost scalar in $D = 2$ which is necessary in order to obtain agreement with the conformal anomaly [257, 276, 217].

The results of the present paper provide further evidence for this edge mode theory. The anomaly in the entropy (3.140) is precisely the duality anomaly of a ghost $(p-1)$-form confined to the entangling surface. Thus the universal differences in the electric versus magnetic prescriptions for the entanglement entropy appears to be captured in the electromagnetic duality anomaly of the edge mode theory. It would be interesting to either confirm or refute this conjecture, for example by calculating the edge mode contribution to the logarithmic divergence of the sphere entanglement entropy in conformal p-form theories. In doing so one ought to pay close attention to the zero modes of the edge system [257, 217, 260].

3.3.5 Discussion

Duality is a rich subject and we have only explored simple, abelian examples. It would be interesting to extend our results to other dualities. In most cases this will not be easy: generically one cannot compute the partition function, and even when one can, the replica manifold may break symmetries (such as supersymmetry) that enabled the computation in the first place. However, any

tractable calculations of entanglement in dual theories would be amenable to an anomaly analysis akin to ours.

It would also be interesting to understand the lattice analogue of the phenomenon we have described. There is an ambiguity in how to cut up the lattice in calculations of entanglement: one can put the cut in the middle of a plaquette, or on a vertex, or on an edge. When such a theory enjoys an electromagnetic duality, its dual lives on the dual lattice and so inherits a different prescription for the entropy. In self-dual theories the entanglement anomaly should record this dependence on the choice of prescription.

Acknowledgements

We are very grateful to Max Metlitski for discussions of many kinds of torsion, to Mark Srednicki for pointing out the universal nature of the anomaly, and to Tarun Grover and Brayden Ware for discussions of the implications for the lattice. We would also like to acknowledge the hospitality of the Perimeter Institute during several phases of this project. BM is supported by NSF Grant PHY13-16748. AW is supported by the Institute for Advanced Study, the Martin A. and Helen Chooljian Membership Fund, the Raymond and Beverly Sackler Foundation, and NSF grant PHY-1314311.

3.3.6 Appendix A: Electromagnetic duality

In this appendix we derive (3.130). For simplicity we first consider the case where the torsion subgroups of the cohomologies $H^k(M, \mathbb{Z})$ are trivial, also deferring discussion of the dimensionful factors μ appearing in the measure (3.109) to

the end.

As described in § 3.3.2, the partition function of p-form Maxwell theory on a manifold M decomposes as

$$Z_p = Z_{\rm osc} Z_{\rm zero} Z_{\rm inst} \tag{3.154}$$

where $Z_{\rm osc}$, $Z_{\rm zero}$ and $Z_{\rm inst}$ can be found in (3.111), (3.116) and (3.120). In the dual $\tilde{p}$ form theory, the oscillator determinants $\tilde{Z}_{\rm osc}$ and zero mode factors $\tilde{Z}_{\rm zero}$ are simply obtained from (3.111) and (3.116) by replacing p with $\tilde{p} = D - p - 2$.

Relating the instanton partition functions is slightly more involved. The trick is to do a Poisson summation and make use of Poincaré duality. The instanton partition function of the original theory is

$$\begin{aligned} Z_{\rm inst} &= \sum_{\mathscr{F} \in H^{p+1}(M,\mathbb{Z})} e^{-I_{\rm cl}(\mathscr{F})} = \sum_{\vec{m} \in \mathbb{Z}^{b_{p+1}}} e^{-\frac{1}{2}\left(\frac{2\pi}{q}\right)^2 \vec{m}\cdot\Gamma_{p+1}\cdot\vec{m}} \\ &= \det\left(\frac{2\pi\Gamma_{p+1}}{q^2}\right)^{-1/2} \cdot \sum_{\vec{n} \in \mathbb{Z}^{b_{\tilde{p}+1}}} e^{-\frac{1}{2}\left(\frac{2\pi}{\tilde{q}}\right)^2 \vec{n}\cdot\Gamma_{\tilde{p}+1}\cdot\vec{n}} \\ &= \det\left(\frac{2\pi\Gamma_{p+1}}{q^2}\right)^{-1/2} \cdot \sum_{\star\mathscr{F} \in H^{\tilde{p}+1}(M,\mathbb{Z})} e^{-\tilde{I}_{\rm cl}(\star\mathscr{F})} \\ &= \det\left(\frac{2\pi\Gamma_{p+1}}{q^2}\right)^{-1/2} \cdot \tilde{Z}_{\rm inst}. \end{aligned} \tag{3.155}$$

In the second line we used

$$\tilde{q} = \frac{2\pi}{q} \tag{3.156}$$

and Poincaré duality, which implies $\Gamma_k = \Gamma_{D-k}^{-1}$.[10]

[10]More precisely, duality implies only that $\Gamma_k = \Gamma_{D-k}^{-1}$ up to a matrix that is invertible over the integers (and therefore has unit determinant). We can set this matrix to the identity by a

Now we compute the ratio of partition functions, i.e. the duality anomaly. Using (3.111), (3.116) and (3.155),

$$\frac{Z_{\rm osc}Z_{\rm zero}Z_{\rm inst}}{\tilde{Z}_{\rm osc}\tilde{Z}_{\rm zero}\tilde{Z}_{\rm inst}} = \frac{\prod_{k=0}^{p}\left[\det(\Delta_k)^{(-1)^{p+1-k}(p+1-k)/2}\cdot\det\left(\frac{2\pi}{q^2}\Gamma_{p-k}\right)^{(-1)^k/2}\right]\det\left(\frac{2\pi}{q^2}\Gamma_{p+1}\right)^{-1/2}}{\prod_{k=0}^{\tilde{p}}\left[\det(\Delta_k)^{(-1)^{\tilde{p}+1-k}(\tilde{p}+1-k)/2}\cdot\det\left(\frac{2\pi}{\tilde{q}^2}\Gamma_{\tilde{p}-k}\right)^{(-1)^k/2}\right]}. \tag{3.157}$$

First we compute the ratio of oscillator determinants, which was first calculated by [243]. It is convenient to decompose $\Delta_k = \delta_k d_k + d_{k-1}\delta_{k-1}$[11]; the spectra of $\delta_k d_k$ and $d_k\delta_k$ agree up to zero modes, so $\det(\delta_k d_k) = \det(d_k\delta_k)$. Then

$$\begin{aligned} Z_{\rm osc} &= \prod_{k=0}^{p}\det(\delta_k d_k)^{(-1)^{p+1-k}(p+1-k)/2}\ \prod_{k=1}^{p}\det(d_{k-1}\delta_{k-1})^{(-1)^{p+1-k}(p+1-k)/2} \\ &= \prod_{k=0}^{p}\det(\delta_k d_k)^{(-1)^{p+1-k}/2}. \end{aligned} \tag{3.158}$$

Letting $E_k = \det(\delta_k d_k)$, the ratio of oscillator partition functions is

$$\begin{aligned} \frac{Z_{\rm osc}}{\tilde{Z}_{\rm osc}} &= \prod_{k=0}^{p}E_k^{(-1)^{p+1-k}/2}\ \prod_{k=0}^{\tilde{p}}E_k^{(-1)^{\tilde{p}-k}/2} \\ &= \prod_{k=0}^{p}E_k^{(-1)^{p+1-k}/2}\ \prod_{k=0}^{D-p-2}E_{D-k-1}^{(-1)^{D-p-2-k}/2} = \left[\prod_{k=0}^{D-1}E_k^{(-1)^{k-1}/2}\right]^{(-1)^p}. \end{aligned} \tag{3.159}$$

This expression is a power of the Ray-Singer analytic torsion

$$\tau_{\rm RS} = \prod_{k=0}^{D-1}E_{D-k-1}^{(-1)^k/2} = \prod_{k=0}^{D}\left(\det\Delta_k\right)^{k(-1)^{k+1}/2}. \tag{3.160}$$

choice of basis.

[11] For $k = 0$ or D, only the nontrivial term contributes.

In even D, $\tau_{\rm RS} = 1$ by Poincaré duality ($E_k = E_{D-k-1}$). Comparing (3.159) and (3.160), the ratio of oscillator partition functions is

$$\log \frac{Z_{\rm osc}}{\tilde{Z}_{\rm osc}} = \begin{cases} 0 & D = 2n \\ (-1)^{p+1} \ \log \tau_{\rm RS} & D = 2n+1 \end{cases} \tag{3.161}$$

which was the result of [243].

Next up are the zero modes and instantons. From (3.157) we get

$$\begin{aligned} \frac{Z_{\rm zero} Z_{\rm inst}}{\tilde{Z}_{\rm zero} \tilde{Z}_{\rm inst}} &= \prod_{k=0}^{p} \det \left(\frac{2\pi}{q^2} \Gamma_k \right)^{(-1)^{p-k}/2} \det \left(\frac{2\pi}{q^2} \Gamma_{p+1} \right)^{-1/2} \prod_{k=0}^{\tilde{p}} \det \left(\frac{\tilde{q}^2}{2\pi} \Gamma_k^{-1} \right)^{(-1)^{\tilde{p}-k}/2} \\ &= \prod_{k=0}^{p+1} \det \left(\frac{2\pi}{q^2} \Gamma_k \right)^{(-1)^{p-k}/2} \prod_{k=0}^{D-p-2} \det \left(\frac{2\pi}{q^2} \Gamma_{D-k} \right)^{(-1)^{D-p-2-k}/2} \\ &= \left[\prod_{k=0}^{D} \left(\frac{2\pi}{q^2} \right)^{(-1)^k b_k} \det \left(\Gamma_k \right)^{(-1)^k} \right]^{(-1)^p/2} \\ &= \left[\left(\frac{2\pi}{q^2} \right)^{\chi} \prod_{k=0}^{D} \det \left(\Gamma_k \right)^{(-1)^k} \right]^{(-1)^p/2} \end{aligned} \tag{3.162}$$

where we used Poincaré duality in the second line, then zeta-function regularization in the third to pull out the charge factors.[12] In even $D = 2n$, the $\det \Gamma$s cancel pairwise, while the ratio in odd dimensions is a power of the Reidemeister torsion:

$$\tau_{\rm Reid} = \prod_{k=0}^{D} \det \Gamma_k^{(-1)^k/2}. \tag{3.163}$$

[12] We define our determinants using zeta-function regularization, $\det \Delta = e^{-\zeta'_\Delta|_{s=0}}$, and so $\det(a\Delta) = a^{\zeta_\Delta|_{s=0}} \det \Delta$ for scalar a. The fact that $\zeta_\Delta|_{s=0} = -\dim \ker \Delta$ [277] (up to a term in even D that can be absorbed into a local counterterm [217]) then implies the equality in the third line of (3.162).

Rewriting (3.162) as

$$\frac{Z_{\rm zero}Z_{\rm inst}}{\tilde{Z}_{\rm zero}\tilde{Z}_{\rm inst}} = \begin{cases} (-1)^{p+1}\ \chi(M)\log\sqrt{\frac{q}{\tilde{q}}} & D=2n \\ (-1)^{p}\ \log\tau_{\rm Reid} & D=2n+1, \end{cases} \tag{3.164}$$

and using the Cheeger-Müller theorem

$$\tau_{\rm RS} = \tau_{\rm Reid}, \tag{3.165}$$

the duality anomaly of a p-form theory on a D-manifold M is

$$\log\frac{Z_p}{\tilde{Z}_{\tilde{p}}} = \begin{cases} (-1)^{p+1}\ \chi(M)\log\sqrt{\frac{q}{\tilde{q}}} & D=2n \\ (-1)^{p+1}\log\frac{\tau_{\rm RS}}{\tau_{\rm Reid}} = 0, & D=2n+1. \end{cases} \tag{3.166}$$

The quantity $q/\tilde{q}$ has mass dimension $2(p+1)-D$. However, the dimensionful factor μ in the measure (3.109) renders the argument of the log dimensionless, after proper consideration of an anomaly in zeta-function regularization discussed at the end of this appendix.

Next we consider the possibility that some of the cohomology groups $H^k(M,\mathbb{Z})$ have nontrivial torsion subgroups.[13] The partition functions for this more general case were explained in § 3.3.2, see (3.117) and (3.120). Separating out the contribution of these factors as

$$Z_p = Z_{p,\rm tors}Z_{p,\rm free} \tag{3.167}$$

[13]We apologize to the reader for introducing a third kind of torsion.

where $Z_{p,\mathrm{tors}}$ contains the contribution of all the torsion subgroups, we have

$$Z_{p,\mathrm{tors}} = \prod_{k=0}^{p+1} |T^{p+1-k}|^{(-1)^k} \tag{3.168}$$

where $|T^k|$ is the size of the torsion subgroup of $H^k(M,\mathbb{Z})$.

We need the duality relation for torsion subgroups to proceed further. It is common mathematical knowledge [278] that $H^k(M,\mathbb{Z}) \cong (H_k/T_k) \oplus T_{k-1}$ and so $T^k = T_{k-1}$. Together with Poincaré duality ($H^k \cong H_{D-k}$) this implies $T^k \cong T^{D-k+1}$. Using (3.168), we find

$$\begin{aligned}
\frac{Z_{p,\mathrm{tors}}}{\tilde{Z}_{\tilde{p},\mathrm{tors}}} &= \frac{\prod_{k=0}^{p+1} |T^{p+1-k}|^{(-1)^k}}{\prod_{k=0}^{\tilde{p}+1} |T^{\tilde{p}+1-k}|^{(-1)^k}} \\
&= \prod_{k=0}^{p+1} |T^{p+1-k}|^{(-1)^k} \prod_{k=0}^{D-p-1} |T^{p+2+k}|^{(-1)^{k+1}} \\
&= \left[\prod_{k=0}^{D} |T^k|^{(-1)^{k+1}} \right]^{(-1)^p}
\end{aligned} \tag{3.169}$$

which is equal to 1 by Poincaré duality in even D. The odd-D duality relation becomes

$$\frac{Z_p}{\tilde{Z}_{\tilde{p}}} = \left(\frac{\tau_{\mathrm{RS}}}{\tau_{\mathrm{Reid}}} \right)^{(-1)^{p+1}} \left[\prod_{k=0}^{D} |T^k|^{(-1)^{k+1}} \right]^{(-1)^p} . \tag{3.170}$$

Thus it appears at first as if we have recovered a nontrivial anomaly. However, Cheeger [248] showed that the relation between Ray-Singer and Reidemeister tor-

sion is modified in the presence of torsion subgroups[14]. The general relation is

$$\frac{\tau_{\rm RS}}{\tau_{\rm Reid}} = \prod_{k=0}^{D} |T^k|^{(-1)^{k+1}} \tag{3.171}$$

and so the odd-dimensional duality remains trivial even on manifolds whose integral cohomologies have nontrivial torsion subgroups. This completes our proof that the odd-dimensional duality anomaly vanishes.

Last we discuss the measure factors in (3.109) which correct the units in (3.166). We define our functional determinants using zeta-function regularization, $\det \Delta = e^{-\zeta'_\Delta|_{s=0}}$, and so $\det(a\Delta) = a^{\zeta_\Delta|_{s=0}} \det \Delta$ for scalar a. In even dimensions there is an anomaly[15] in $\zeta_\Delta|_{s=0}$: it is given by [277]

$$\zeta_{\Delta_k}|_{s=0} = \mathcal{A}_k - \dim \ker \Delta_k \tag{3.172}$$

where $\mathcal{A}_k = a^{(k)}_{D/2}$ is the coefficient of t^0 in the asymptotic expansion of the trace of the heat kernel $\mathrm{Tr}\, e^{-t\Delta_k}$ as $t \to 0$. They satisfy a Betti number-like relation to the Euler character, the McKean-Singer formula [270]:

$$\sum_{k=0}^{D} (-1)^k \mathcal{A}_k = \chi(M). \tag{3.173}$$

In odd dimensions $\mathcal{A}_k = 0$ for all k.

[14] We found appendix E of [237] particularly helpful in understanding these results.
[15] We apologize to the reader for introducing a third kind of anomaly.

The full partition with all the factors of μ was given in (3.121):

$$Z_p = \prod_{k=0}^{p} \left(\frac{\det\left(\frac{2\pi\mu^{2(p-k+1)}}{q^2}\Gamma_k\right)^{1/2}}{|T^k| \det\left(\frac{\Delta_k}{\mu^2}\right)^{(p-k+1)/2}} \right)^{(-1)^{p-k}} |T^{p+1}| \sum_{\mathscr{F}\in\mathbb{Z}^{b_{p+1}}} e^{-I[\mathscr{F}]}. \tag{3.174}$$

Scaling out the factors of μ in the functional determinants using (3.172), it is obvious that the nonanomalous pieces in the rescaling cancel with the factors of μ from the zero modes. However, we must deal with the anomalous piece. We find

$$Z_{\rm osc} Z_{\rm zero} = \prod_{k=0}^{p} \mu^{-\mathcal{A}_k(p+1-k)(-1)^{p+1-k}} \left(\frac{\det\left(\frac{2\pi}{q^2}\Gamma_k\right)^{1/2}}{|T^p| \det(\Delta_k)^{(p-k+1)/2}} \right)^{(-1)^{p-k}}. \tag{3.175}$$

Note that we have defined the dimensionful measure factors for the ghosts to be given by the same μ as appeared in the measure for the p-forms; this is part of our definition of the theory. We will also take the dual measure factors $\tilde{\mu}$ equal to the original μ. With these choices, the net effect of the anomaly is to multiply the ratio of partition functions (3.157) by

$$\prod_{k=0}^{p} \mu^{-\mathcal{A}_k(p+1-k)(-1)^{p+1-k}} \prod_{k=0}^{\tilde{p}} \mu^{\mathcal{A}_k(\tilde{p}+1-k)(-1)^{\tilde{p}+1-k}}. \tag{3.176}$$

The log of (3.157) picks up a piece

$$\sum_{k=0}^{D} (-1)^{p-k} \mathcal{A}_k (p+1-k) = (-1)^p \left[(p+1) \sum_{k=0}^{D} (-1)^k \mathcal{A}_k - \sum_{k=0}^{D} (-1)^k \frac{k}{2} \mathcal{A}_k - \sum_{k=0}^{D} (-1)^k \frac{(D-k)}{2} \mathcal{A}_k \right]$$

$$= (-1)^p \left(p + 1 - \frac{D}{2} \right) \chi(M) \tag{3.177}$$

multiplied by $\log \mu$. The LHS of (3.177) follows from (3.176) after using Poincaré duality, $\mathcal{A}_k = \mathcal{A}_{D-k}$. The first equality follows from Poincaré duality, the second from (3.173). Thus the net effect of the measure factors is to correct the even-dimensional duality relation to

$$\log \frac{Z_p}{\tilde{Z}_{\tilde{p}}} = (-1)^{p+1} \ \chi(M) \log \sqrt{\frac{q/\tilde{q}}{\mu^{2(p+1)-D}}}, \tag{3.178}$$

which is dimensionless as expected.

3.3.7 Appendix B: With a θ term

In this appendix we extend our analysis of the duality anomaly of p-form Maxwell theories to include a topological term. We focus on terms of the form

$$\theta \int_M F \wedge F \tag{3.179}$$

which only exist when $p + 1 = D/2$ and vanish unless D is a multiple of 4. For concreteness we describe the case of $p = 1$ duality with a θ term, which was studied by [244] (and earlier [279]); our analysis extends to all $D = 4n$.

We start with the action

$$I = \frac{2\pi^2}{q^2} \left(\frac{1}{8\pi^2} \int_M F_{\mu\nu} F^{\mu\nu} \right) + \frac{i\theta}{2} \left(\frac{1}{8\pi^2} \int_M \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} \right). \tag{3.180}$$

The terms in parentheses are the metric on the middle cohomology and the inter-

section form, respectively. The intersection form calculates the winding number of F and depends only on its cohomology. On a spin manifold, the intersection form is even, and the action is invariant under modular transformations.[16]

The θ term only affects the instanton contribution to the partition function. As described in § 3.3.2, the instanton contribution to the action is [17]

$$\begin{aligned} I_{\text{inst}} &= \frac{2\pi^2}{q^2}\vec{m}\cdot\Gamma\cdot\vec{m} + \frac{i\theta}{2}\vec{m}\cdot P\cdot\vec{m} \\ &= \vec{m}\cdot\Gamma\left(\frac{2\pi^2}{q^2} + \frac{i\theta}{2}S\right)\cdot\vec{m} \\ &= i\pi(\bar{\tau}\ \vec{m}_+\cdot\Gamma_+\cdot\vec{m}_+ - \tau\ \vec{m}_-\cdot\Gamma_-\cdot\vec{m}_-) \end{aligned} \tag{3.181}$$

In the first line we defined the matrix $P_{ij} = \int w_i \wedge w_j$. In the second line we related P to Γ via $P = \Gamma S$, where S is defined by $\star w_i = S_{ij}w_j$. This notation follows [263]. In the last line we used the orthogonal decomposition of the middle homology into self-dual and anti-self-dual forms, writing $\vec{m} = (\vec{m}_+, \vec{m}_-)$, where $\vec{m}_\pm$ are basis vectors for $H^2_\pm(M,\mathbb{Z})$ and have $b_2^\pm$ components each. In this basis S takes the form diag(1,-1).

Now we can get cracking. Using Poisson summation, the instanton partition function is

$$\begin{aligned} Z_{\text{inst}} &= \sum_{\vec{m}\in H^2(M,\mathbb{Z})} e^{-\vec{m}\cdot\Gamma\left(\frac{2\pi^2}{q^2}+\frac{i\theta}{2}S\right)\cdot\vec{m}} \\ &= \det\left(\frac{\pi}{\Gamma(2\pi^2/q^2 + i\theta S/2)}\right)^{1/2} \sum_{\vec{n}\in H_2(M,\mathbb{Z})} e^{-\vec{n}\cdot\Gamma^{-1}\left(\frac{2\pi^2}{q^2}+\frac{i\theta}{2}S\right)^{-1}\cdot\vec{n}\cdot\pi^2} \end{aligned} \tag{3.182}$$

[16] If the manifold does not admit a spin structure, the invariance is under the Hecke group generated by S and the T-transformation $\tau \to \tau + 2$ [280].

[17] In this section Γ refers to Γ_2.

Consider the determinant out front first:

$$\begin{aligned}\det\left(\frac{\pi}{\Gamma(2\pi^2/q^2+i\theta S/2)}\right)^{1/2} &= \det\Gamma^{-1/2}\,\det{}^{+}\left(\frac{2\pi^2}{q^2}+\frac{i\theta S}{2}\right)^{-1/2}\det{}^{-}\left(\frac{2\pi^2}{q^2}+\frac{i\theta S}{2}\right)^{-1/2}\\ &= \det\Gamma^{-1/2}\ \left(\frac{2\pi^2}{q^2}+\frac{i\theta}{2}\right)^{-b_2^+/2}\left(\frac{2\pi^2}{q^2}-\frac{i\theta}{2}\right)^{-b_2^-/2}\\ &= \det\Gamma^{-1/2}(i\bar\tau)^{-b_2^+/2}(-i\tau)^{-b_2^-/2}. \qquad (3.183)\end{aligned}$$

In the first line we used the decomposition $H^2 = H^2_+ \oplus H^2_-$.

Next, the exponent of (3.182), which is

$$\begin{aligned}\pi^2\vec{n}\cdot\Gamma^{-1}\left(\frac{2\pi^2}{q^2}+\frac{i\theta}{2}S\right)^{-1}\cdot\vec{n} &= \pi(\vec{n}_+\ \vec{n}_-)\Gamma^{-1}\begin{pmatrix}\frac{2\pi}{q^2}+\frac{i\theta}{2\pi} & 0\\ 0 & \frac{2\pi}{q^2}-\frac{i\theta}{2\pi}\end{pmatrix}^{-1}\begin{pmatrix}\vec{n}_+\\ \vec{n}_-\end{pmatrix}\\ &= i\pi\left[\left(\frac{-1}{\bar\tau}\right)\vec{n}_+\cdot\Gamma_+\cdot\vec{n}_+ - \left(\frac{-1}{\tau}\right)\vec{n}_-\cdot\Gamma_-\cdot\vec{n}_-\right] \qquad (3.184)\\ &= \tilde{I}_{\rm inst}(\star\mathscr{F}) \qquad (3.185)\end{aligned}$$

i.e. the action of an instanton in the dual theory. The full p-form partition function is

$$\begin{aligned}Z_p &= Z_{\rm osc}Z_{\rm osc}Z_{\rm inst}\\ &= Z_{\rm osc}\tilde{Z}_{\rm inst}\det\left(\frac{2\pi}{q^2}\Gamma_0\right)^{-1/2}\det\left(\frac{2\pi}{q^2}\Gamma_1\right)^{1/2}\det\Gamma_2^{-1/2}(i\bar\tau)^{-b_2^+/2}(-i\tau)^{-b_2^-/2}\\ &= Z_{\rm osc}\tilde{Z}_{\rm inst}\det\Gamma_0^{-1/2}\det\Gamma_1^{1/2}\det\Gamma_2^{-1/2}\left(\frac{2\pi}{q^2}\right)^{(b_1-b_0)/2}(i\bar\tau)^{-b_2^+/2}(-i\tau)^{-b_2^-/2}\end{aligned}$$

$$= \tilde{Z}_{\rm osc}\tilde{Z}_{\rm inst}\det\Gamma_0^{-1/2}\det\Gamma_1^{1/2}\det\Gamma_2^{-1/2}\,(\mathrm{Im}\;\tau)^{(b_1-b_0)/2}\,(i\bar{\tau})^{-b_2^+/2}(-i\tau)^{-b_2^-/2} \tag{3.186}$$

while the dual partition function is

$$\begin{aligned}\tilde{Z}_{\tilde{p}} &= \tilde{Z}_{\rm osc}\tilde{Z}_{\rm zero}\tilde{Z}_{\rm inst} \\ &= \tilde{Z}_{\rm osc}\tilde{Z}_{\rm inst}\det\left(\frac{2\pi}{\tilde{q}^2}\Gamma_0\right)^{-1/2}\det\left(\frac{2\pi}{\tilde{q}^2}\Gamma_1\right)^{1/2} \\ &= \tilde{Z}_{\rm osc}\tilde{Z}_{\rm inst}\det\Gamma_0^{-1/2}\det\Gamma_1^{1/2}\left(\mathrm{Im}\;\frac{-1}{\tau}\right)^{(b_1-b_0)/2}.\end{aligned} \tag{3.187}$$

It then follows from $\mathrm{Im}\;\frac{-1}{\tau} = \frac{\mathrm{Im}\;\tau}{\tau\bar{\tau}}$ and $\det\Gamma_2 = 1$ that[18]

$$\begin{aligned}\tilde{Z} = Z\left(\frac{-1}{\tau}\right) &= e^{i\pi\sigma/4}\bar{\tau}^{(b_2^+-b_1+b_0)/2}\tau^{(b_2^--b_1+b_0)/2}Z(\tau) \\ &= e^{i\pi\sigma/4}\bar{\tau}^{(\chi+\sigma)/4}\tau^{(\chi-\sigma)/4}Z(\tau).\end{aligned} \tag{3.188}$$

From eq. (3.188) we see that the partition function transforms as a modular form, up to a phase. This agrees with Witten's result for the duality anomaly [244] except for the phase, which also appears in the calculation of the anomaly by Alvarez and Olive in [280], and in Metlitski's calculation in [237]. It seems that the partition function is a modular form only up to this phase, whose presence allowed us to make a physical argument about the topological signature of the replica manifold in § 3.3.4.

While $\theta\int F\wedge F$ vanishes in 0-form theory in two dimensions, in 1-form theory there is a term $\theta\int F$ which plays a similar role to the θ term in $D = 4n$ dimensions

[18] Γ_2 is a positive matrix, and P^{-1} is an integer matrix (by Poincaré duality), so $|\det\Gamma_2| = 1$.

(i.e. it assigns a phase linear in n to instantons in the path integral). However, since it is linear in the field strength, it enters differently into the duality relation. The action

$$S = \frac{1}{2}\int F \wedge \star F + \frac{i\theta q}{2\pi}\int F \tag{3.189}$$

leads to the bundle sum

$$\begin{aligned} Z_{\text{inst}} &= \sum_{m\in\mathbb{Z}} e^{-\frac{1}{2}\left(\frac{2\pi}{q}\right)^2 m^2 + i\theta m} \\ &= \det\left(\frac{2\pi}{q^2}\Gamma_2\right)^{-1/2} \sum_{n\in\mathbb{Z}} e^{-\frac{1}{2}\left(\frac{2\pi}{\tilde{q}}\right)^2 \left(n+\frac{\theta}{2\pi}\right)^2}. \end{aligned} \tag{3.190}$$

The shift by θ in the lattice of integer charges in the dual theory is the Witten effect [281]. However, since it does not enter into the determinant, θ does not affect the duality anomaly in $D = 2$.

3.3.8 Appendix C: One-dimensional Maxwell theory

Maxwell theory in one spacetime dimension is a particularly trivial theory. It has only a vacuum state, so the canonical partition function is simply

$$Z_{\text{canonical}} = \sum_{\text{states}} e^{-\beta E} = 1. \tag{3.191}$$

We can use the sledgehammer forged in the body of this work to reproduce this trivial result. This exercise serves to clarify the role of the non-oscillator contributions in a simple example.

The oscillator partition function is

$$Z_{\text{osc}} = (\det \Delta_1)^{-1/2} \left[(\det \Delta_0)^{1/2}\right]^2 = (\det \Delta_0)^{1/2} \tag{3.192}$$

where the last equality follows from Poinaré duality. As usual, we will calculate the functional determinant using ζ-function regularization: $\det \Delta = e^{-\zeta'_\Delta(0)}$. Noting that the eigenfunctions f_n of the scalar Laplacian take the form $f_n(\theta) \sim e^{\frac{2\pi i n \theta}{\beta}}$, the relevant zeta function reads

$$\zeta_\Delta(s) = \sum_{n\neq 0} \frac{1}{\lambda_n^s} = 2\left[-\left(\frac{\beta}{2\pi}\right)^2\right] \zeta_R(2s) \tag{3.193}$$

where in the second equality we introduced the Riemann zeta function $\zeta_R(s) = \sum_{n\in\mathbb{Z}^+} \frac{1}{n^s}$.

Evaluation then yields

$$\zeta'_\Delta(0) = 4\zeta'_R(0) + 2\log\left[-\left(\frac{\beta}{2\pi}\right)^2\right] \zeta_R(0) = -\beta^2 \tag{3.194}$$

and so

$$Z_{\text{osc}} = \beta \neq Z_{\text{canonical}}. \tag{3.195}$$

However, the zero mode partition function is

$$\begin{aligned} Z_{\text{zero}} &= \det\left(\frac{2\pi}{q^2}\Gamma_1\right)^{1/2} \det\left(\frac{2\pi}{q^2}\Gamma_0\right)^{-1/2} \\ &= \det \Gamma_0^{-1} = \beta^{-1} \end{aligned} \tag{3.196}$$

where in the last line we used Poincaré duality. There are no instantons in this

simple scenario so we are done. We now see that

$$Z_{\text{osc}} \cdot Z_{\text{zero}} = 1 = Z_{\text{canonical}}, \tag{3.197}$$

i.e. the gauge theory partition function agrees with the canonical partition function only once the functional determinants are combined with the contribution from the zero modes.

3.4 Remarks on Rindler Quantization

In § 3.2 and § 3.3 we described new effects that must be taken into account in a complete calculation of entanglement entropy. One might wonder whether these effects are also relevant for the proper calculation of more traditional field theory quantities, such as correlation functions. The issue arises, in principle, when one attempts to calculate correlation functions using a quantization that splits the Hilbert space into multiple pieces, such as Rindler coordinates. In such a quantization the Minkowski vacuum is [284]

$$|\Omega\rangle \sim \sum_n e^{-\pi E_n} |\bar{\psi}_n\rangle_L \, |\psi_n\rangle_R. \tag{3.198}$$

where the sum runs over the energy eigensates $|\psi_n\rangle$ of the theory, and left/right subscripts refer to states living in the Hilbert space defined on the left/right halves of space at $t = 0$. However, in gauge theory the proper calculation of entanglement [218, 217] mandates instead the use of the state

$$|\Omega\rangle \sim \bigoplus_{E_\perp} \sqrt{p\,(E_\perp)} \sum_n e^{-\pi E_n} |\bar{\psi}_n,\ E_\perp\rangle_L \, |\psi_n,\ E_\perp\rangle_R. \tag{3.199}$$

where the state $|\psi_n,\ E_\perp\rangle$ lives in the Hilbert space with fixed $E_\perp$ where the two halves meet, and the distribution $\{p(E_\perp)\}$ of fluxes through the Rindler horizon makes the necessary extra contribution to the entanglement.

I showed that one obtains the correct gauge theory correlation functions using the naive state (3.198) without any consideration of the horizon fluxes. This follows simply from a computation of the two-point function in the scalar theory

in the state (3.198): one obtains the Minkowski vacuum result while neglecting the need to apply boundary conditions, even for trans-horizon correlators. Working in Feynman gauge, one can trivially show that this implies that all gauge-invariant correlators in the gauge theory also assume their Minkowski vacuum values.

3.4.1 Introduction

One peculiar feature of relativistic quantum theories is the freedom to quantize in any time coordinate. This leads to interesting physics even in flat space, such as the Unruh effect [282]: accelerating observers quantize in Rindler time τ rather than Minkowski time t and experience the Minkowski vacuum as a thermal state, a fact encoded in the Bogoliubov transformations between the accelerated and inertial mode functions.

Quantization in Rindler modes was first described by Fulling in his 1972 PhD. thesis [283] and is most often used as a tool to understand the thermal nature of the Minkowski vacuum [282, 284]. However it is also a perfectly valid way to describe operators in either Rindler wedge, and like any other quantization can be used to compute correlation functions wherever the coordinates are well-defined. At the Unruh temperature these computations ought to reproduce Minkowski expectation values, while at the physics at other temperatures can be quite different.

Correlation functions in Rindler quantization were the subject of intensive study in the 1970s and 1980s. However, the connection to Minkowski quantization is somewhat hidden in the literature, and has not to our knowledge been discussed in the context of the emerging links between entanglement and spacetime. It is the purpose of this note to highlight these old results, offer a new and more explicit

derivation using the thermofield double state and explore the connection to issues relevant to entanglement entropy and the black hole information paradox.

Boulware was the first to show equivalence between the quantization of a free scalar in the Minkowski vacuum and Rindler quantization in the thermal state at the Unruh temperature, which is often called the Hartle-Hawking state [285]. The argument is in appendix B of [286]: an integral representation of the Hartle-Hawking two-point function can be manipulated into an integral representation of the Minkowski vacuum two-point function, using integral representations for the concomitant modified Bessel functions and an analytic continuation. This is enough to conclude equivalence.

More general thermal states are also interesting, and away from the Unruh temperature they are singular on the horizon. The thermal two-point function was computed by Dowker in $d = 4$ [287, 288]. His main tool, Wick rotation, has played a prominent role in studies of Rindler space: at inverse temperature β the Wick rotation of Rindler space is a Euclidean cone of opening angle β, which is amenable to classical techniques. In [287] Dowker used Carslaw's result from 1919 [289] for the propagator on the infinite cover of the plane before imposing periodicity with the method of images to obtain a contour integral representation of the Euclidean two-point function, which can be evaluated [288] and continued [290] to obtain Lorentzian correlators.

The renormalized thermal stress-energy tensor can be computed either by conformal methods or by direct differentiation of Dowker's correlator; using the latter method, Brown and Ottewill [291] found that the boost energy density diverges at the horizon except when β is the inverse of the Unruh temperature.

Candelas and Deutsch [292] described a similar effect when a brick-wall boundary condition is imposed just outside the horizon, even at the Unruh temperature. Candelas [293] also studied some of these questions in the Schwarzschild geometry.

While the results of this note largely recapitulate the literature, our methods are different and it is instructive in any case to review these calculations since they bear on questions that had not arisen when they were first performed. In § 3.4.2 we give a new proof of the equivalence between the Minkowski vacuum and the Hartle-Hawking state using canonical quantization in the thermofield double state via explicit Lorentzian computation of the Wightman function. This evaluation makes manifest convergence issues which demand an $i\epsilon$ prescription in agreement with the corresponding Minkowski correlator. Our computation does not simply extend to arbitrary β but we study $\beta \to \infty$, the Boulware vacuum, and reproduce the results of Dowker in a tractable coincidence limit. We find a straightforward way to extend our discussion to U(1) gauge theory, which simplifies previous work [294, 295, 296]. In § 3.4.3 we discuss non-Hartle-Hawking states. After reviewing the behavior of the stress tensor at the horizon in § 3.4.4, following [291] and [292], we close in § 3.4.5 by addressing questions that arise from calculations of entanglement entropy: the necessity of flux and boundary condition sums at the horizon.

3.4.2 Minkowski correlations from Rindler quantization

The Hartle-Hawking state is the Minkowski vacuum [282, 284] and so Rindler correlation functions in the Hartle-Hawkingstate should reproduce their Minkowski vacuum expectation values. Our point of comparison will be the two-point (Wight-

man) function of a free massless scalar in the Minkowski vacuum of flat d-dimensional spacetime, which is [297]

$$\langle 0|\, \varphi(x^M, t)\varphi(x'^M, t')\, |0\rangle = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}} \frac{1}{\left[(x^M - x'^M)^2 - (t - t' - i\epsilon)^2\right]^{(d-2)/2}}. \quad (3.200)$$

The index $M = 1 \ldots d-1$ and $\epsilon > 0$. Rindler quantization in the Hartle-Hawking state must reproduce this result.

Rindler coordinates in the right (left) wedge $(\tau, z, \vec{x})$ are related to the Minkowski coordinates (t, x^M) by

$$t = (-)z \sinh \alpha\tau, \quad x^1 = (-)z \cosh \alpha\tau, \quad \vec{x} = (x^2, \ldots, x^{d-1}). \quad (3.201)$$

z ranges from 0 to ∞ while the other coordinates are unbounded; with this choice z is positive in each wedge. When written unadorned, x refers to $\vec{x}$. We will set $\alpha \to 1$ for the remainder, but dependence on the acceleration can be restored by dimensional analysis.

The solutions to the Rindler equation of motion

$$(\partial_z^2 + z^{-1}\partial_z + \vec{\partial}^2 - z^{-2}\partial_\tau^2)\varphi = 0 \quad (3.202)$$

involve the modified Bessel functions $I(z)$ and $K(z)$, the latter of which is finite as $z \to \infty$.[1] In the right and left wedges respectively, the mode functions

$$f_{\omega k}(z, \tau, x) = K_{i\omega}(|k|z)e^{i(kx - \omega\tau)}, \qquad \tilde{f}_{\omega k}(z, \tau, x) = K_{i\omega}(|k|z)e^{i(kx + \omega\tau)} \quad (3.203)$$

[1] In the past and future wedges the argument becomes complex, and the relevant solutions involve the second Hankel function $H^{(2)}$ in place of K.

parametrize the solutions to (3.202) with the appropriate asymptotics. Solutions with $\omega > 0$ are positive-frequency modes with respect to the generator of Rindler time translations in each wedge and accompany the creation/annihilation operators of a field expanded in Rindler coordinates just as planar mode functions accompany the Cartesian field expansion in flat space:

$$\varphi^R(z,\tau,x) = \int_0^\infty d\omega \int d^{d-2}k \; N_{\omega k} \left[f_{\omega k}(z,\tau,x) a^R_{\omega k} + \text{h.c.} \right] \tag{3.204}$$

in the right wedge, and

$$\varphi^L(z,\tau,x) = \int_0^\infty d\omega \int d^{d-2}k \; N_{\omega k} \left[\tilde{f}_{\omega k}(z,\tau,x) a^L_{\omega k} + \text{h.c.} \right] \tag{3.205}$$

in the left. The range of the ω integral reflects the division into positive and negative frequency modes. With the normalization

$$N^2_{\omega k} = (2\pi)^{d-2} \frac{\sinh \pi\omega}{\pi^2} \tag{3.206}$$

the modefunctions are Klein-Gordon orthonormal (which can be shown using the orthogonality relation $\int_0^\infty \frac{dx}{x} K_{i\omega}(x) K_{i\omega'}(x) = \frac{\pi^2}{2\omega \sinh \pi\omega} \delta_{\omega\omega'}$ – proved quite recently, see [298]) and

$$[a^i_{\omega k}, a^{j\dagger}_{\omega' k'}] = \delta^{ij} \delta_{\omega\omega'} \delta_{kk'} \tag{3.207}$$

follows from the canonical commutation relation for φ. Here i, j are either L or R.

We will compute correlation functions in the thermofield double state

$$|\mathrm{TFD},\beta\rangle = \bigotimes_{\omega,k} Z_{\omega,k} \sum_n e^{-\frac{\beta n\omega}{2}} |\Theta(n,\omega,k)\rangle_L \; |n,\omega,k\rangle_R. \tag{3.208}$$

Θ is a CPT conjugation picked out by the path integral preparation, with its P operator acting only on the (τ, z) plane:

$$|\Theta(n,\omega,k)\rangle = |n,\omega,-k\rangle, \tag{3.209}$$

and the $Z_{\omega,k}$ are chosen such that $\langle \mathrm{TFD},\beta|\,\mathrm{TFD},\beta\rangle = 1$. The Hartle-Hawking state $|\Omega\rangle$ is the thermofield state (3.208) at $\beta = 2\pi$ [282, 284],

$$|\Omega\rangle = |\mathrm{TFD}, 2\pi\rangle\,. \tag{3.210}$$

The details of the computation depend on whether or not the operators are on the same side of the horizon. First we study one-sided correlators, putting both operators in the right wedge. It will be useful to separate the product of fields into the commutator and the anticommutator:

$$\varphi(X)\varphi(X') = \frac{1}{2}[\varphi(X),\varphi(X')] + \frac{1}{2}\{\varphi(X),\varphi(X')\}. \tag{3.211}$$

The commutator is independent of the state and so the interesting piece is the anticommutator, a hermitean operator and the subject of our first computation. We have denoted the coordinates collectively by X and will also abbreviate $(\omega, \vec{k}) \equiv S$, $\int_0^\infty d\omega \int d^{d-2}k \equiv \int dS$. Taking the expectation value in the Hartle-Hawking state

(3.210) and using the field expansion (3.204), one finds

$$\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle = \langle\Omega|\left[\int dSdS' N_S N'_S \left(a_S f_S(X) + a_S^\dagger f_S^*(X)\right)\right.$$
$$\left.\times\left(a_{S'} f_{S'}(X') + a_{S'}^\dagger f_{S'}^*(X')\right)\right]|\Omega\rangle + (X \leftrightarrow X') \quad (3.212)$$

The aa and $a^\dagger a^\dagger$ terms do not contribute because the left and right occupation numbers must match, while the $a^\dagger a$ and $aa^\dagger$ terms can be computed using the properties of the Bose-Einstein distribution, since $|\Omega\rangle$ is a thermal state in ω with $\beta = 2\pi$:

$$\langle\Omega|\, a_{k\omega}^\dagger a_{k'\omega'}\,|\Omega\rangle = \frac{1}{e^{2\pi\omega}-1}\delta_{SS'}, \quad \langle\Omega|\, a_{k\omega} a_{k'\omega'}^\dagger\,|\Omega\rangle = \frac{e^{2\pi\omega}}{e^{2\pi\omega}-1}\delta_{SS'}. \quad (3.213)$$

Their sum is $\coth\pi\omega$. After a short computation one finds

$$\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle = \int dS \coth\pi\omega\ N_{\omega k}^2\left[f_S(X)f_S^*(X') + f_S^*(X)f_S(X')\right]$$
$$= \int_0^\infty d\omega \int \frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{\cosh\pi\omega}{\pi^2}\left[e^{i[k\Delta x - \omega\Delta\tau]}K_{i\omega}(|k|z)K_{i\omega}^*(|k|z')\right.$$
$$\left.+(X \leftrightarrow X')\right]. \quad (3.214)$$

The evaluation of this integral is left to the appendix. The basic idea is to first do the integral without the cosh, using integral representations for the Bessel functions and the fact that the integrand is even in ω. The cosh can then be accomodated via analytic continuation, which is only well-defined if we implement

the $i\epsilon$ prescription below. The result is

$$\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\left[\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'-i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right.$$
$$\left.+\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'+i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right]. \tag{3.215}$$

The computation of the commutator yields a similar expression:

$$[\varphi^R(X),\varphi^R(X')] = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\left[\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'-i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right.$$
$$\left.-\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'+i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right]. \tag{3.216}$$

This vanishes except on the light cone. Together they imply the one-sided Wightman function

$$\langle\Omega|\,\varphi^R(X)\varphi^R(X')\,|\Omega\rangle = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'-i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}} \tag{3.217}$$

which is just the Minkowski expression (3.200) after the change of coordinates (3.201). The ϵs in (3.217) and (3.200) are related by a z-dependent redefinition, but since $z>0$ this does not affect the sign.

The computation of the two-sided correlator is similar but the matrix elements

take more work, again left to the appendix. The correlator

$$\langle\Omega|\,\varphi^R(X)\varphi^L(X')\,|\Omega\rangle = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\frac{1}{[z^2+z'^2+2zz'\cosh(\tau-\tau')+(\vec{x}-\vec{x}')^2]^{(d-2)/2}} \tag{3.218}$$

is once more the Minkowski result (3.200) with the change of coordinates (3.201), after accounting for the extra sign in the transformation to the left wedge (recall we defined z to be positive in both wedges). As explained in the appendix there is no need for an $i\epsilon$ prescription, which is consistent because points in different wedges are spacelike separated. The two-sided commutator vanishes, as required by causality:

$$[\varphi^R(X),\varphi^L(X')]=0. \tag{3.219}$$

These results extend to gauge theory. As reviewed in the appendix, the only difference between the scalar and gauge calculations is the presence of a polarization sum. The thermal trace includes a sum over longitudinal modes, but the same modes appear in the Minkowski calculation and drop out of gauge-invariant correlators. In Feynman gauge the propagator is just the scalar propagator times the metric and so

$$\langle\Omega|\,A_\mu(X)A_\nu(X')\,|\Omega\rangle = g_{\mu\nu}\,\langle\Omega|\,\varphi(X)\varphi(X')\,|\Omega\rangle\,. \tag{3.220}$$

It follows that correlation functions of gauge invariant (and in Feynman gauge, gauge variant) operators match their Minkowski vacuum expectation values.

This correlation structure implies the familiar fact that the Rindler horizon is invisible in the Hartle-Hawking state, since correlators match their Minkowski

vacuum expectation values. All trans-horizon probes necessarily are well-behaved: for example, the two-point function changes smoothly as an operator is dragged across the horizon, and the expectation values of all operators everywhere assume their Minkowski vacuum values.

3.4.3 Non-Hartle-Hawking states

In other thermal states $|\text{TFD}, \beta\rangle$ the Rindler horizon is far from invisible. The anticommutator in the thermofield state (3.208) is

$$\begin{aligned}\langle \text{TFD}, \beta | \{\varphi^R(X), \varphi^R(X')\} | \text{TFD}, \beta \rangle = \int dS \coth(\beta\omega/2) \frac{\sinh(\pi\omega)}{\pi^2} e^{i(k\Delta x - \omega\Delta\tau)} \\ \times K_{i\omega}(|k|z) K^*_{i\omega}(|k|z') \\ + (X \leftrightarrow X').\end{aligned} \tag{3.221}$$

At $\beta \neq 2\pi$ we will not obtain Minkowski vacuum expectation values. The simplest state with $\beta \neq 2\pi$ is the Boulware vacuum [286]

$$|B\rangle \equiv |\text{TFD}, \infty\rangle = |0\rangle_L \otimes |0\rangle_R \tag{3.222}$$

obtained by taking $\beta \to \infty$ in (3.208). It is annihilated by the $a^i_{\omega k}$. As in any product state there are no trans-horizon correlations, and their absence necessarily implies a firewall.

The argument is straightforward. The renormalized stress tensor $\overline{T}_{\mu\nu}$ is defined by subtracting off the Minkowski vacuum expectation value $\langle 0 | T_{\mu\nu} | 0 \rangle$, which is divergent everywhere due to the short-distance correlation structure. When all

trans-horizon correlators are zero,

$$\langle B|\,\overline{T}_{\mu\nu}(z=0,\tau,x)\,|B\rangle = -\,\langle 0|\,T_{\mu\nu}(z=0,\tau,x)\,|0\rangle\,, \tag{3.223}$$

since $\langle B|\,T_{\mu\nu}(z=0,\tau,x)\,|B\rangle = 0$ follows immediately from the lack of trans-horizon correlations and the point-splitting [297] definition of $T_{\mu\nu}$. The renormalized stress tensor at the horizon is therefore proportionate to the singular unrenormalized Minkowski vacuum expectation value.

Actually, any non-Hartle-Hawking thermal state is singular on the horizon [291] despite the thermal entanglement (albeit at the wrong temperature). Dowker computed the two-point function at arbitrary inverse temperature in $d = 4$ using Euclidean techniques [288]:

$$\langle\varphi(z,\tau,x)\varphi(z',0,0)\rangle_\beta = \frac{i}{4\pi\beta zz'\sinh\gamma}\frac{\sinh\left(2\pi\gamma/\beta\right)}{\cosh\left(2\pi\gamma/\beta\right)-\cosh\left(2\pi\tau/\beta\right)} \tag{3.224}$$

where

$$\cosh\gamma = \frac{z^2+z'^2+|x|^2}{2zz'}. \tag{3.225}$$

As we review in the next section, these states have divergent stress-energy as $z \to 0$ whenever $\beta \neq 2\pi$. On the other hand, taking $\beta = 2\pi$ and doing the appropriate Lorentzian continuation, it is easy to see that (3.224) agrees with the Hartle-Hawking result (3.217). This is our point of comparison for the general thermal calculation in $d = 4$.

The $\beta \to \infty$ limit of (3.221) is harder to compute since the integrand is odd in ω, so the tricks applied to the Hartle-Hawking state in the previous section will

not work here.[2] However it is relatively easy to evaluate the integrals in the limit of coincident $\vec{x}$ in $d = 4$, where we obtain

$$\langle B|\, \varphi^R(z,\tau,x)\varphi^R(z',\tau',x)\, |B\rangle = \frac{1}{4\pi^2(z^2-z'^2)}\left[\frac{1}{\Delta\tau + \log\frac{z}{z'} - i\epsilon} - \frac{1}{\Delta\tau - \log\frac{z}{z'} - i\epsilon}\right]. \tag{3.226}$$

This agrees with Dowker's result (3.224) when $x = 0$ (i.e. $\gamma = \log(z/z')$). As either of the operators is taken to the horizon (z or $z' \to 0$), (3.226) vanishes. By contrast the Wightman function in the Hartle-Hawking state (3.217) approaches $\frac{1}{z^2}$ when $z' \to 0$ at coincident x, which is just a reflection of the fact that all points on a boost orbit have the same distance from the origin.

3.4.4 Stress-energy at the horizon

Since the Wightman function in a general state in the thermofield double differs from the Wightman function in the Minkowski vacuum, it corresponds to a non-Minkowski vacuum distribution of stress-energy. The stress tensor can be computed either directly from (3.224) [299], or by computing the response of the effective action to a change in the metric. Brown and Ottewill [291] took the latter approach in $d = 4$ and computed the stress tensor using the dimensionally regularized effective action; in this scheme one obtains a vanishing Minkowski vacuum stress tensor and so the method agrees with the canonical computation in the vacuum subtraction scheme. When the manifold is conformally related to one on which the trace of the stress tensor vanishes, such as Rindler to Minkowski, the stress tensor on the original manifold is determined in terms of the (a, c)

[2]The integral representation $K_{i\omega}(|k|z)K^*_{i\omega}(|k|z') = \frac{1}{2}\int_{-\infty}^{\infty} d\lambda e^{i\omega\lambda} K_0(|k|\upsilon)$, where $\upsilon = z^2 + z'^2 + 2zz'\cosh\lambda$, might be useful; finite β and general d are likely tractable.

anomalies of the theory and geometric data of the conformal relation. They find for the thermal expectation value of the stress tensor in the Rindler theory at inverse temperature β

$$\langle \mathrm{TFD}, \beta | \, \overline{T}_{\mu\nu}(z, \tau, x) \, | \mathrm{TFD}, \beta \rangle = \frac{1}{1440\pi^2 z^4} \left[\left(\frac{2\pi}{\beta} \right)^4 - 1 \right] (g_{\mu\nu} + 4 v_\mu v_\nu) \quad (3.227)$$

where $v_\mu = (z, \vec{0})$ is a unit vector proportionate to the boost Killing vector ∂_τ. This is a renormalized stress tensor (it vanishes in the Minkowski vacuum) and a zero-temperature term has been omitted [291]. From (3.227) it is clear that the stress-energy diverges at the Rindler horizon at any non-Unruh temperature: there is a firewall. The states with $\beta \neq 2\pi$ are still thermally entangled and have two-sided correlations at finite β but correspond to Euclidean path integrals with a conical deficit inserted at the origin on the $t = 0$ slice.

It is interesting to contrast this behavior with the stress tensor obtained from a brick-wall quantization of the scalar field, where a boundary condition such as

$$\varphi(z = z_0, \tau, \vec{x}) = 0 \quad (3.228)$$

is imposed on the "stretched horizon" at $z = z_0$ and the field expansion (3.204) modified to satisfy the boundary conditions. This quantization was studied in detail (in $d = 4$) by Candelas and Deutsch [292]. Letting G_0 denote the Hartle-Hawking Wightman function (3.217), they found

$$D(X, X') = G_0(X, X') - \frac{i}{\pi} \int_0^\infty \frac{d\omega}{2\pi} e^{-i\omega\Delta\tau} \int \frac{d^2k}{(2\pi)^2} e^{ik\Delta x} \frac{K_{i\omega}(e^{i\pi}|k|z_0)}{K_{i\omega}(|k|z_0)} K_{i\omega}(|k|z) K_{i\omega}(|k|z') \quad (3.229)$$

for the Wightman function in the Hartle-Hawking state with Dirichlet boundary conditions, and

$$N(X,X') = G_0(X,X') + \frac{i}{\pi}\int_0^\infty \frac{d\omega}{2\pi} e^{-i\omega\Delta\tau}\int \frac{d^2k}{(2\pi)^2} e^{ik\Delta x}\frac{K'_{i\omega}(e^{i\pi}|k|z_0)}{K'_{i\omega}(|k|z_0)} K_{i\omega}(|k|z) K_{i\omega}(|k|z') \tag{3.230}$$

with Neumann. The expectation value of the renormalized stress tensor is

$$\langle \text{TFD (brick wall)}, 2\pi | \, \overline{T}_{\mu\nu}(z,\tau,\vec{x}) \, | \text{TFD (brick wall)}, 2\pi \rangle = \text{diag}(c_1, c_2, c_3, c_3) \tag{3.231}$$

where $c_{1,2,3}$ depend on the choice of boundary conditions.[3] Candelas and Deutsch give integral representations for the c_i and evaluate the stress tensor near the brick wall at $z = z_0$:

$$c_1 \sim -\frac{z^2}{360\pi^2 z_0 (z-z_0)^3}, \qquad c_2 \sim \frac{(z-z_0)c_1}{2z_0 z^2}, \qquad c_3 \sim \frac{c_1}{2z^2} \tag{3.232}$$

for both Dirichlet and Neumann conditions. The independence of the near-horizon stress tensor on the choice of boundary conditions is unexpected. They find a similar result for gauge fields.

The Hartle-Hawking state with a brick wall is qualitatively similar to a thermal state at $\beta \neq 2\pi$, as both have divergent stress-energy at the horizon. However, the degree of divergence is different. The divergence structure of the brick-wall stress-energy follows from dimensional analysis and the fact that the stress tensor

[3] This expectation value is computed in the analog of the Hartle-Hawking state in the brick-wall quantization. The boundary condition imposes a quantization condition on the frequencies, and so the tensor product in (3.208) runs over a discrete set that depends on the boundary condition.

must reproduce the result for an unaccelerated barrier as $z_0 \to \infty$.

3.4.5 Discussion: Correlations vs. entanglement

It is useful to contrast these results with calculations of entanglement entropy, which can be quite subtle [300, 220], especially in gauge theory [300, 216, 218]. Gauge invariance implies [218] that the state that leads to the correct calculation of the entropy in ordinary Maxwell theory is the extended thermofield state

$$|\text{eTFD}, \beta\rangle = \bigoplus_{E_\perp} \sqrt{p\left(E_\perp\right)} \bigotimes_{\lambda,\omega,k} Z_{\omega k} \sum_n e^{-\frac{\beta n \omega}{2}} |n, \lambda, \omega, -k; E_\perp\rangle_L \; |n, \lambda, \omega, k; E_\perp\rangle_R. \tag{3.233}$$

Here $E_\perp$ is a configuration of the normal components of the electric field on the horizon, $p\left(E_\perp\right)$ are a set of probabilities that were computed for $d = 4$ Maxwell theory in [257, 217], and the polarizations λ and an omitted ghost dressing are discussed around eq. (3.261) in the appendix. The general structure of the state (3.233) was described by Donnelly [218] in $d = 2$, where $E_\perp$ is the only quantum number. The n^{th} term in the sum (3.233) is a product of states with n photons on top of the coherent state $|0; E_\perp\rangle$ in which the normal electric field at the horizon is $E_\perp$.

The block-diagonal structure of (3.233) is required by Gauss's law, which equates gauge-invariant operators at the boundary of the subregion with a sum of operators localized outside the subregion. Consequently, by causality the boundary operators must commute with any operators localized to the subregion, and the algebra of the subregion decomposes into superselection sectors: in the Rindler wedge they are labeled by the normal electric field $E_\perp$ at the horizon. The block-

diagional structure in (3.233) makes an additional Shannon contribution to the vacuum entanglement entropy of the Rindler wedge [218] which must be included in order to obtain a result consistent with conformal symmetry in $d = 4$ Maxwell theory [217]. In conformal theories the individual blocks are BCFT states, and their contribution to the entropy is a weighted sum of Affleck-Ludwig entropies [301, 220].

Computations of correlation functions, or the stress tensor – actual observables, unlike vacuum entanglement – could in principle share this structure: it might have been necessary to do a flux sum in order to obtain the correct value. However, we computed eq. (3.220), which demonstrated the equivalence between the Rindler and Minkowski quantizations of Maxwell theory, using the naive thermofield state described in the appendix

$$|\text{TFD}, \beta\rangle = \bigotimes_{\lambda,\omega,k} Z_{\omega k} \sum_n e^{-\frac{\beta n\omega}{2}} |n, \lambda, \omega, -k\rangle_L \; |n, \lambda, \omega, k\rangle_R \tag{3.234}$$

which has no flux sum. It is possible that the right computation in the block-diagonal state (3.233) leads to (3.220) as well, but the correct correlator can be obtained more simply using (3.234) sans flux sum. In the scalar theory there is no gauge invariance to complicate the factorization of the Hilbert space, but the issue of boundary conditions at the horizon remains [220]. One might have considered a sum over horizon field values motivated by the vacuum path integral, but (3.217) shows that correlators can be correctly computed without such a horizon sum.

Acknowledgements

I am grateful to Will Donnelly, Don Marolf, Eric Mintun, Joe Polchinski and Mark Srednicki for extensive discussions. I especially thank Joe for suggesting the project, and Will and Mark for comments on the draft. I was supported by the NSF Graduate Research Fellowship Grant DGE-1144085 and NSF Grant PHY13-16748 over the course of this work.

3.4.6 Appendix

This appendix describes the canonical quantization computation of the massless scalar two-point function in the Hartle-Hawking state. We also discuss the Boulware vacuum and the quantization of Maxwell theory.

Our ultimate goal is to evaluate the Wightman function but we will have to begin with the anticommutator:

$$\begin{aligned}\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle &= \int_0^\infty \frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{\cosh\pi\omega}{\pi^2}\left[e^{i(k\Delta x-\omega\Delta\tau)}\right.\\ &\quad\left.+(X\leftrightarrow X')\right].\end{aligned} \tag{3.235}$$

The first step is to evaluate the integral without the $\cosh\pi\omega$:

$$\begin{aligned}&\frac{1}{\pi^2}\int_0^\infty\int\frac{d^{d-2}k}{(2\pi)^{d-2}}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')+(X\leftrightarrow X')\right]\\ &\qquad=\frac{1}{2\pi^2}\int_{-\infty}^\infty d\omega\int\frac{d^{d-2}k}{(2\pi)^{d-2}}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')+(X\leftrightarrow X')\right]\\ &\qquad\equiv C(X,X')+(X\leftrightarrow X').\end{aligned} \tag{3.236}$$

We proceed by using an integral representation of the modified Bessel function

$$K_\alpha(t) = \frac{1}{2}(t/2)^\alpha \int_0^\infty \frac{du}{u^{\alpha+1}} e^{-u-\frac{t^2}{4u}} \tag{3.237}$$

and then doing the ω integral:

$$\begin{aligned}
C(X,X') &= \frac{1}{8\pi^2}\int_{-\infty}^{\infty} d\omega \int \frac{d^{d-2}k}{(2\pi)^{d-2}} e^{i(k\Delta x - \omega\Delta\tau)} \int_0^\infty \frac{dudu'}{uu'} \left(\frac{zu'}{z'u}\right)^{i\omega} e^{-(u+u')} e^{-\frac{k^2}{4}\left(\frac{z^2}{u}+\frac{z'^2}{u'}\right)} \\
&= \frac{1}{8\pi^2}\int \frac{d^{d-2}k}{(2\pi)^{d-3}} e^{ik\Delta x} \int_0^\infty \frac{dudu'}{uu'} e^{-(u+u')} e^{-\frac{k^2}{4}\left(\frac{z^2}{u}+\frac{z'^2}{u'}\right)} \delta\left(\log\frac{zu'}{z'u} - \Delta\tau\right) \\
&= \frac{1}{8\pi^2}\int_0^\infty \frac{du}{u} e^{-u\left(1+\frac{z'}{z}e^{\Delta\tau}\right)} \int \frac{d^{d-2}k}{(2\pi)^{d-3}}\, e^{ik\Delta x} e^{-\frac{k^2 z^2}{4u}\left(1+\frac{z'}{z}e^{-\Delta\tau}\right)} \\
&= \frac{(2\pi)^{2-\frac{d}{2}}}{8\pi^2}\int_0^\infty \frac{du}{u} e^{-u\left(1+\frac{z'}{z}e^{\Delta\tau}\right)} \left(\frac{z^2+zz'e^{-\Delta\tau}}{2u}\right)^{-\frac{(d-2)}{2}} e^{-u\left(\frac{x^2}{z^2+zz'e^{-\Delta\tau}}\right)} \\
&= \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}} \left(z^2+z'^2+2zz'\cosh\Delta\tau+\Delta x^2\right)^{-\frac{d-2}{2}}.
\end{aligned} \tag{3.238}$$

Note that the original integrand had to be even in ω in order for this method to work.

Now we return to the anticommutator, which differs from (3.238) by a $\cosh\pi\omega$ factor. This factor can be almost be accomodated by analytically continuing the integral to $\Delta\tau \to \Delta\tau \pm i\pi$, but the integral diverges outside a strip in the complex $\Delta\tau$ plane and will require regulation. The problematic regions of the integral are at $\omega \to \pm\infty$. At large imaginary order, the modified Bessel functions have the expansion [302]

$$K_{i\omega}(x) = -\left(\frac{\pi}{\omega\sinh\pi\omega}\right)^{1/2} \sin\left(\omega\log(x/2) - \gamma_\omega\right) + O(x^2), \tag{3.239}$$

where γ_ω is the phase of $\Gamma(1+i\omega)$. Therefore the integral (3.238) is only finite at $\omega \to \pm\infty$ if

$$-\pi < \text{Im } \Delta\tau < \pi. \tag{3.240}$$

Using symmetry in ω, the anticommutator (3.235) is

$$\begin{aligned}\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle &= \int_{-\infty}^{\infty}\frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{e^{\pi\omega}+e^{-\pi\omega}}{4\pi^2}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')\right. \\ &\left.+(X\leftrightarrow X')\right].\end{aligned} \tag{3.241}$$

The only effect of the cosh is to multiply the integrand in (3.238) by an $e^{\pm\pi\omega}$ factor but this renders the integral divergent. The integral can be regulated by inserting a factor $e^{\mp\epsilon\omega}$ into the integrand, the sign depending on whether the offending divergence is at $\pm\infty$; $\epsilon > 0$ will be taken to zero at the end of the calculation. The integral with the $e^{\pm\pi\omega}$ factor and the necessary regulator $e^{\mp\epsilon\omega}$ can then be computed by continuing $\Delta\tau \to \Delta\tau \pm i\pi \mp i\epsilon$ in (3.238) ($\Delta\tau \to \Delta\tau \mp i\pi \pm i\epsilon$ in the $X \leftrightarrow X'$ term in (3.236)). This procedure is well-defined since the continuation stays within the convergence bounds (3.240). One evaluates

$$\begin{aligned}&\int_{-\infty}^{\infty}\frac{d^{d-2}k}{(2\pi)^{d-2}}e^{\pi\omega}e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z') \\ &\qquad\to\int_{-\infty}^{\infty}\frac{d^{d-2}k}{(2\pi)^{d-2}}e^{\pi\omega}e^{-\omega\epsilon}e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z') \\ &\qquad=\int_{-\infty}^{\infty}\frac{d^{d-2}k}{(2\pi)^{d-2}}e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')\Big|_{\Delta\tau\to\Delta\tau+i\pi-i\epsilon}\end{aligned} \tag{3.242}$$

and similarly for the other terms. Combined with the result (3.238) for the integral

to be continued,

$$\langle\Omega|\left\{\varphi^R(X),\varphi^R(X')\right\}|\Omega\rangle = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\left[\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'-i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right.$$
$$\left.+\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'+i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right]. \tag{3.243}$$

Now we compute the commutator. Using the mode expansion (3.204),

$$[\varphi^R(X),\varphi^R(X')] = \int_0^\infty \frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{\sinh\pi\omega}{\pi^2}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')-(X\leftrightarrow X')\right]. \tag{3.244}$$

This differs from (3.235) by the sinh instead of the cosh, and the relative sign on the $(X\leftrightarrow X')$ term. It follows that the integrand is still even in ω and so we can use the result (3.238), obtaining

$$[\varphi^R(X),\varphi^R(X')] = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\left[\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'-i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right.$$
$$\left.-\frac{1}{\left[z^2+z'^2-2zz'\cosh(\tau-\tau'+i\epsilon)+(\vec{x}-\vec{x}')^2\right]^{(d-2)/2}}\right]. \tag{3.245}$$

after doing the requisite regulations and analytic continuations as described above. From (3.235) and (3.244) we obtain the Wightman function (3.217) in the Hartle-Hawking state. We could not have computed the Wightman function directly using the methods of this appendix, since the integrand is not even in ω.

Now we consider operators inserted on opposite sides of the Rindler horizon.

The anticommutator in the Hartle-Hawking state is

$$\begin{aligned}\langle\Omega|\left\{\varphi^R(X),\varphi^L(X')\right\}|\Omega\rangle = \quad & \langle\Omega|\left[\int dSdS' N_S N'_S \left(a^R_S f_S(X) + a^{R\dagger}_S f^*_S(X)\right)\right. \\ & \left.\times\left(a^L_{S'}\tilde{f}_{S'}(X') + a^{L\dagger}_{S'}\tilde{f}^*_{S'}(X')\right)\right]|\Omega\rangle + (X\leftrightarrow X')\end{aligned} \tag{3.246}$$

The $aa^\dagger$ and $a^\dagger a$ terms do not contribute because the left and right boson numbers must match, while the aa and $a^\dagger a^\dagger$ terms can be computed from the thermofield double:

$$\begin{aligned}\langle\Omega|\, a^R_{\omega k}a^L_{\omega' k'}\,|\Omega\rangle &= \frac{1}{1-e^{-\beta\omega/2}}\sum_{n,n'}\langle n'\omega k|_R\,\langle n'\omega(-k)|_L\, e^{-\frac{\beta\omega}{2}(n+n')}a^R_{k\omega}a^L_{k'\omega'} \\ &\qquad\times\ |n\omega(-k)\rangle_L\,|n\omega k\rangle_R \\ &= \frac{\delta_{\omega\omega'}\delta_{k(-k')}}{1-e^{-\beta\omega/2}}\sum_{n,n'} n\,\langle n'\omega k|_R\,\langle n'\omega(-k)|_L\, e^{-\frac{\beta\omega}{2}(n+n')} \\ &\qquad\times\ |n-1,\omega(-k)\rangle_L\,|n-1,\omega k\rangle_R \\ &= \frac{1}{2}\,\mathrm{csch}\left(\frac{\beta\omega}{2}\right)\delta_{\omega\omega'}\delta_{k(-k')} \\ &= \left(\langle\Omega|\, a^{R\,\dagger}_{\omega k}a^{L\,\dagger}_{\omega k}\,|\Omega\rangle\right)^\dagger.\end{aligned} \tag{3.247}$$

The first equality is slightly nontrivial: inserting $a^R_{\omega k}a^L_{\omega' k'}$ into the thermofield state (3.208), there are two tensor products over the momentum quantum numbers in the bra and the ket (call them $S_{\rm bra}$ and $S_{\rm ket}$) which have been omitted here. However, unless $\delta_{\omega\omega'}\delta_{k(-k')} = 1$ there will be a mode number mismatch between the left and the right; the orthonormality of the $|n,\omega,k\rangle$ ensures that $S_{\rm bra} = S_{\rm ket} \equiv (\bar\omega,\bar k)$, and except for $\bar\omega = \omega$, the normalization factor $Z_{\bar\omega} = (1-e^{-\beta\bar\omega/2})^{-1}$ cancels against the thermal sum since there are no insertions with those quantum

numbers. This implies the first equality of (3.247).

Plugging (3.247) into (3.246) one obtains for the anticommutator

$$\langle\Omega|\left\{\varphi^R(X),\varphi^L(X')\right\}|\Omega\rangle=\int_0^\infty\frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{1}{\pi^2}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')+(X\leftrightarrow X')\right] \tag{3.248}$$

which is the integral without the cosh evaluated above. Thus there is no need for analytic continuation and the answer is just

$$\langle\Omega|\left\{\varphi^R(X),\varphi^L(X')\right\}|\Omega\rangle=\frac{\Gamma\left(\frac{d-2}{2}\right)}{4\pi^{d/2}}\left(z^2+z'^2+2zz'\cosh\Delta\tau+\Delta x^2\right)^{-\frac{d-2}{2}}. \tag{3.249}$$

Since we did not analytically continue there is no need to regulate the integrals with an $i\epsilon$ prescription, which is perfectly consistent with the Minkowski description since insertions in different wedges are necessarily at spacelike separation. This also implies that the commutator vanishes:

$$[\varphi^R(X),\varphi^L(X')]=0 \tag{3.250}$$

as required by causality.

Next we will attempt to compute the Wightman function in the Boulware vacuum $|B\rangle$. Using the expansion (3.204) and the fact that $a|B\rangle=0$, one finds for the anticommutator

$$\begin{aligned}\langle B|\left\{\varphi^R(X),\varphi^R(X')\right\}|B\rangle &= \int_0^\infty\frac{d^{d-2}k}{(2\pi)^{d-2}}\frac{\sinh\pi\omega}{\pi^2}\left[e^{i(k\Delta x-\omega\Delta\tau)}K_{i\omega}(|k|z)K^*_{i\omega}(|k|z')\right.\\ &\left.+(X\leftrightarrow X')\right].\end{aligned} \tag{3.251}$$

This integral cannot be evaluated using the methods of this appendix, since the integrand is odd in ω. However, the computation simplifies at $x = x'$ in $d = 4$:

$$\begin{aligned}\langle B|\, \varphi^R(X)\varphi^R(X')\, |B\rangle &= \int_0^\infty \frac{d^2k}{(2\pi)^2}\frac{\sinh \pi\omega}{\pi^2} e^{-i\omega\Delta\tau} K_{i\omega}(|k|z)K^*_{i\omega}(|k|z') \\ &= \frac{1}{4\pi^3}\int_0^\infty d\omega \sinh\pi\omega\ e^{-i\omega\Delta\tau}\left(\frac{z}{z'}\right)^{i\omega} \\ &\quad \times \int_0^\infty du\ du' ds \left(\frac{u'}{u}\right)^{i\omega} e^{-(u+u')} e^{-s(z^2u'+z'^2u)} \\ &= \frac{1}{4\pi^2(z^2-z'^2)}\left[\frac{1}{\Delta\tau + \log\frac{z}{z'} - i\epsilon} - \frac{1}{\Delta\tau - \log\frac{z}{z'} - i\epsilon}\right]\end{aligned} \tag{3.252}$$

after some manipulation. Again the $i\epsilon$ prescription is obtained by demanding convergence. This agrees with Dowker's result (3.224) when $\vec{x} = 0$, and approaches $(z\Delta\tau)^{-2}$ as $z \to z'$; in the same limit the thermofield correlator (3.217) approaches $(z(1-\cosh\Delta\tau))^{-2}$.

Last, we study the Rindler quantization of abelian Maxwell theory. Our goal will be to show that correlation functions of gauge-invariant operators in the Hartle-Hawking state, such as

$$\langle\Omega|\, F_{\mu\nu}(X)F_{\rho\sigma}(X')\, |\Omega\rangle\,, \tag{3.253}$$

agree with their Minkowski expectation values. We will do this by showing that the Wightman function

$$\langle\Omega|\, A_\mu(X)A_\nu(X')\, |\Omega\rangle \tag{3.254}$$

agrees with the Minkowski vacuum Wightman function

$$\langle 0|\, A_\mu(X) A_\nu(X')\, |0\rangle \tag{3.255}$$

in Feynman gauge.

Ignoring the ghosts, the Lagrangian in Feynman gauge is

$$\mathcal{L} = -\frac{1}{2}\left(\partial_\mu A_\nu\right)^2 \tag{3.256}$$

and so the quantization of the gauge field is just that of a d scalars. In the right wedge the d scalar fields can be expanded as

$$A_\mu^R = \sum_{\lambda=0}^{d-1} \int_0^\infty d\omega \int d^{d-2}k\ N_{\omega k} \left[f_{\omega k}(z,\tau,x) a^R_{\lambda,\omega k} \epsilon^\lambda_{\mu,\omega k} + \text{h.c.} \right] \tag{3.257}$$

where the $\epsilon^\lambda_{\mu,\omega k}$ are polarization vectors and λ runs over d polarizations, including two unphysical. Choosing $N_{\omega k}$ as for the scalar gives

$$[a_{\lambda,\omega k}, a^\dagger_{\lambda',\omega' k'}] = g_{\lambda\lambda'} \delta_{\omega\omega'} \delta_{kk'} \tag{3.258}$$

if we require the ϵ to obey

$$\epsilon^\lambda_{\mu,\omega k} (\epsilon^{\lambda'}_{\mu,\omega k})^* = g^{\lambda\lambda'} \tag{3.259}$$

and

$$\sum_\lambda \epsilon^\lambda_{\mu,\omega k} (\epsilon^\lambda_{\nu,\omega k})^* = g_{\mu\nu}. \tag{3.260}$$

We take the gauge theory analog of the thermofield state (3.208) to be

$$|\mathrm{TFD}\rangle = \bigotimes_{\lambda,\omega,k} Z_{\omega k} \sum_n e^{-\frac{\beta\omega n}{2}} |n,\omega,-k,\lambda\rangle_L \; |n,\omega,k,\lambda\rangle_R. \tag{3.261}$$

Note the absence of a flux sum. The state (3.261) lives in a Hilbert space that is not obviously physical, as it includes states of negative norm. However, such states are BRST-exact and thus equivalent to zero in the cohomology of physical states; they do not contribute to gauge-invariant correlation functions. In order to describe an actual state in the gauge-fixed Hilbert space (3.261) must technically be dressed with a tensor factor describing the ghost fields, but in the abelian theory the ghosts decouple and their tensor factor does not even affect the correlation functions of gauge-variant operators.

As usual our object of interest will be the Wightman function in the state (3.261) at $\beta = 2\pi$. After a short computation entirely analogous to the scalar case, one finds the anticommutator

$$\langle\Omega|\,\{A^R_\mu(X), A^R_\nu(X')\}\,|\Omega\rangle = \int_0^\infty d\omega \int \frac{d^{d-2}k}{(2\pi)^{d-2}} \frac{\cosh \pi\omega}{\pi^2} \big[e^{i(k\Delta x - \omega\Delta\tau)} \cdot K_{i\omega}(|k|z) K^*_{i\omega}(|k|z') \sum_\lambda \epsilon^\lambda_{\mu,\omega k} (\epsilon^\lambda_{\nu,\omega k})^* + (X \leftrightarrow X')\big]. \tag{3.262}$$

This differs from the scalar anticommutator (3.214) only by the presence of the polarization sum. The Feynman gauge expression (3.260) for the polarization sum leads immediately to the result (3.220) and the conclusion that gauge-invariant correlation functions reproduce their Minkowski expectation values. The argu-

ment for trans-horizon correlators proceeds identically.

3.5 Living on the Edge: A Toy Model for Holographic Reconstruction of Algebras with Centers

Recent studies of holography have led to an intriguing new connection between AdS/CFT and quantum information: the partial holographic reconstruction of the bulk from boundary subregions shares features with quantum error-correcting codes [303]. Perhaps the most prominent manifestation is the fact that a given local bulk operator can be reconstructed on any boundary region that contains the bulk operator in its entanglement wedge [365]. This redundant boundary encoding of bulk information can be precisely analogized with such codes [350, 351], where a certain set of physical (viz. boundary) qubits encodes a smaller set of logical (viz. bulk) qubits. The encoding of the logical bits in the physical ones is robust against the deletion of some of the physical bits, or a change in their state due to noise, and so the encoding scheme is protected against such "errors". This is strongly remniscent of the ability to reconstruct the bulk operator using only a subregion of the boundary.

A few years prior to [303] it was realized [366] that a particular class of tensor network descriptions of quantum states (dubbed MERA, for multiscale entanglement renormalization ansatz [367]) captures certain properties of holography. Most prominently, the entanglement structure in the tensor network reflects the bulk geometry and so the network automatically implements the Ryu-Takayanagi formula. After it was realized that holography is also error-correcting, it was natural to ask if there was a tensor network description that captured this behavior as

well, especially since error-correcting codes are often realized on tensor networks. This question was answered affirmatively by [304], who showed that a particular tensor network (built on the pentagon code) exhibits the quantum-error-correcting features of holography. In this network, dubbed the HaPPY code, a local bulk operator is redundantly encoded in multiple boundary regions, with each such encoding robust to the deletion of the rest of the boundary.

However, bulk physics is not wholly local: there are bulk gauge fields, and of course the graviton, which cannot be described by a local tensor model, and so we proposed an extension of the HaPPY code that models bulk gauge fields. Gauge fields naturally live on the links of a graph, rather than the vertices, as in lattice gauge theory. In the HaPPY code the vertices are the tensors of the network, and the crux of our extension is the addition of degrees of freedom that live on the links instead. Our construction adds a tensor above each link that makes two copies of its input (in a basis specified by hand) which are passed to the two tensors at the vertices adjoining the link.

As a simple example, taking the Hilbert space of each tensor leg to be that of a single spin, we could take the copying tensor to map

$$|+\rangle \to |++\rangle\,,\ |-\rangle \to |--\rangle \tag{3.263}$$

where $|\pm\rangle$ are σ_z eigenstates. With this choice of copying code, the action of σ_z on the link is encoded (i.e. can be reconstructed) on any boundary region whose entanglement wedge contains both of the vertices connected by the link, as in the original HaPPY code. However, if the entanglement wedge contains only one of the vertices adjoining the link, σ_z is encoded on the boundary region but σ_x acting

on the same link is not. On such a boundary region σ_z lies in the center of the algebra of operators on the subregion since it is also encoded on a complementary boundary region, while σ_x is absent from the algebra entirely. One can treat more general systems with a modification of the copying code but even this simple example is enough to describe an (emergent) gauge theory via the toric code, with σ_z mapping to the flux operator and σ_x to the conjugate holonomy.

These reconstruction properties mirror the structure of operator algebras on subregions in lattice gauge theories [216], where the reduced Hilbert space decomposes into superselection sectors labeled by the flux through the entangling surface (the same decomposition that gives rise to the edge mode theory of [218, 217]). Essentially, the flux operators on links piercing the entangling surface must be included in the algebra of operators in the entangling region in order to allow fluctuations of the electric field on the entangling surface, but since the Gauss's law constraint expresses them entirely in terms of operators outside the region they must commute with (i.e are in the center of) the algebra of operators on that region. These flux operators are analogous to σ_z in our construction. On the other hand, the holonomies conjugate to the flux operators on the entangling surface are not elements of the algebra of operators in the entangling region (by necessity, since the conjugate flux is in the center), like σ_x in our model.

It has been suggested [306] that this operator algebra structure leads naturally to the Ryu-Takayanagi formula [307] for the entanglement entropy of subregion of the boundary. Indeed, the R-T formula implies (in the absence of black holes) that the area operator – the flux of gravitons across the R-T surface – can be reconstructed either on the subregion or its complement, and so lies in the center of

the algebra on either subregion. While our model lives on a fixed background network and so cannot model the spacetime fluctuations necessary for a description of quantum gravity, we show that our copying code can be modified to capture contributions to the entanglement entropy that, like the area of the R-T surface, are the dominant contributions in a certain large-N limit.

3.5.1 Introduction

Recent works [303, 304, 305, 306] have introduced models of gauge/gravity duality based on quantum error correcting codes and thus provided a new paradigm for studying holographic systems. The models implement their codes via tensor networks that map bulk logical operators to operators on a code subspace of a larger boundary Hilbert space. Such representations were termed "holographic codes" in [304] and have been shown to exhibit key properties of the AdS/CFT correspondence such as bulk reconstruction and the Ryu-Takayanagi (RT) relation between entanglement in the boundary theory and the area of bulk minimal surfaces [307, 308].

Indeed, as noted in [304], such holographic codes also reproduce an important part of the $1/N^2$ corrections to RT found by Faulkner, Lewkowycz, and Maldacena (FLM) [309]. Recall [309] that with such corrections the entropy S_A of a boundary region A takes the interesting form

$$S_A = \frac{\text{Area}}{4G_N} + S_{\text{bulk}}(\rho_{W(A)}) + \frac{\delta\text{Area}}{4G_N} + \dots \tag{3.264}$$

The first term on the right is the leading-order Ryu-Takayanagi piece, which is

local on the entangling surface and independent of the state. The second accounts for bulk entropy in the entanglement wedge $W(A)$ defined by the RT minimal surface, and is thus generally both non-local and non-linearly dependent on the bulk state. The third is an additional effect from quantum corrections to the RT area, which is distinguished by being both local on the RT surface and linear in the bulk state; i.e. it is an expectation value. The ... denote higher order terms in the $1/N$ expansion. In the codes from [304], the analogous result contains the first two terms on the right-hand side.

It was suggested in [306] that the remaining $\frac{\delta \text{Area}}{4G_N}$ term would also arise naturally from a quantum error correction model containing operators $\mathcal{O}$, associated with the boundary between the entanglement wedge of A and that of its complement $\bar{A}$, that are reconstructible from both A and $\bar{A}$. Such $\mathcal{O}$ must lie in the center of either reconstructed algebra. The terms $\frac{\text{Area}}{4G_N} + \frac{\delta \text{Area}}{4G_N}$ then naturally correspond to aspects of the code that are, in some sense, dependent on the values of operators in this center.

Much of the above structure is familiar from analyses [310, 274, 216] of entropy in lattice gauge theories. In that context, the (electric [216]) algebra of operators acting in a bulk subregion contains the electric fields $E_\ell|_{\partial A}$ along the links ℓ at the boundary of the subregion [310, 274, 216, 218]. And since Gauss's law equates the $E_\ell|_{\partial A}$ with operators spatially separated from A, the boundary electric fields commute with the entire subalgebra on A. In particular, the canonical conjugates of the E_ℓ are closed Wilson loops that pass through ℓ, which are not elements of the subalgebras on either A or $\bar{A}$ when $\ell \in \partial A$.

It is thus natural to consider modifications of the HaPPY code inspired by

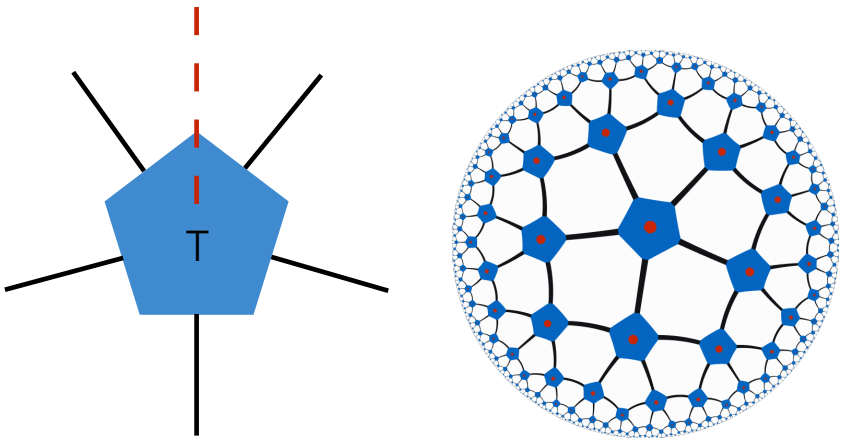
lattice gauge theory and having additional degrees of freedom that live on the links of the bulk lattice. This is done in section 3.5.2 building on the HaPPY pentagon code [304]. As desired, a key feature of our model is the existence of bulk operators that are reconstructible on both a boundary region and its complement. Such properties are derived in section 3.5.3 and follow directly from results of [304]. We then demonstrate in section 3.5.4 that such central elements do indeed endow our model with an FLM-like relation containing analogues of all three terms shown explicitly on the right-hand side of (3.264). Section 3.5.5 concludes with some final discussion. In particular, comparison with lattice gauge theory constructions suggests that the FLM $\frac{\delta \text{Area}}{4G_N}$ term might be usefully reinterpreted as part of the bulk entropy of metric fluctuations in an appropriate extension of the physical bulk Hilbert space.

3.5.2 Edge Mode Construction

The fact that gauge theories are described canonically by a connection and a conjugate electric flux makes it natural to describe these degrees of freedom as living on the links of a discrete graph-like model, as is common in lattice gauge theory. This allows holonomies to be described as paths through the lattice and the Gauss law constraint to be imposed by requiring the electric fields on links attached to any vertex v to sum to the charge at v.

Since we wish to extend holographic codes in a manner reminiscent of bulk gauge theories, we will introduce degrees of freedom below on the links of the tensor network corresponding to the pentagon code of [304]. We will first review the relevant features of this code and then describe the desired augmentations.

The pentagon code is a tiling of a hyperbolic disk where the fundamental unit is a six index tensor T drawn in figure 3.1. The disk has finite size as the code is to be thought of as a model of a holographic CFT with a cutoff. Except at the boundary of the disk, five of the legs of the tensor are connected to adjacent tensors as depicted in figure 3.1. Even at the boundary, we refer to these five as network legs. Each such tensor has one uncontracted index representing a local bulk degree of freedom. If T is chosen to be a perfect tensor, meaning that it describes an isometry from any subset of at most 3 legs to the rest, an operator $\mathcal{O}$ acting on any bulk input can be "pushed" along three of the output legs to three adjacent tensors: the action of $\mathcal{O}$ on T can be replaced by the action of $\mathcal{O}' = T^\dagger O T$ on one of the adjacent tensors.



(a) (b)

Figure 3.1: (a) The fundamental tensor T of the pentagon code showing the bulk leg (dashed line, red in color version) and the network legs (solid lines). (b) These units are contracted along their networks legs to form a pentagonal tiling of the hyperbolic plane.

This procedure allows us to push local bulk operators to the boundary, as the negative curvature of the hyperbolic plane ensures that each tensor has at least

three legs pointing toward the boundary in a suitable sense. Since T is a perfect tensor, one can also show [304] that the entropy is given by an FLM-like formula having analogues of the first two terms on the right-hand side of (3.264).

We wish to introduce additional degrees of freedom modeling bulk gauge fields in a way that largely preserves these properties. As a first guess, one might add to each of the non-bulk legs of the fundamental unit a three index tensor G_{ijk}, whose role in the network is to link two adjacent tensors to a common input modelling the electric flux of some bulk gauge field. One might then choose the tensor structure

$$G = \delta_{ij}\delta_{jk}\delta_{ik} \tag{3.265}$$

to impose flux conservation along each each link in the network. This new fundamental unit is drawn in figure 3.2. However, the values on all the network legs are

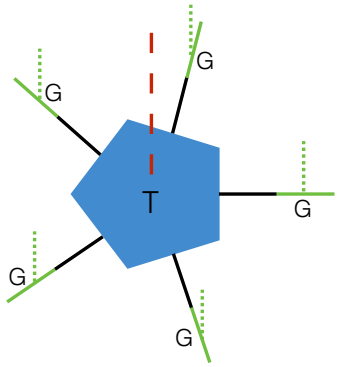


Figure 3.2: An unsuccessful first attempt to add edge degrees of freedom. A copy of the tensor G has been attached to each of the 5 network legs of the tensor T from figure 3.1. The bulk input leg of G is drawn in small dashes. This attempt does not succeed, as the tensor annihilates bulk states lacking particular correlations among the 6 bulk inputs.

then determined by the inputs to the associated Gs, so there is no room for further input from the bulk leg of T. Indeed, the network just described will annihilate

all bulk states orthogonal to a space in which the T inputs are determined by the G inputs (and where the G inputs also satisfy a further set of constraints).

This unfortunate issue can be resolved by considering a model in which the above 6 bulk inputs are manifestly independent. We do so by extending the fundamental unit T to the 6-fold tensor product $\otimes_{m=1}^{6} T = T \otimes T \otimes T \otimes T \otimes T \otimes T$ and again connecting these units as in the pentagonal tiling of the hyperbolic disk shown in figure 3.1 (b). Each factor in the resulting tensor product will be called a "copy" of the network: the first copy will be treated as an independent HaPPY network, while the additional copies will be contracted with our 3-legged tensor (or indeed any isometry) G as described below.

Thus far our network has 6 bulk input legs at each vertex. We will turn 5 of these into inputs associated with edges instead. Consider some particular edge in the interior of the disk and choose one input leg from each of the two vertices it connects (to simplify the figures, both input legs are chosen from the same copy). Our edge-mode code is constructed by contracting these legs with two legs of the tensor G; see figure 3.3. We will treat these two legs of G as output legs; the remaining input is naturally associated with the edge under consideration. Doing so for each edge uses 5 of the bulk legs at each vertex, leaving the 6th free to serve as a normal bulk input at each vertex just as in the original code from [304]. To be concrete, we take this 6th bulk input to live in the first copy of the network. Figure 3.4 shows a pictorial representation of the full edge-mode code including all six copies the pentagon code. Note that we have added one G for every two bulk legs, and thus also for every two Ts.

The resulting code defines an isometry from the bulk degrees of freedom to

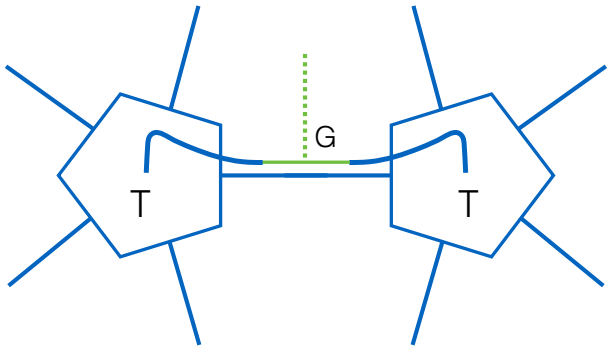
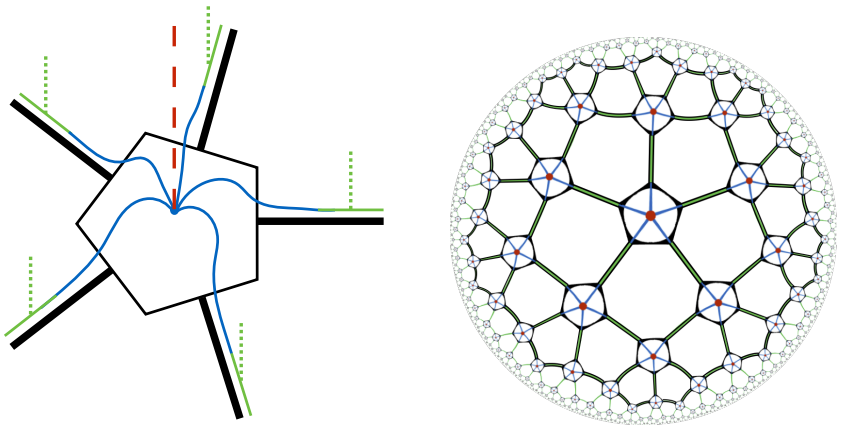


Figure 3.3: Our code is built from 6 copies of the code from [304] by contracting the tensor G with a pair of neighboring bulk inputs. The relevant two T-tenors are shown here, where we have chosen them both to be part of the same copy of the pentagon code.



(a) (b)

Figure 3.4: (a) The structure near each vertex of our edge-mode code. The thick black legs carry 5 indices. The central input (long dashes, red in color version) corresponds to a bulk matter field as in [304] while the inputs on each edge (short dashes, green in color version) are to be interpreted as degrees of freedom of a bulk gauge field. (b) A sketch of the full edge-mode code.

the boundary, and therefore has many of the same features as the code described in [304]. This is because one may view this edge-mode code as six copies of the HaPPY pentagon code together with a set of G tensors interposed between these codes and the bulk state. Since the G tensors are isometries, composing them in

this way with the original HaPPY network yields another isometry. As described below, this observation allows us to import all of the main technology from [304] including operator pushing, the greedy entanglement wedge construction, and the Ryu-Takayanagi formula for entanglement entropy. However, the additional tensors G introduce certain subtleties which we will discuss in depth.

3.5.3 Operators in the Center and Bulk Reconstruction

We now consider properties of our code associated with subregion duality, showing that our model leads to the bulk reconstruction of algebras with centers. This reproduces the structure suggested in [306]. Here we view the tensor network of our edge-mode code as an isometry from a bulk Hilbert space $\mathcal{H}_{\rm bulk}$ defined by the set of all bulk inputs (both edge and vertex) to a boundary Hilbert space $\mathcal{H}_{\rm bndy}$ defined by the set of network links that reach the boundary of the hyperbolic disk. For simplicity of notation, we follow standard practice and use the above isometry to identify $\mathcal{H}_{\rm bulk}$ with its image $\mathcal{H}_{\rm code}$ in $\mathcal{H}_{\rm bndy}$. Bulk operators are then maps from $\mathcal{H}_{\rm bulk} = \mathcal{H}_{\rm code}$ to itself.

As in [303, 304, 306], we shall say that a bulk operator $\mathcal{O}$ lies in the algebra $\mathcal{M}_A$ that can be reconstructed from a region A of the boundary if (and only if) there exists an operator $\mathcal{O}^{(A)}$ with support in A such that

$$\mathcal{O}^{(A)} \left|\psi\right\rangle = \mathcal{O} \left|\psi\right\rangle \qquad \forall \left|\psi\right\rangle \in \mathcal{H}_{\rm code} \,. \tag{3.266}$$

Note that, with this definition, bulk operators $\mathcal{O}_1 \in \mathcal{M}_{A_1}$ and $\mathcal{O}_2 \in \mathcal{M}_{A_2}$ for

non-intersecting regions A_1 and A_2 must commute. In detail, on $\mathcal{H}_{\text{code}}$ we have

$$\begin{aligned}[\mathcal{O}_1, \mathcal{O}_2]\,|\psi\rangle &= \mathcal{O}_1^{(A_1)}\mathcal{O}_2\,|\psi\rangle - \mathcal{O}_2^{(A_2)}\mathcal{O}_1\,|\psi\rangle \\ &= \mathcal{O}_1^{(A_1)}\mathcal{O}_2^{(A_2)}\,|\psi\rangle - \mathcal{O}_2^{(A_2)}\mathcal{O}_1^{(A_1)}\,|\psi\rangle \\ &= \left[\mathcal{O}_1^{(A_1)}, \mathcal{O}_2^{(A_2)}\right]|\psi\rangle = 0. \end{aligned} \tag{3.267}$$

In the first step we have used the fact that bulk operators preserve $\mathcal{H}_{\text{code}}$, while the final step uses the fact that all operators in A_1 commute with those in A_2. As a result, any bulk $\mathcal{O}$ lying in both $\mathcal{M}_{A_1}$ and $\mathcal{M}_{A_2}$ must be a central element of both algebras.

This is precisely the structure suggested by [306] as the natural quantum-error-correction model of FLM corrections to the Ryu-Takayanagi relation. It is useful to contrast this situation with that of the HaPPY code, where reconstruction on A succeeds for any operator in the greedy entanglement wedge $w^*(A)$ (or greedy wedge for short) defined by the greedy algorithm of [304].[1] The boundary of this greedy wedge consists of two parts, one lying on the boundary of our hyperbolic disk and the other in the interior of the disk. We refer to the latter as the greedy entangling surface γ_A^*. In our edge-mode code, we define a corresponding greedy wedge and γ_A^* using only the tensors T associated with our 6 copies of the pentagon code. When we then add the additional G tensors, some bulk operators (G-inputs) act on links that straddle the resulting γ_A^*. Only certain bulk operators acting on such links can be reconstructed on A, and those operators will typically lie in the

[1] The greedy wedge associated with a region A on the boundary is constructed by first taking all tensors with at least three legs contained in A. Next all tensors with at least three legs contracted with this set of tensors are included, and this procedure continues until there are no more tensors to add.

center of the algebra of operators on A.

The essential point can be illustrated by considering only a pair of Ts ($T_{L/R}$) that are linked by a single G as in figure 3.3. We take the one bulk edge input to be a single qubit that feeds into G, and we take G to map

$$|0\rangle \to |00\rangle \quad \text{and} \, |1\rangle \to |11\rangle \tag{3.268}$$

as in (3.265). For this reason we refer to G as the copying tensor below. The perfect tensors $T_{L/R}$ each have 4 uncontracted legs which we treat as proxies for the left and right halves of the boundary.

One bulk operator of interest is the Pauli σ_z defined by $\sigma_z|0\rangle = -|0\rangle$, $\sigma_z|1\rangle = |1\rangle$ acting on the bulk edge input. The structure of G allows one to push σ_z through G onto either output leg of G: writing $G = |00\rangle\langle 0| + |11\rangle\langle 1|$, it follows that

$$G\sigma_z = -|00\rangle\langle 0| + |11\rangle\langle 1| = \sigma_z^{(L)} G = \sigma_z^{(R)} G. \tag{3.269}$$

Here $\sigma_z^{(L/R)}$ denotes a corresponding Pauli matrix acting on the output of G that feeds into $T_{L/R}$ as depicted in figure 3.5. It follows that we can reconstruct σ_z in

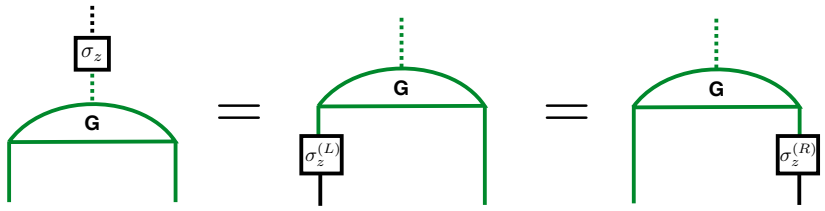


Figure 3.5: Pushing σ_z through G

the left boundary by pushing $\sigma_z^{(L)}$ through T_L. But the same procedure allows us to reconstruct σ_z as an operator acting only on R by pushing $\sigma_z^{(R)}$ through T_R.

So as above σ_z must lie in the center of $\mathcal{M}_L$, and also of $\mathcal{M}_R$.

As a result, the other Pauli operators σ_x, σ_y at our bulk edge input cannot be reconstructed from either L or R alone. But these operators can still be reconstructed if we are granted simultaneous access to both sets of boundary sites. Indeed, σ_x satisfies

$$G\sigma_x = |00\rangle\langle 1| + |11\rangle\langle 0| = \sigma_x^{(L)}\sigma_x^{(R)}G. \tag{3.270}$$

We may then push each of $\sigma_x^{(L/R)}$ through its respective T to the boundary, and so any boundary region whose greedy entanglement wedge includes both T_L and T_R will be able to reconstruct the σ_x that acts between them. σ_y will also be reconstructible on the region, and σ_z will not be central.

Returning to the full edge-mode code, we may consider the greedy wedge for any region A on the boundary. Bulk operators in this wedge may be generated by taking sums and products of the following three types of 'local' operators: (i) operators that act on inputs at a single vertex, (ii) operators acting on a single link that lies in the interior of the greedy wedge, and (iii) operators acting on a single link that straddles the corresponding greedy entangling surface.

Operators of type (i) are precisely the bulk operators defined in [304] and act on the first copy of the code constructed in section 3.5.2, so as in [304] such operators lie in $\mathcal{M}_A$. Operators of type (ii) were shown above to be equivalent to a pair of operators acting on tensors T on either side of the link, thus they also lie in the interior of the greedy wedge. In particular, each member of the pair acts on the leg of T that was interpreted in [304] as a bulk input of a pentagon code.

So again such operators lie in $\mathcal{M}_A$.

For operators of type (iii), there are two cases. When the operator acts like σ_z above it can be replaced by a single operator acting on the adjacent tensor T lying inside the greedy wedge. It acts on a leg that was interpreted in [304] as a bulk input of a pentagon code and so can be reconstructed in A. However, other operators on this edge input cannot generally be reconstructed in A. Indeed, when the relevant edge e also straddles the greedy entangling surface of the complementary ($\bar{A}$) boundary region[2] it follows as above that σ_z lies in the center of both $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$.

We close this section by noting that our model admits a broad class of generalizations preserving the above properties. The point is that the above arguments depended only on G copying the input into both outputs. In particular, this makes G and isometry from any one leg to the remaining two, allowing us to push operators of type (ii) as above to a two site operator completely contained in the greedy wedge. This remains the case if we break the symmetry between the first (input) leg of G and the output legs (second and third) by replacing (3.268) with any map of the form

$$G : H_{\text{in}} \to (H_{\text{in}} \otimes H_{\text{aux}})_L \otimes (H_{\text{aux}} \otimes H_{\text{in}})_R, \quad G|\alpha\rangle \mapsto |\alpha\rangle|\psi(\alpha)\rangle|\alpha\rangle. \tag{3.271}$$

Here $\{|\alpha\rangle\}$ is a basis for the input Hilbert space H_{in} (which we call the copying

[2]For the entanglement wedges defined by minimal surfaces in gauge/gravity duality, this would always be true in a pure state as the entanglement wedge of A is the complement of that for $\bar{A}$. In contrast, in the model of [304] there can be a region that lies in neither the greedy wedge for A nor that for $\bar{A}$. The existence of such a region is to be regarded as an artifact of the model associated with discretization of the bulk spacetime; see section 3.5.4 for further discussion.

basis), $|\psi(\alpha)\rangle$ is a state on an auxiliary product Hilbert space $H_{\rm aux} \otimes H_{\rm aux}$, and the tensor factors marked L, R in (3.271) correspond respectively to the two output legs of G. As a concrete example, one may consider

$$G : |0\rangle \to |0000\rangle\,, \qquad |1\rangle \to |1\rangle \otimes \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) \otimes |1\rangle\,, \tag{3.272}$$

which has

$$|\psi(0)\rangle = |00\rangle\,, \qquad |\psi(1)\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right). \tag{3.273}$$

The generalization (3.271) allows us to make direct contact with the extended-lattice discussion [218] of entropy in lattice gauge theories. Ref. [218] considered regions of a lattice with the boundary γ of the region taken to intersect only links (i.e., no vertices lie on γ). Each link was then divided into two parts, with separate Hilbert spaces defined on the parts on either side of γ. The original link carries a Hilbert space $H_{\rm in}$, which may be thought of as defined by the 'electric' basis $\{|R\rangle\}$ with R ranging over all representations of the gauge group. The entropy of the chosen region is defined in [218] by taking, for each R, the state on the corresponding two half-links to be the maximally-entangled pure state $|\tilde{\psi}\rangle_{RR}$ on two copies (one for each half-link) of the representation R. The operation mapping the original lattice to the extended lattice of half-links may then be cast in the form (3.271) by taking $\{|\alpha\rangle\} = \{|R\rangle\}$ and $|\psi(R)\rangle$ to be supported on some $\dim(R)$ dimensional subspaces of each copy of $H_{\rm aux}$, and in which it is unitarily equivalent to $|\tilde{\psi}\rangle_{RR}$. Here $H_{\rm in}$ has dimension equal to the (potentially infinite) number of representations, and the dimension of $H_{\rm aux}$ is correspondingly infinite. It is thus

natural to apply this version of our edge-mode code when the bulk theory on the links is a lattice gauge theory.

3.5.4 Subsystem Entropy and Edge-Mode Codes

We now turn to the issue of how our code relates to various discussions of holographic entropy. In particular, following [304] we show that the entropy of a boundary region A can be written in a form analogous to the explicit terms in the FLM formula (3.264), but in constrast to the original HaPPY code, for general codes of the form (3.271) there is a non-trivial term playing the role of $\frac{\delta \text{Area}}{4G_N}$. This is to be expected from the algebraic entropy analysis of [306], though we give a direct calculation in section 3.5.4. To allow proper comparison, we first review the results of [306] before beginning our computation. We save discussion of the result for section 3.5.5, where examination of the lattice gauge theory edge-mode codes described below (3.271) and comparison with [218], [216] will suggest that, in a corresponding analysis of linearized gravity, the FLM $\frac{\delta \text{Area}}{4G_N}$ term would be precisely the difference between an extended-lattice bulk entropy analogous to that of [218] and an algebraic entropy analogous to that of [216].

The general structure of entropy in holographic codes was studied in [306]. For an error-correcting code with complementary recovery[3] it takes the form

$$S_A = S\left(\rho_A, \mathcal{M}_A\right) + \text{tr}\left(\rho \mathcal{L}_A\right) \tag{3.274}$$

in a code state ρ, where A is a subsystem of the bulk Hilbert space. The first term is the entropy of ρ with respect to the von Neumann algebra $\mathcal{M}_A$ associated

[3] As for [304], our code satisfies this requirement for a region A when $\gamma_A^* = \gamma_{\bar{A}}^*$.

to A, while the second is the expectation value of a linear operator ($\mathcal{L}_A$) lying in the center of $\mathcal{M}_A$. It was proposed that the $\mathrm{tr}\,(\rho\mathcal{L}_A)$ corresponds to the first and third ($\frac{\mathrm{Area}}{4G_N}$ and $\frac{\delta\mathrm{Area}}{4G_N}$) terms in (3.264), while the $S\,(\rho_A, \mathcal{M}_A)$ term corresponds to the second ($S(\rho_{W(A)})$) term in (3.264).

If $\mathcal{M}_A$ has a non-trivial center, any operator in $\mathcal{M}_A$ is block-diagonal, with blocks labeled by eigenvalues α of the center operators:

$$\mathcal{O}_A = \bigoplus_\alpha \mathcal{O}_A^{(\alpha)}. \tag{3.275}$$

This is also true of any state ρ_A, which we write as $\rho_A = \bigoplus_\alpha p_\alpha \rho_A^{(\alpha)}$ in terms of normalized states $\rho_A^{(\alpha)}$. The entropy of ρ_A with respect to $\mathcal{M}_A$ can then be defined as $-\mathrm{tr}\,(\rho_A \ln \rho_A)$ using the representation (3.275), so we have

$$S\,(\rho_A, \mathcal{M}_A) = -\sum_\alpha p_\alpha \ln p_\alpha + \sum p_\alpha S\left(\rho_A^{(\alpha)}\right), \tag{3.276}$$

with $S\left(\rho_A^{(\alpha)}\right) = -\mathrm{tr}\left(\rho_A^{(\alpha)} \ln \rho_A^{(\alpha)}\right)$. This is the same as the entanglement entropy we would obtain by tracing the full system state ρ over $\bar{A}$ and then decohering the blocks by deleting all the matrix elements that are not block-diagonal:

$$S\,(\rho_A, \mathcal{M}_A) = S\,([\rho_A]_{\mathrm{dec}}) \tag{3.277}$$

where $[\rho_A]_{\mathrm{dec}}$ is the reduced density matrix $\rho_A = \mathrm{tr}_{\bar{A}}\rho$ after decohering the blocks.

By contrast, the second contribution to the entanglement in (3.274) depends

on the details of the code. Since $\mathcal{L}_A$ lies in the center of $\mathcal{M}_A$, it takes the form

$$\mathcal{L}_A = \bigoplus_\alpha s_\alpha \mathbb{1}_\alpha. \tag{3.278}$$

$\mathbb{1}_\alpha$ is the identity matrix on the block α and the s_α are just numbers.

Since the algebras relevant to the original HaPPY code had trivial centers, a primary goal of our work is to manifest this structure in a holographic code. When the center is trivial there is only one value for the index α and (3.276) becomes the tautalogy $S(\rho_A) = S(\rho_A)$ and the $\mathrm{tr}(\rho\mathcal{L}_A)$ term becomes independent of the state (though it still depends on A). As reviewed in section 3.5.4 below, in the HaPPY code this constant is given by the logarithm of the bond dimension χ times the length of the minimal surface anchored to the boundary of A. This term is then interpreted as a model for the leading Ryu-Takayanagi term in the holographic entanglement entropy, so that it is natural to think of χ as large. Similarly, the $S(\rho_A, \mathcal{M}_A)$ term is interpreted as modeling the second term in (3.264).

Entropy from Edge Modes

We now compute the entropy of an arbitrary boundary region A in our edge-mode code with an arbitrary bulk state. To the extent that our code satisfies the assumptions of [306], the result must take the form (3.274). The goal of the calculation is thus to determine the explicit form of $\mathcal{L}_A$ for our edge-mode code, as well as to take into account small violations of the complementary recovery assumption of [306] (which are also present in the original code [304]). We will see that $\mathcal{L}_A$ takes the form of a local density on the entangling surface, as appropriate for the term that gives rise to the $\frac{\mathrm{Area}}{4G}$ and $\frac{\delta\mathrm{Area}}{4G}$ pieces of the holographic

entanglement.

It is useful to begin by reviewing results for the pentagon code of [304] on which our edge-mode code is strongly based. In order to arrive an an explicit FLM-like form, we rewrite the original arguments of [304] as a tensor network computation. We begin by thinking of the code as a map from bulk states $|\psi\rangle_{bulk}$ to boundary states $|\psi\rangle_{bndy}$. In figure 3.6, this map is thought of as a tensor network built out of 3 parts: the greedy wedges $w^*(A)$, $w^*(\bar{A})$, and a residual bulk region $X = \overline{\left(w^*(A) \cup w^*(\bar{A})\right)}$ excluded from both $w^*(A)$ and $w^*(\bar{A})$. Here the long overline denotes the complement in the bulk.

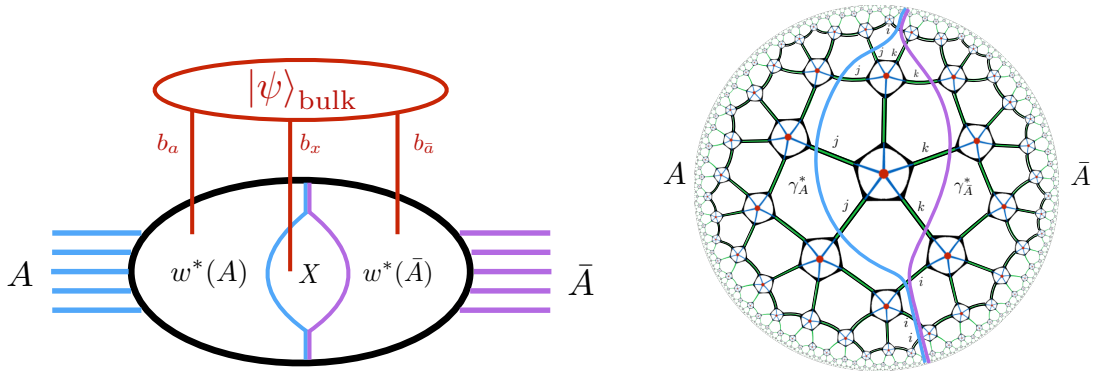


Figure 3.6: In (a) the network that computes the boundary state, broken into pieces corresponding to the wedges $w^*(A)$, $w^*(\bar{A})$, and the residual region X. In general all three regions receive inputs from bulk legs. In (b) we label the links cut by $\gamma^*(A)$ and $\gamma^*(\bar{A})$ with j and k respectively, and the links cut by both with i.

We may of course also use this network to map bulk density matrices ρ_{bulk} to boundary density matrices ρ_{bndy}. Although ρ_{bndy} is defined on the entire boundary, tracing over $\bar{A}$ gives a reduced density matrix ρ_A on boundary region A. The associated tensor network is shown in figure 3.7, displaying the different roles

performed by network links cut by γ_A^*, those cut by $\gamma_{\bar{A}}^*$, and those cut by both. We then recall that $w^*(\bar{A})$ defines an isometry (which we will also denote as $w^*(\bar{A})$)

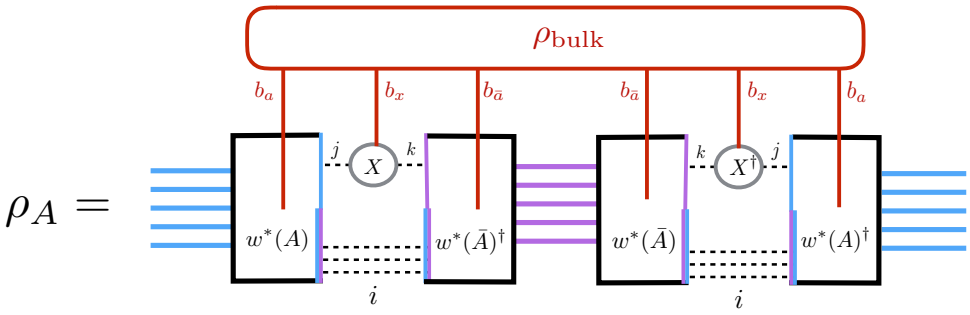


Figure 3.7: The reduced density matrix ρ_A on boundary region A described as a circuit and broken into pieces corresponding to $w^*(A)$, $w^*(\bar{A})$, and X. The j links describe network edges cut by γ_A^*, the k links describe network edges cut by $\gamma_{\bar{A}}^*$, and the i links describe those cut by both.

from the associated bulk indices and the network edges cut by $\gamma_{\bar{A}}^*$ to $\bar{A}$, i.e. it satisfies

$$[w^*(\bar{A})]^\dagger w^*(\bar{A}) = \mathbb{1}, \tag{3.279}$$

with $\mathbb{1}$ being the identity on the space of inputs. As a result, the tensor network may be simplified to that shown in figure 3.8.

A key part of this network representation of ρ_A is given by the state $\rho_{\gamma_A^*}$, defined on the Hilbert space $H_{w^*(A)} \otimes H_{\gamma_A^*}$ as show in figure 3.9. Here $H_{w^*(A)}$ is the space associated with bulk inputs to $w^*(A)$, and $H_{\gamma_A^*}$ is the space defined by network edges cut by γ_A^*. Figure 3.8 then implies that we may write

$$\rho_A = w^*(A)\rho_{\gamma_A^*} w^*(A)^\dagger \tag{3.280}$$

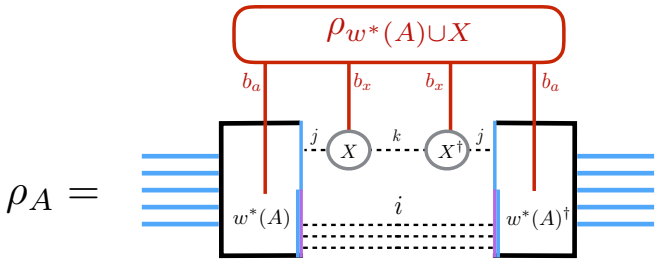


Figure 3.8: A simpler network for ρ_A obtained from figure 3.7 by using the fact that $w^*(\bar{A})$ defines an isometry from the indices k, i, and $b_{\bar{a}}$ to the boundary $\bar{A}$.

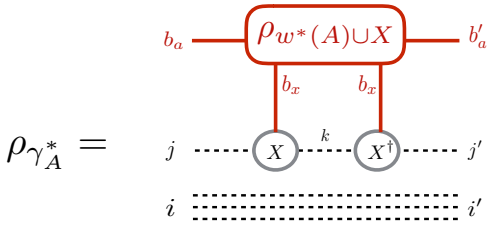


Figure 3.9: The state $\rho_{\gamma_A^*}$ used to compute S_A.

in terms of the isometry $w^*(A)$. Since isometries preserve von Neumann entropy, the entropy of the boundary region A may be written

$$S_A = -\mathrm{tr}\left(\rho_A \ln \rho_A\right) = -\mathrm{tr}\left(\rho_{\gamma_A^*} \ln \rho_{\gamma_A^*}\right). \tag{3.281}$$

To the extent that we can ignore the excluded region X, we have that $\rho_{w^*(A)\cup X} \approx \rho_{w^*(A)}$ and the density matrix $\rho_{\gamma_A^*}$ is

$$\rho_{\gamma_A^*} \sim \rho_{w^*(\bar{A})} \otimes \mathbb{1}_i \otimes \mathbb{1}_j \tag{3.282}$$

up to normalization. The von Neumann entropy S_A is then precisely

$$S_A = S(\rho_{w^*(A)}) + |\gamma_A^*| \ln \chi \,, \tag{3.283}$$

in terms of the bond dimension χ of each network edge and the number $|\gamma_A^*|$ of edges cut by γ_A^* (i.e. the total number of i and j indices).

More generally, the region X introduces further corrections to (3.283). While such corrections are difficult to compute explicitly, subadditivity and the Araki-Lieb inequality ($|S_B - S_C| \leq S_{BC} \leq S_B + S_C$) can be used to bound departures from this estimate in terms of χ and the number of j edges as defined in figure 3.7 (those cut by γ_A^* but not by $\gamma_{\bar{A}}^*$). In special cases X can be quite large, but this is not generally the case [304]. Indeed, X can often be made to vanish by moving a small number of boundary points from A to $\bar{A}$ and/or from $\bar{A}$ to A. It is thus natural to think of think of X as an artifact of the discrete toy model used here. As a result, although there is no limit of our model in which the region X can be systematically neglected in all cases, as in [304] we choose to ignore the region X when considering implications for gauge/gravity duality. We thus consider only cases with trivial $X = \emptyset$ below.

With this technology in hand, we now turn to our edge mode code. Our code attaches bulk states $|\psi\rangle_{\text{bulk, EMC}}$ to a product of HaPPY codes via an isometry $\mathscr{G}$ built from the copying tensors G. The end result is thus just what would be obtained by feeding the state

$$|\psi\rangle_{\text{bulk, HaPPY}} = \mathscr{G} \, |\psi\rangle_{\text{bulk, EMC}} \tag{3.284}$$

into a product of HaPPY codes. Thus all that remains is to replace $\rho_{w^*(A)\cup X}$ in figure 3.9 with the corresponding density matrix obtained from the state with Gs inserted.

Now, since G is an isometry from its single input to its pair of outputs, adding G-tensors with both outputs in $w^*(A)$ will not change the entropy of figure 3.9 and may be ignored. Furthermore, the trace over $w^*(\bar{A})$ in passing from figure 3.6 to figure 3.8 removes all G's with both outputs in $w^*(\bar{A})$. When $X = \emptyset$, G has no outputs in X.

Thus the only remaining Gs to consider are those with one output in each of $w^*(A)$, $w^*(\bar{A})$. They will act on bulk indices which we may organize in pairs containing one unprimed index (a ket index of $\rho_{w^*(A)}$) and the corresponding primed index (a bra index of $\rho_{w^*(A)}$). As shown in figure 3.10, the $w^*(\bar{A})$ output legs of the pair of G-tensors acting on these bulk indices are contracted by the above-mentioned trace over $w^*(\bar{A})$.

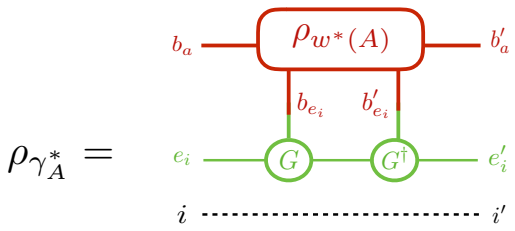


Figure 3.10: A state $\rho_{\gamma_A^*}$ that may be used to compute S_A in our edge-mode code when $X = \emptyset$. Every leg i cut by γ_A^* has an associated b_{e_i} and e_i, although only one such leg is drawn for simplicity.

The effect of these final Gs on the entropy is easy to understand by thinking about the action of a single G on any pure-state input $|\psi\rangle = \sum_\alpha c_\alpha |\alpha\rangle$. Due to

the copying property of G in (3.271), the action of G on $|\psi\rangle$ may be thought of as a von Neumann measurement; i.e., as a unitary transformation that entangles the original system (here $(H_{\text{in}})_L$, which we take to lie in $w^*(A)$) with a 'measuring apparatus' $(H_{\text{in}} \otimes H_{\text{aux}})_R$ in $w^*(\bar{A})$. At the same time, it also creates further entanglement with $(H_{\text{aux}})_L$. Tracing over the $w^*(\bar{A})$-output then decoheres the state into the copying basis $\{\alpha\}$ so that $\rho_{\gamma_A^*}$ is block-diagonal in this basis, and the entanglement with $(H_{\text{aux}})_L$ means that the $|\alpha\rangle\langle\alpha|$-blocks appear tensored with the state

$$\rho_{\text{aux}}(\alpha) = \text{tr}_{(H_{\text{aux}})_R} \left(|\psi(\alpha)\rangle\langle\psi(\alpha)| \right). \tag{3.285}$$

on $(H_{\text{aux}})_L$. Thus we have

$$\rho_{\gamma_A^*} = \bigoplus_{\{\alpha\}} p_{\{\alpha\}} \left(\left[\rho_{\gamma_A^*}\right]_{\{\alpha\}} \otimes \rho_{\text{aux}}(\{\alpha\}) \right), \tag{3.286}$$

with $p_{\{\alpha\}} = \prod_i |c_{\alpha_i}|^2$, $\rho_{\text{aux}}(\{\alpha\}) = \otimes_i \, \rho_{\text{aux}}(\alpha_i)$, and i again ranging over all links cut by both γ_A^* and $\gamma_{\bar{A}}^*$. The relevant $\left[\rho_{\gamma_A^*}\right]_{\{\alpha\}}$ may be described by introducing the α-decohered bulk state

$$\left[\rho_{w^*(A)}\right]_{\text{dec}} = \bigoplus_{\{\alpha\}} p_{\{\alpha\}} \left[\rho_{w^*(A)}\right]_{\{\alpha\}}. \tag{3.287}$$

The $\left[\rho_{\gamma_A^*}\right]_{\{\alpha\}}$ are then given by figure 3.9 (here with $X = \emptyset$ and thus no j, j' indices) in terms of the $\left[\rho_{w^*(A)}\right]_{\{\alpha\}}$ defined by (3.287). Putting everything together, we draw the network representation of (3.286) in figure 3.11.

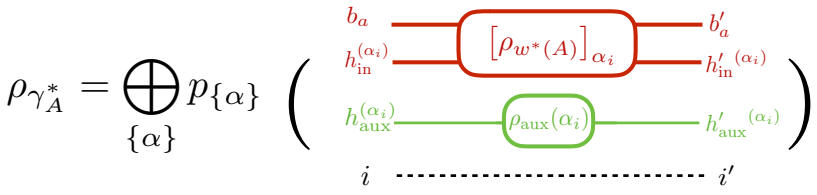


Figure 3.11: The state $\rho_{\gamma_A^*}$ in terms of the states $\left[\rho_{\gamma_A^*}\right]_{\{\alpha\}}$ obtained by decohering the tensors G in figure 3.10 with respect to the α basis. The index $h_{\text{in}}^{(\alpha_i)}$ refers to the subspace of $(H_{\text{in}})_L$ at link i associated with the eigenvalue α, and similarly for $h_{\text{aux}}^{(\alpha_i)}$. Again we have drawn only one link i for simplicity.

This gives

$$\begin{aligned} S_A &= |\gamma_A^*| \ln \chi + \sum_{\{\alpha\}} p_{\{\alpha\}} \left(-\ln p_{\{\alpha\}} + S(\left[\rho_{w^*(A)}\right]_{\{\alpha\}}) + S\left(\rho_{\text{aux}}(\{\alpha\})\right)\right) \\ &= S(\rho_A, \mathcal{M}_A) + |\gamma_A^*| \ln \chi + \sum_{\{\alpha\}} p_{\{\alpha\}} S\left(\rho_{\text{aux}}(\{\alpha\})\right) , \end{aligned} \tag{3.288}$$

where χ now denotes the total bond dimension associated with the T tensors (i.e. 6 times that of the HaPPY code for the code described in section 3.5.2). Here we have used (3.276) to identify the "bulk entanglement" term $S(\rho_A, M)$. Note that (3.288) takes Harlow's form (3.274) if we define $\mathcal{L}_A$ by (3.278) and make the further identification

$$s_{\{\alpha\}} = |\gamma_A^*| \ln \chi + S\left(\rho_{\text{aux}}(\{\alpha\})\right) = \sum_i \left[\ln \chi + S\left(\rho_{\text{aux}}(\alpha_i)\right)\right] . \tag{3.289}$$

This manifestly takes the form of a local density on the entangling surface. The contribution of the first term in (3.289) to S_A is independent of the state, analogous to the leading Ryu-Takayanagi piece in the entropy. The contribution of

the second piece to S_A linearly second depends on the bulk state, like the $\frac{\delta \text{Area}}{4G}$ correction of [309].

As a concrete example, consider the copying tensor defined in (3.272). It is easy to show that $s_0 = 0$ and $s_1 = \ln 2$.

3.5.5 Discussion

We have constructed edge-mode holographic codes by composing (copies of) the HaPPY pentagon code [304] with certain 'copying tensors' G (3.271). The results provide toy models for holography that implement the structure described in [306]. In particular, subregions A of the boundary allow the reconstruction of bulk algebras M having a non-trivial center associated with the interior boundary of the (greedy) entanglement wedge $w^*(A)$, i.e. with the holographic code analogue of the Ryu-Takayanagi minimal surface. As a result, subject to the same caveats as for the original HaPPY code [304], our model gives rise to an FLM-like relation (3.274). We expect that a similar edge-mode extension can be applied to other holography-inspired codes including [311, 312, 305].

In particular, the linear operator $\mathcal{L}_A$ of [306] receives a contribution that depends on the choice of copying tensor G. For general G this term depends non-trivially on the bulk state. This behavior is in contrast to that of contributions from the $|\gamma_A^*| \ln \chi$ term in $\chi_{\{\alpha\}}$. Because the $|\gamma_A^*| \ln \chi$ term does not depend on α, it contributes $\sum_{\{\alpha\}} p_{\{\alpha\}} |\gamma_A^*| \ln \chi = |\gamma_A^*| \ln \chi$ to the entropy for any bulk state.

Due to this distinction, and following the spirit of [306], it is natural to think of the $|\gamma_A^*| \ln \chi$ term in (3.288) as modeling the Ryu-Takayanagi term in (3.264), the $S(\rho_A, \mathcal{M}_A)$ term as corresponding to $S_{bulk}(\rho_{w(R)})$, and the remaining term

$\sum_{\{\alpha\}} p_{\{\alpha\}} S\left(\rho_{\rm aux}(\{\alpha\})\right)$ from $\mathrm{tr}\left(\rho \mathcal{L}_A\right)$ as modeling FLM's $\frac{\delta \mathrm{Area}}{4G_N}$. We expect that this is roughly correct, though it remains to be verified in detail due to a subtle difference between [306] and the approach of FLM. The point here is that to identify a part of $\mathrm{tr}\left(\rho \mathcal{L}_A\right)$ with $\frac{\delta \mathrm{Area}}{4G_N}$ we must take the bulk state to include propagating metric fluctuations. But the analysis [309] of FLM treated the bulk as a 'normal' quantum field theory which could be defined on any metric background, and in particular on backgrounds with conical defects. This is not the case for metric fluctuations which propagate consistently only when the background is on-shell, and so a conclusive result awaits a more complete re-analysis of FLM including a careful treatment of bulk gravitons.[4]

However, perturbative gravity is a gauge theory having much in common with Yang-Mills theory. It is thus interesting to consider in detail the form of (3.288) when the bulk degrees of freedom are taken to describe Yang-Mills. As described at the end of section 3.5.3, it is natural to do so by introducing a lattice gauge theory on the links of the network links and attaching the bulk state to 5 pentagon codes using copying tensors that transform this lattice into an extended lattice as in [274, 218]. Only 5 pentagon codes are required as, for the moment, we suppose that the gauge theory defines all bulk degrees of freedom.

The resulting system can now be viewed in two different ways. The viewpoint

[4] A further source of confusion, though not real difficulty, is the fact that [306] allows an arbitrary bulk state to be considered independent of any choice of background, while the semi-classical approach of [309] naturally correlates the bulk state with the background. In particular, this approach generally selects bulk state in which deviations from the background metric have vanishing expectation value, so that there is no explicit first-order contribution to $\frac{\delta \mathrm{Area}}{4G_N}$ from linearized gravitons and, instead, this term receives contributions only from back-reaction and quadratic terms at the next order. But the freedom to expand around a different background makes clear that, in principle, this $\frac{\delta \mathrm{Area}}{4G_N}$ would indeed receive a linear contribution from linearized metric fluctuations.

used thus far is that the bulk system consists of lattice gauge theory on the network links and that we act on this system with our edge-mode code to obtain the associated boundary state. This leads to (3.288) and the above identification of terms. However, the result can equally-well be viewed as a bulk system defined by an *extended* lattice, in which each link has been replaced by a pair of half-links, acted on by a code that is precisely the tensor product of 5 copies of the HaPPY pentagon code. We may then use the result (3.283), taking the first term on the right to be the entropy $S(\rho^{ext}_{w^*(A)})$ defined by the extended bulk lattice. The entropy of such extended lattice states was discussed in [218], and letting α_i range over representations R_i of the gauge group as at the end of section 3.5.3, takes the form

$$\begin{aligned} S(\rho^{ext}_{w^*(A)}) &= \sum_{\{R\}} p_{\{R\}} \left(-\ln p_{\{R\}} + S(\left[\rho_{w^*(A)}\right]_{\{R\}}) + \sum_i \ln(\dim\, R_i) \right) \\ &= S(\rho_A, \mathcal{M}_A) + \sum_{\{\alpha\}} p_{\{\alpha\}} \sum_i S\left(\rho_{\text{aux}}(\alpha_i)\right), \end{aligned} \tag{3.290}$$

as expected for agreement with our previous computation of S_A.

The final result (3.290) becomes particularly interesting if we assume that corresponding results for perturbative gravitons are given by a naive extrapolation. Given the identification in that context of the FLM $\frac{\delta \text{Area}}{4G_N}$ term with the second term on the right hand side of (3.290) as described above, such extrapolation suggests that – at least to some order in the bulk Newton constant – the FLM relation may be rewritten as simply

$$S_A = \frac{\text{Area}}{4G} + S(\rho^{ext}_{w(A)}), \tag{3.291}$$

with the first term computed in some classical background and the second defined by an appropriate extended lattice construction for the perturbative metric fluctuations. The form (3.291) is particularly natural given the reliance of FLM on the replica trick and the agreement between the replica trick and extended lattice constructions noted in [257, 217]. It would thus be very interesting to explore such a construction directly in linearized gravity, either on a lattice or in the continuum with an appropriate corresponding extension of the Hilbert space, and to establish any relation to the fully non-linear extended classical phase space for gravity described recently in [313].

Acknowledgements

It is a pleasure to thank Horacio Casini, Zicao Fu, Brianna Grado-White, Dan Harlow, Veronika Hubeny, Aitor Lewkowycz, and Mukund Rangamani for useful conversations. DM was supported in part by the Simons Foundation and by funds from the University of California. BM was supported by NSF Grant PHY13-16748. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research and Innovation.

Bibliography

[1] S. D. Mathur, "The Information paradox: A Pedagogical introduction," Class. Quant. Grav. **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]].

[2] E. J. Martinec and V. Sahakian, "Black holes and five-brane thermodynamics," Phys. Rev. D **60**, 064002 (1999) [hep-th/9901135].

[3] N. Itzhaki, J. M. Maldacena, J. Sonnenschein and S. Yankielowicz, "Supergravity and the large N limit of theories with sixteen supercharges," Phys. Rev. D **58**, 046004 (1998) [hep-th/9802042].

[4] O. Lunin and S. D. Mathur, "AdS / CFT duality and the black hole information paradox," Nucl. Phys. B **623**, 342 (2002) [hep-th/0109154].

[5] S. D. Mathur, "The Fuzzball proposal for black holes: An Elementary review," Fortsch. Phys. **53**, 793 (2005) [hep-th/0502050].

[6] L. Motl, "Proposals on nonperturbative superstring interactions," hep-th/9701025.

[7] T. Banks and N. Seiberg, "Strings from matrices," Nucl. Phys. B **497**, 41 (1997) [hep-th/9702187].

[8] R. Dijkgraaf, E. P. Verlinde and H. L. Verlinde, "Matrix string theory," Nucl. Phys. B **500**, 43 (1997) [hep-th/9703030].

[9] O. Lunin, J. M. Maldacena and L. Maoz, "Gravity solutions for the D1-D5 system with angular momentum," hep-th/0212210.

[10] I. Kanitscheider, K. Skenderis and M. Taylor, "Fuzzballs with internal excitations," JHEP **0706**, 056 (2007) [arXiv:0704.0690 [hep-th]].

[11] V. S. Rychkov, "D1-D5 black hole microstate counting from supergravity," JHEP **0601**, 063 (2006) [hep-th/0512053].

[12] O. Lunin and S. D. Mathur, "Statistical interpretation of Bekenstein entropy for systems with a stretched horizon," Phys. Rev. Lett. **88**, 211303 (2002) [hep-th/0202072].

[13] C. G. Callan, J. M. Maldacena and A. W. Peet, "Extremal black holes as fundamental strings," Nucl. Phys. B **475**, 645 (1996) [hep-th/9510134].

[14] A. Dabholkar, J. P. Gauntlett, J. A. Harvey and D. Waldram, "Strings as solitons and black holes as strings," Nucl. Phys. B **474**, 85 (1996) [hep-th/9511053].

[15] A. Dabholkar, "Exact counting of black hole microstates," Phys. Rev. Lett. **94**, 241301 (2005) [hep-th/0409148].

[16] A. Dabholkar, R. Kallosh and A. Maloney, "A Stringy cloak for a classical singularity," JHEP **0412**, 059 (2004) [hep-th/0410076].

[17] A. Sen, "Black hole solutions in heterotic string theory on a torus," Nucl. Phys. B **440**, 421 (1995) [hep-th/9411187].

[18] A. Sen, "Two Charge System Revisited: Small Black Holes or Horizonless Solutions?," JHEP **1005**, 097 (2010) [arXiv:0908.3402 [hep-th]].

[19] K. Skenderis and M. Taylor, "Fuzzball solutions and D1-D5 microstates," Phys. Rev. Lett. **98**, 071601 (2007) [hep-th/0609154].

[20] I. Kanitscheider, K. Skenderis and M. Taylor, "Holographic anatomy of fuzzballs," JHEP **0704**, 023 (2007) [hep-th/0611171].

[21] M. Baggio, J. de Boer and K. Papadodimas, "A non-renormalization theorem for chiral primary 3-point functions," JHEP **1207**, 137 (2012) [arXiv:1203.1036 [hep-th]].

[22] S. Giusto and R. Russo, "Entanglement Entropy and D1-D5 geometries," arXiv:1405.6185 [hep-th].

[23] V. Balasubramanian, B. Czech, V. E. Hubeny, K. Larjo, M. Rangamani and J. Simon, "Typicality versus thermality: An Analytic distinction," Gen. Rel. Grav. **40**, 1863 (2008) [hep-th/0701122].

[24] J. M. Deutsch, Phys. Rev. A 43, 2046 (1991);

M. Srednicki, Phys. Rev. E 50, 888 (1994) [arXiv:cond-mat/9406056]; J. Phys. A 29, L75 (1996) [arXiv:chao-dyn/9511001]; J. Phys. A 32, 1163 (1999) [arXiv:cond-mat/9809360].

[25] R. Bousso, "Firewalls From Double Purity," Phys. Rev. D **88**, 084035 (2013) [arXiv:1308.2665 [hep-th]].

[26] D. Marolf and J. Polchinski, "Gauge/Gravity Duality and the Black Hole Interior," Phys. Rev. Lett. **111**, 171301 (2013) [arXiv:1307.4706 [hep-th]].

[27] J. M. Maldacena and L. Maoz, "Desingularization by rotation," JHEP **0212**, 055 (2002) [hep-th/0012025].

[28] O. Lunin and S. D. Mathur, "Metric of the multiply wound rotating string," Nucl. Phys. B **610**, 49 (2001) [hep-th/0105136].

[29] N. Iizuka and M. Shigemori, "A Note on D1-D5-J system and 5-D small black ring," JHEP **0508**, 100 (2005) [hep-th/0506215].

[30] V. Balasubramanian, P. Kraus and M. Shigemori, "Massless black holes and black rings as effective geometries of the D1-D5 system," Class. Quant. Grav. **22**, 4803 (2005) [hep-th/0508110].

[31] H. Elvang, R. Emparan, D. Mateos and H. S. Reall, "A Supersymmetric black ring," Phys. Rev. Lett. **93**, 211302 (2004) [hep-th/0407065]; "Supersymmetric black rings and three-charge supertubes," Phys. Rev. D **71**, 024033 (2005) [hep-th/0408120].

[32] I. Bena and N. P. Warner, "One ring to rule them all ... and in the darkness bind them?," Adv. Theor. Math. Phys. **9**, 667 (2005) [hep-th/0408106].

[33] J. P. Gauntlett and J. B. Gutowski, "General concentric black rings," Phys. Rev. D **71**, 045002 (2005) [hep-th/0408122].

[34] A. Dabholkar, N. Iizuka, A. Iqubal, A. Sen and M. Shigemori, "Spinning strings as small black rings," JHEP **0704**, 017 (2007) [hep-th/0611166].

[35] C. V. Johnson, A. W. Peet and J. Polchinski, "Gauge theory and the excision of repulson singularities," Phys. Rev. D **61**, 086001 (2000) [hep-th/9911161].

[36] J. Polchinski and M. J. Strassler, "The String dual of a confining four-dimensional gauge theory," hep-th/0003136.

[37] I. R. Klebanov and A. A. Tseytlin, "Gravity duals of supersymmetric SU(N) x SU(N+M) gauge theories," Nucl. Phys. B **578**, 123 (2000) [hep-th/0002159].

[38] I. R. Klebanov and M. J. Strassler, "Supergravity and a confining gauge theory: Duality cascades and chi SB resolution of naked singularities," JHEP **0008**, 052 (2000) [hep-th/0007191].

[39] A. Sen, "Unification of string dualities," Nucl. Phys. Proc. Suppl. **58**, 5 (1997) [hep-th/9609176].

[40] S. Hawking, *Particle Creation by Black Holes*, Commun.Math.Phys. **43** (1975) 199–220

[41] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black Holes: Complementarity or Firewalls?*, JHEP **1302** (2013) 062, 1207.3123

[42] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford and J. Sully, *An Apologia for Firewalls*, JHEP **1309** (2013) 018, 1304.6483

[43] C. G. Callan and J. M. Maldacena, *D-brane approach to black hole quantum mechanics*, Nuclear Physics B **472** (jul, 1996) 591–608

[44] G. T. Horowitz and J. Polchinski, *A Correspondence principle for black holes and strings*, Phys. Rev. **D55** (1997) 6189–6197, hep-th/9612146

[45] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, Phys.Lett. **B379** (1996) 99–104, hep-th/9601029

[46] I. Bena and N. P. Warner, *Black holes, black rings and their microstates*, Lect. Notes Phys. **755** (2008) 1–92, hep-th/0701216

[47] V. Balasubramanian, J. de Boer, S. El-Showk and I. Messamah, *Black Holes as Effective Geometries*, Class. Quant. Grav. **25** (2008) 214004, 0811.0263

[48] K. Skenderis and M. Taylor, *The fuzzball proposal for black holes*, Phys. Rept. **467** (2008) 117–171, 0804.0552

[49] S. D. Mathur, *Fuzzballs and the information paradox: A Summary and conjectures*, 0810.4525

[50] B. D. Chowdhury and A. Virmani, *Modave Lectures on Fuzzballs and Emission from the D1-D5 System*, 1001.1444

[51] I. Bena and N. P. Warner, *Resolving the Structure of Black Holes: Philosophizing with a Hammer*, 1311.4538

[52] F. C. Eperon, H. S. Reall and J. E. Santos, *Instability of supersymmetric microstate geometries*, 1607.06828

[53] G. Gibbons and N. Warner, *Global structure of five-dimensional fuzzballs*, Class.Quant.Grav. **31** (2014) 025016, 1305.0957

[54] V. Jejjala, O. Madden, S. F. Ross and G. Titchener, *Non-supersymmetric smooth geometries and D1-D5-P bound states*, Phys.Rev. **D71** (2005) 124030, hep-th/0504181

[55] V. Cardoso, O. J. Dias, J. L. Hovdebo and R. C. Myers, *Instability of non-supersymmetric smooth geometries*, Phys.Rev. **D73** (2006) 064031, hep-th/0512277

[56] I. Bena, D. R. Mayerson, A. Puhm and B. Vercnocke, *Tunneling into Microstate Geometries: Quantum Effects Stop Gravitational Collapse*, JHEP **07** (2016) 031, 1512.05376

[57] F. Chen, B. Michel, J. Polchinski and A. Puhm, *Journey to the Center of the Fuzzball*, JHEP **02** (2015) 081, 1408.4798

[58] R. Dijkgraaf, *Instanton strings and hyperKahler geometry*, Nucl. Phys. **B543** (1999) 545–571, hep-th/9810210

[59] N. Seiberg and E. Witten, *The D1 / D5 system and singular CFT*, JHEP **04** (1999) 017, hep-th/9903224

[60] J. Breckenridge, R. C. Myers, A. Peet and C. Vafa, *D-branes and spinning black holes*, Phys.Lett. **B391** (1997) 93–98, hep-th/9602065

[61] O. Lunin and S. D. Mathur, *Rotating deformations of AdS(3) x S**3, the orbifold CFT and strings in the pp wave limit*, Nucl. Phys. **B642** (2002) 91–113, hep-th/0206107

[62] O. Lunin, S. D. Mathur and A. Saxena, *What is the gravity dual of a chiral primary?*, Nucl. Phys. **B655** (2003) 185–217, hep-th/0211292

[63] A. Schwimmer and N. Seiberg, *Comments on the N=2, N=3, N=4 Superconformal Algebras in Two-Dimensions*, Phys. Lett. **B184** (1987) 191–196

[64] P. K. Townsend and M. Zamaklar, *The First law of black brane mechanics*, Class.Quant.Grav. **18** (2001) 5269–5286, hep-th/0107228

[65] P. de Lange, D. R. Mayerson and B. Vercnocke, *Structure of Six-Dimensional Microstate Geometries*, JHEP **09** (2015) 075, 1504.07987

[66] B. C. Palmer and D. Marolf, *Counting supertubes*, JHEP **06** (2004) 028, hep-th/0403025

[67] D. Bak, Y. Hyakutake and N. Ohta, *Phase moduli space of supertubes*, Nucl. Phys. **B696** (2004) 251–262, hep-th/0404104

[68] D. Bak, Y. Hyakutake, S. Kim and N. Ohta, *A geometric look on the microstates of supertubes*, Nucl. Phys. **B712** (2005) 115–138, hep-th/0407253

[69] E. Witten, *A new proof of the positive energy theorem*, Communications in Mathematical Physics **80** (sep, 1981) 381–402

[70] P. C. Aichelburg and R. U. Sexl, *On the gravitational field of a massless particle*, General Relativity and Gravitation **2** (1971), no. 4, 303–312

[71] D. Berenstein, J. Maldacena and H. Nastase, *Strings in flat space and pp waves from Script N = 4 Super Yang Mills*, Journal of High Energy Physics **2002** (apr, 2002) 013–013

[72] V. Balasubramanian, J. de Boer, E. Keski-Vakkuri and S. F. Ross, *Supersymmetric conical defects: Towards a string theoretic description of black hole formation*, Phys. Rev. **D64** (2001) 064011, hep-th/0011217

[73] M. Banados, M. Henneaux, C. Teitelboim and J. Zanelli, *Geometry of the (2+1) black hole*, Phys. Rev. **D48** (1993) 1506–1525, gr-qc/9302012, [Erratum: Phys. Rev.D88,069902(2013)]

[74] M. Banados, C. Teitelboim and J. Zanelli, *The Black hole in three-dimensional space-time*, Phys. Rev. Lett. **69** (1992) 1849–1851, hep-th/9204099

[75] I. Bena, S. Giusto, E. J. Martinec, R. Russo, M. Shigemori, D. Turton and N. P. Warner, *Smooth horizonless geometries deep inside the black-hole regime*, Phys. Rev. Lett. **117** (2016), no. 20, 201601, 1607.03908

[76] O. Lunin, *Adding momentum to D-1 - D-5 system*, JHEP **04** (2004) 054, hep-th/0404006

[77] S. Giusto, S. D. Mathur and A. Saxena, *Dual geometries for a set of 3-charge microstates*, Nucl.Phys. **B701** (2004) 357–379, hep-th/0405017

[78] S. Giusto and S. D. Mathur, *Geometry of D1-D5-P bound states*, Nucl. Phys. **B729** (2005) 203–220, hep-th/0409067

[79] S. Giusto, S. D. Mathur and A. Saxena, *3-charge geometries and their CFT duals*, Nucl. Phys. **B710** (2005) 425–463, hep-th/0406103

[80] A. Strominger, *AdS(2) quantum gravity and string theory*, JHEP **01** (1999) 007, hep-th/9809027

[81] N. Iizuka and J. Polchinski, "A Matrix Model for Black Hole Thermalization," *JHEP* **10** (2008) 028, arXiv:0801.3657 [hep-th].

[82] G. 't Hooft, "A Two-Dimensional Model for Mesons," *Nucl. Phys.* **B75** (1974) 461.

[83] N. Iizuka, T. Okuda, and J. Polchinski, "Matrix Models for the Black Hole Information Paradox," *JHEP* **02** (2010) 073, arXiv:0808.0530 [hep-th].

[84] G. Festuccia and H. Liu, "The Arrow of time, black holes, and quantum mixing of large N Yang-Mills theories," *JHEP* **12** (2007) 027, arXiv:hep-th/0611098 [hep-th].

[85] J. M. Maldacena, "Eternal black holes in anti-de Sitter," *JHEP* **04** (2003) 021, arXiv:hep-th/0106112 [hep-th].

[86] A. Kitaev, "Hidden Correlations in the Hawking Radiation and Thermal Noise," *talk given at Fundamental Physics Prize Symposium, Nov. 10, 2014* .

[87] A. Kitaev, "A simple model of quantum holography," *KITP strings seminar and Entanglement 2015 program* (Feb. 12, April 7, and May 27, 2015) . http://online.kitp.ucsb.edu/online/entangled15/.

[88] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," arXiv:1503.01409 [hep-th].

[89] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," *JHEP* **03** (2014) 067, arXiv:1306.0622 [hep-th].

[90] S. H. Shenker and D. Stanford, "Multiple Shocks," *JHEP* **12** (2014) 046, arXiv:1312.3296 [hep-th].

[91] S. Leichenauer, "Disrupting Entanglement of Black Holes," *Phys. Rev.* **D90** (2014) no. 4, 046009, arXiv:1405.7365 [hep-th].

[92] D. A. Roberts, D. Stanford, and L. Susskind, "Localized shocks," *JHEP* **03** (2015) 051, arXiv:1409.8180 [hep-th].

[93] D. A. Roberts and D. Stanford, "Two-dimensional conformal field theory and the butterfly effect," *Phys. Rev. Lett.* **115** (2015) no. 13, 131603, 1412.5123 [hep-th].

[94] S. H. Shenker and D. Stanford, "Stringy effects in scrambling," *JHEP* **05** (2015) 132, arXiv:1412.6087 [hep-th].

[95] J. Polchinski, "Chaos in the black hole S-matrix," arXiv:1505.08108 [hep-th].

[96] D. Berenstein and A. M. Garcia-Garcia, "Universal quantum constraints on the butterfly effect," arXiv:1510.08870 [hep-th].

[97] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels," arXiv:1511.04021 [hep-th].

[98] G. Gur-Ari, M. Hanada, and S. H. Shenker, "Chaos in Classical D0-Brane Mechanics," arXiv:1512.00019 [hep-th].

[99] D. Stanford, "Many-body chaos at weak coupling," arXiv:1512.07687 [hep-th].

[100] E. Berkowitz, M. Hanada, and J. Maltz, "Chaos in Matrix Models and Black Hole Evaporation," arXiv:1602.01473 [hep-th].

[101] J. Polchinski and V. Rosenhaus, "The Spectrum in the Sachdev-Ye-Kitaev Model," arXiv:1601.06768 [hep-th].

[102] A. L. Fitzpatrick and J. Kaplan, "A Quantum Correction To Chaos," arXiv:1601.06164 [hep-th].

[103] A. Larkin and Y. N. Ovchinnikov, "Quasiclassical method in the theory of superconductivity," *JETP* **28** (1969) no. 6, 1200.

[104] S. Sachdev and J. Ye, "Gapless spin-fluid ground state in a random quantum Heisenberg magnet," *Phys. Rev. Lett.* **70** (May, 1993) 3339–3342. http://link.aps.org/doi/10.1103/PhysRevLett.70.3339.

[105] G. T. Horowitz and V. E. Hubeny, "Quasinormal modes of AdS black holes and the approach to thermal equilibrium," *Phys. Rev.* **D62** (2000) 024027, arXiv:hep-th/9909056 [hep-th].

[106] U. H. Danielsson, E. Keski-Vakkuri, and M. Kruczenski, "Black hole formation in AdS and thermalization on the boundary," *JHEP* **02** (2000) 039, arXiv:hep-th/9912209 [hep-th].

[107] M.-J. Giannoni, A. Voros, and J. Zinn-Justin, *Les Houches: Chaos and Quantum Physics*. Elsevier, 1991.

[108] V. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*. Addison-Wesley, 1968.

[109] O. Parcollet and A. Georges, "Non-Fermi-liquid regime of a doped Mott insulator," *Phys. Rev. B* **59** (Feb., 1999) 5341–5360, cond-mat/9806119.

[110] S. Sachdev, "Bekenstein-Hawking Entropy and Strange Metals," *Phys. Rev.* **X5** (2015) no. 4, 041025, arXiv:1506.05111 [hep-th].

[111] J. Maldacena and D. Stanford, *to appear*.

[112] A. Kitaev, *to appear*.

[113] C. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. Springer, 1999.

[114] S. Weinberg, "Phenomenological Lagrangians," Physica A **96**, 327 (1979).

[115] W. D. Goldberger and M. B. Wise, "Renormalization group flows for brane couplings," Phys. Rev. D **65**, 025011 (2002) [hep-th/0104170].

[116] T. Damour, Nuovo Cim. B26, 157 (1975);

B. Allen, B. S. Kay and A. C. Ottewill, "Long range effects of cosmic string structure," Phys. Rev. D **53**, 6829 (1996) [gr-qc/9510058];

E. A. Mirabelli and M. E. Peskin, "Transmission of supersymmetry breaking from a four-dimensional boundary," Phys. Rev. D **58**, 065002 (1998) [hep-th/9712214];

S. N. Solodukhin, "Exact solution for a quantum field with delta - like interaction," Nucl. Phys. B **541**, 461 (1999) [hep-th/9801054];

H. Georgi, A. K. Grant and G. Hailu, "Brane couplings from bulk loops," Phys. Lett. B **506**, 207 (2001) [hep-ph/0012379].

[117] U. Danielsson, S. Massai and T. Van Riet, "SUSY-breaking in warped throats," to appear.

[118] O. DeWolfe, S. Kachru and H. L. Verlinde, "The Giant inflaton," JHEP **0405**, 017 (2004) [hep-th/0403123].

[119] J. Blaback, U. H. Danielsson and T. Van Riet, "Resolving antibrane singularities through time-dependence," JHEP **1302**, 061 (2013) [arXiv:1202.1132 [hep-th]].

[120] I. Bena, A. Buchel and O. J. C. Dias, "Horizons cannot save the Landscape," Phys. Rev. D **87**, no. 6, 063012 (2013) [arXiv:1212.5162 [hep-th]].

[121] I. Bena, J. Blaback, U. H. Danielsson and T. Van Riet, "Antibranes cannot become black," Phys. Rev. D **87**, no. 10, 104023 (2013) [arXiv:1301.7071 [hep-th]].

[122] J. Blabeck, U. H. Danielsson, D. Junghans, T. Van Riet and S. C. Vargas, "Localised antibranes in non-compact throats at zero and finite T," arXiv:1409.0534 [hep-th].

[123] S. Kachru, J. Pearson and H. L. Verlinde, "Brane / flux annihilation and the string dual of a nonsupersymmetric field theory," JHEP **0206**, 021 (2002) [hep-th/0112197].

[124] E. Mintun, J. Polchinski and S. Sun, "The Field Theory of Intersecting D3-branes," arXiv:1402.6327 [hep-th].

[125] R. Jackiw, "Delta function potentials in two-dimensional and three-dimensional quantum mechanics," In *Jackiw, R.: Diverse topics in theoretical and mathematical physics* 35-53, Singapore: World Scientific (1995).

[126] A. V. Manohar, "Effective field theories," Lect. Notes Phys. **479** , 311 (1997) [hep-ph/9606222];

H. Georgi, "Effective field theory," Ann. Rev. Nucl. Part. Sci. **43**, 209 (1993).

[127] S. S. Gubser, A. Hashimoto, I. R. Klebanov and J. M. Maldacena, "Gravitational lensing by p-branes," Nucl. Phys. B **472**, 231 (1996) [hep-th/9601057];

M. R. Garousi and R. C. Myers, "Superstring scattering from D-branes," Nucl. Phys. B **475**, 193 (1996) [hep-th/9603194];

M. R. Garousi and R. C. Myers, "World volume interactions on D-branes," Nucl. Phys. B **542**, 73 (1999) [hep-th/9809100];

C. P. Bachas, P. Bain and M. B. Green, "Curvature terms in D-brane actions and their M theory origin," JHEP **9905**, 011 (1999) [hep-th/9903210];

R. C. Myers, "Dielectric branes," JHEP **9912**, 022 (1999) [hep-th/9910053];

A. Fotopoulos, "On (alpha-prime)**2 corrections to the D-brane action for nongeodesic world volume embeddings," JHEP **0109**, 005 (2001) [hep-th/0104146];

P. Koerber, "Abelian and non-Abelian D-brane effective actions," Fortsch. Phys. **52**, 871 (2004) [hep-th/0405227].

M. R. Garousi and E. Hatefi, "On Wess-Zumino terms of Brane-Antibrane systems," Nucl. Phys. B **800**, 502 (2008) [arXiv:0710.5875 [hep-th]].

K. Becker, G. Guo and D. Robbins, "Four-Derivative Brane Couplings from String Amplitudes," JHEP **1112**, 050 (2011) [arXiv:1110.3831 [hep-th]];

E. Hatefi, "On D-brane anti D-brane effective actions and their corrections to all orders in alpha-prime," JCAP **1309**, 011 (2013) [arXiv:1211.5538, arXiv:1211.5538 [hep-th]].

[128] S. S. Gubser, "Curvature singularities: The Good, the bad, and the naked," Adv. Theor. Math. Phys. **4**, 679 (2000) [hep-th/0002160].

[129] T. Banks, W. Fischler, S. H. Shenker and L. Susskind, "M theory as a matrix model: A Conjecture," Phys. Rev. D **55**, 5112 (1997) [hep-th/9610043].

[130] E. Silverstein, "(A)dS backgrounds from asymmetric orientifolds," hep-th/0106209.

[131] S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, "De Sitter vacua in string theory," Phys. Rev. D **68**, 046005 (2003) [hep-th/0301240].

[132] R. Kallosh and T. Wrase, "Emergence of Spontaneously Broken Supersymmetry on an Anti-D3-Brane in KKLT dS Vacua," JHEP **1412**, 117 (2014) [arXiv:1411.1121 [hep-th]].

[133] P. McGuirk, G. Shiu and Y. Sumitomo, "Non-supersymmetric infrared perturbations to the warped deformed conifold," Nucl. Phys. B **842**, 383 (2011) [arXiv:0910.4581 [hep-th]].

[134] I. Bena, M. Graña and N. Halmagyi, "On the Existence of Meta-stable Vacua in Klebanov-Strassler," JHEP **1009**, 087 (2010) [arXiv:0912.3519 [hep-th]].

[135] A. Dymarsky and S. Massai, "Uplifting the baryonic branch: a test for backreacting anti-D3-branes," JHEP **1411**, 034 (2014) [arXiv:1310.0015 [hep-th]].

[136] U. H. Danielsson and T. Van Riet, "Fatal attraction: more on decaying antibranes," arXiv:1410.8476 [hep-th].

[137] J. Blaback, U. H. Danielsson, D. Junghans, T. Van Riet, T. Wrase and M. Zagermann, "The problematic backreaction of SUSY-breaking branes," JHEP **1108**, 105 (2011) [arXiv:1105.4879 [hep-th]];

J. Blaback, U. H. Danielsson, D. Junghans, T. Van Riet, T. Wrase and M. Zagermann, "(Anti-)Brane backreaction beyond perturbation theory," JHEP **1202**, 025 (2012) [arXiv:1111.2605 [hep-th]].

I. Bena, D. Junghans, S. Kuperstein, T. Van Riet, T. Wrase and M. Zagermann, "Persistent antibrane singularities," JHEP **1210**, 078 (2012) [arXiv:1205.1798 [hep-th]].

[138] J. Polchinski, "Open heterotic strings," JHEP **0609**, 082 (2006) [hep-th/0510033].

[139] D. Marolf, "Chern-Simons terms and the three notions of charge," hep-th/0006117.

[140] M. J. Bowick, S. B. Giddings, J. A. Harvey, G. T. Horowitz and A. Strominger, "Axionic Black Holes and a Bohm-Aharonov Effect for Strings," Phys. Rev. Lett. **61**, 2823 (1988).

[141] R. I. Nepomechie, "Magnetic Monopoles from Antisymmetric Tensor Gauge Fields," Phys. Rev. D **31**, 1921 (1985);

C. Teitelboim, "Monopoles of Higher Rank," Phys. Lett. B **167**, 69 (1986).

[142] G. S. Hartnett, "Localised Anti-Branes in Flux Backgrounds," arXiv:1501.06568 [hep-th].

[143] U. H. Danielsson, "Perturbative decay of anti-branes in flux backgrounds due to space time instabilities," arXiv:1502.01234 [hep-th].

[144] I. Bena, M. Graña, S. Kuperstein and S. Massai, "Giant Tachyons in the Landscape," arXiv:1410.7776 [hep-th].

[145] N. P. Warner, "Renormalization group flows from five-dimensional supergravity," Class. Quant. Grav. **17**, 1287 (2000) [hep-th/9911240].

[146] I. Bena, A. Puhm and B. Vercnocke, "Non-extremal Black Hole Microstates: Fuzzballs of Fire or Fuzzballs of Fuzz ?," JHEP **1212**, 014 (2012) [arXiv:1208.3468 [hep-th]];

I. Bena, A. Puhm and B. Vercnocke, "Metastable Supertubes and non-extremal Black Hole Microstates," JHEP **1204**, 100 (2012) [arXiv:1109.5180 [hep-th]].

[147] A. L. Fitzpatrick, J. Kaplan and M. T. Walters, *Virasoro Conformal Blocks and Thermality from Classical Background Fields*, *JHEP* **11** (2015) 200, [1501.05315].

[148] A. L. Fitzpatrick, J. Kaplan, M. T. Walters and J. Wang, *Hawking from Catalan*, *JHEP* **05** (2016) 069, [1510.00014].

[149] A. L. Fitzpatrick and J. Kaplan, *Conformal Blocks Beyond the Semi-Classical Limit*, *JHEP* **05** (2016) 075, [1512.03052].

[150] A. L. Fitzpatrick and J. Kaplan, *A Quantum Correction To Chaos*, *JHEP* **05** (2016) 070, [1601.06164].

[151] A. L. Fitzpatrick, J. Kaplan, D. Li and J. Wang, *On information loss in AdS_3/CFT_2*, *JHEP* **05** (2016) 109, [1603.08925].

[152] H. Chen, A. L. Fitzpatrick, J. Kaplan, D. Li and J. Wang, *Degenerate Operators and the $1/c$ Expansion: Lorentzian Resummations, High Order Computations, and Super-Virasoro Blocks*, 1606.02659.

[153] A. Liam Fitzpatrick and J. Kaplan, *On the Late-Time Behavior of Virasoro Blocks and a Classification of Semiclassical Saddles*, 1609.07153.

[154] D. A. Roberts and D. Stanford, *Two-dimensional conformal field theory and the butterfly effect*, *Phys. Rev. Lett.* **115** (2015) 131603, [1412.5123].

[155] T. Hartman, *Entanglement Entropy at Large Central Charge*, 1303.6955.

[156] T. Hartman, C. A. Keller and B. Stoica, *Universal Spectrum of 2d Conformal Field Theory in the Large c Limit*, *JHEP* **09** (2014) 118, [1405.5137].

[157] C. T. Asplund, A. Bernamonti, F. Galli and T. Hartman, *Holographic Entanglement Entropy from 2d CFT: Heavy States and Local Quenches*, *JHEP* **02** (2015) 171, [1410.1392].

[158] C. T. Asplund, A. Bernamonti, F. Galli and T. Hartman, *Entanglement Scrambling in 2d Conformal Field Theory*, *JHEP* **09** (2015) 110, [1506.03772].

[159] E. Perlmutter, *Bounding the Space of Holographic CFTs with Chaos*, *JHEP* **10** (2016) 069, [1602.08272].

[160] T. Anous, T. Hartman, A. Rovai and J. Sonner, *Black Hole Collapse in the 1/c Expansion*, *JHEP* **07** (2016) 123, [1603.04856].

[161] P. Kovtun and A. Ritz, *Black holes and universality classes of critical points*, *Phys. Rev. Lett.* **100** (2008) 171606, [0801.2785].

[162] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, *Holography from Conformal Field Theory*, *JHEP* **10** (2009) 079, [0907.0151].

[163] S. El-Showk and K. Papadodimas, *Emergent Spacetime and Holographic CFTs*, *JHEP* **10** (2012) 106, [1101.4163].

[164] Z. Komargodski and A. Zhiboedov, *Convexity and Liberation at Large Spin*, *JHEP* **11** (2013) 140, [1212.4103].

[165] A. L. Fitzpatrick, J. Kaplan, D. Poland and D. Simmons-Duffin, *The Analytic Bootstrap and AdS Superhorizon Locality*, *JHEP* **12** (2013) 004, [1212.3616].

[166] A. L. Fitzpatrick and J. Kaplan, *AdS Field Theory from Conformal Field Theory*, *JHEP* **02** (2013) 054, [1208.0337].

[167] A. L. Fitzpatrick, J. Kaplan and M. T. Walters, *Universality of Long-Distance AdS Physics from the CFT Bootstrap*, *JHEP* **08** (2014) 145, [1403.6829].

[168] A. L. Fitzpatrick, J. Kaplan, M. T. Walters and J. Wang, *Eikonalization of Conformal Blocks*, *JHEP* **09** (2015) 019, [1504.01737].

[169] X. O. Camanho, J. D. Edelstein, J. Maldacena and A. Zhiboedov, *Causality Constraints on Corrections to the Graviton Three-Point Coupling*, *JHEP* **02** (2016) 020, [1407.5597].

[170] T. Hartman, S. Jain and S. Kundu, *Causality Constraints in Conformal Field Theory*, *JHEP* **05** (2016) 099, [1509.00014].

[171] J. Maldacena, D. Simmons-Duffin and A. Zhiboedov, *Looking for a bulk point*, 1509.03612.

[172] L. F. Alday, A. Bissi and T. Lukowski, *Lessons from crossing symmetry at large N*, *JHEP* **06** (2015) 074, [1410.4717].

[173] N. Benjamin, M. C. N. Cheng, S. Kachru, G. W. Moore and N. M. Paquette, *Elliptic Genera and 3d Gravity*, *Annales Henri Poincare* **17** (2016) 2623–2662, [1503.04800].

[174] J. L. Cardy, *Operator Content of Two-Dimensional Conformally Invariant Theories*, *Nucl. Phys.* **B270** (1986) 186–204.

[175] E. Shaghoulian, *Modular forms and a generalized Cardy formula in higher dimensions*, *Phys. Rev.* **D93** (2016) 126005, [1508.02728].

[176] E. Shaghoulian, *Black hole microstates in AdS*, 1512.06855.

[177] L. J. Dixon, D. Friedan, E. J. Martinec and S. H. Shenker, *The Conformal Field Theory of Orbifolds*, *Nucl. Phys.* **B282** (1987) 13–73.

[178] R. Dijkgraaf, G. W. Moore, E. P. Verlinde and H. L. Verlinde, *Elliptic genera of symmetric products and second quantized strings*, *Commun. Math. Phys.* **185** (1997) 197–209, [hep-th/9608096].

[179] P. Bantay, *Characters and modular properties of permutation orbifolds*, *Phys. Lett.* **B419** (1998) 175–178, [hep-th/9708120].

[180] P. Bantay, *Symmetric products, permutation orbifolds and discrete torsion*, *Lett. Math. Phys.* **63** (2003) 209–218, [hep-th/0004025].

[181] O. Lunin and S. D. Mathur, *Correlation functions for M^N/S_N orbifolds*, *Commun. Math. Phys.* **219** (2001) 399–442, [hep-th/0006196].

[182] A. Pakman, L. Rastelli and S. S. Razamat, *Diagrams for Symmetric Product Orbifolds*, *JHEP* **10** (2009) 034, [0905.3448].

[183] C. A. Keller, *Phase transitions in symmetric orbifold CFTs and universality*, *JHEP* **03** (2011) 114, [1101.4937].

[184] J. de Boer, *Six-dimensional supergravity on $S^3 \times AdS_3$ and 2-D conformal field theory*, *Nucl. Phys.* **B548** (1999) 139–166, [hep-th/9806104].

[185] S. M. Trott, *A pair of generators for the unimodular group*, *Canad. Math. Bull.* **5** (1962) .

[186] R. C. Myers, *Stress tensors and Casimir energies in the AdS / CFT correspondence*, *Phys.Rev.* **D60** (1999) 046002, [hep-th/9903203].

[187] G. T. Horowitz and R. C. Myers, *The AdS / CFT correspondence and a new positive energy conjecture for general relativity*, *Phys. Rev.* **D59** (1998) 026005, [hep-th/9808079].

[188] A. Belin, C. A. Keller and A. Maloney, *Permutation Orbifolds in the large N Limit*, 1509.01256.

[189] F. M. Haehl and M. Rangamani, *Permutation orbifolds and holography*, *JHEP* **03** (2015) 163, [1412.2759].

[190] A. Belin, C. A. Keller and A. Maloney, *String Universality for Permutation Orbifolds*, *Phys. Rev.* **D91** (2015) 106005, [1412.7159].

[191] N. Benjamin, S. Kachru, C. A. Keller and N. M. Paquette, *Emergent space-time and the supersymmetric index*, *JHEP* **05** (2016) 158, [1512.00010].

[192] P. H. Ginsparg, *APPLIED CONFORMAL FIELD THEORY*, in *Les Houches Summer School in Theoretical Physics: Fields, Strings, Critical Phenomena Les Houches, France, June 28-August 5, 1988*, pp. 1–168, 1988. hep-th/9108028.

[193] P. J. Cameron, *Oligomorphic permutation groups*, vol. 152 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1990, 10.1017/CBO9780511549809.

[194] P. J. Cameron, *Oligomorphic permutation groups*, in *Perspectives in mathematical sciences. II*, vol. 8 of *Stat. Sci. Interdiscip. Res.*, pp. 37–61. World Sci. Publ., Hackensack, NJ, 2009. DOI.

[195] P. J. Cameron, *Transitivity of permutation groups on unordered sets*, *Math. Z.* **148** (1976) 127–139.

[196] D. Goldfeld, *Automorphic Forms and L-Functions for the Group GL(n,R)*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[197] M. J. Park, C. Fang, B. A. Bernevig and M. J. Gilbert, *Modular Anomalies in (2+1) and (3+1)-D Edge Theories*, 1604.00407.

[198] J. L. Cardy, *Operator content and modular properties of higher dimensional conformal field theories*, *Nucl. Phys.* **B366** (1991) 403–419.

[199] S. Hellerman, *A Universal Inequality for CFT and Quantum Gravity*, *JHEP* **08** (2011) 130, [0902.2790].

[200] S. Hellerman and C. Schmidt-Colinet, *Bounds for State Degeneracies in 2D Conformal Field Theory*, *JHEP* **08** (2011) 127, [1007.0756].

[201] D. Friedan and C. A. Keller, *Constraints on 2d CFT partition functions*, *JHEP* **10** (2013) 180, [1307.6562].

[202] J. D. Qualls and A. D. Shapere, *Bounds on Operator Dimensions in 2D Conformal Field Theories*, *JHEP* **05** (2014) 091, [1312.0038].

[203] J. D. Qualls, *Universal Bounds in Even-Spin CFTs*, *JHEP* **12** (2015) 001, [1412.0383].

[204] N. Benjamin, E. Dyer, A. L. Fitzpatrick and S. Kachru, *Universal Bounds on Charged States in 2d CFT and 3d Gravity*, *JHEP* **08** (2016) 041, [1603.09745].

[205] S. Collier, Y.-H. Lin and X. Yin, *Modular Bootstrap Revisited*, 1608.06241.

[206] L.-Y. Hung, R. C. Myers, M. Smolkin and A. Yale, *Holographic Calculations of Renyi Entropy*, *JHEP* **1112** (2011) 047, [1110.1084].

[207] L.-Y. Hung, R. C. Myers and M. Smolkin, *Twist operators in higher dimensions*, *JHEP* **1410** (2014) 178, [1407.6429].

[208] P. Calabrese and J. L. Cardy, *Entanglement entropy and quantum field theory*, *J. Stat. Mech.* **0406** (2004) P06002, [hep-th/0405152].

[209] H. Casini and M. Huerta, *Entanglement entropy in free quantum field theory*, *J. Phys.* **A42** (2009) 504007, [arXiv:0905.2562].

[210] J. Eisert, M. Cramer, and M. B. Plenio, *Area laws for the entanglement entropy - a review*, *Rev. Mod. Phys.* **82** (2010) 277–306, [arXiv:0808.3773].

[211] H. Casini, *Relative entropy and the Bekenstein bound*, *Class. Quant. Grav.* **25** (2008) 205021, [arXiv:0804.2182].

[212] H. Casini and M. Huerta, *Entanglement entropy for the n-sphere*, *Phys. Lett.* **B694** (2011) 167–171, [arXiv:1007.1813].

[213] H. Casini, M. Huerta, and R. C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **05** (2011) 036, [arXiv:1102.0440].

[214] I. R. Klebanov, S. S. Pufu, S. Sachdev, and B. R. Safdi, *Renyi Entropies for Free Field Theories*, *JHEP* **04** (2012) 074, [arXiv:1111.6290].

[215] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *J. Phys.* **A42** (2009) 504005, [arXiv:0905.4013].

[216] H. Casini, M. Huerta, and J. A. Rosabal, *Remarks on entanglement entropy for gauge fields*, *Phys. Rev.* **D89** (2014), no. 8 085012, [arXiv:1312.1183].

[217] W. Donnelly and A. C. Wall, *Geometric entropy and edge modes of the electromagnetic field*, *Phys. Rev.* **D94** (2016), no. 10 104053, [arXiv:1506.0579].

[218] W. Donnelly, *Entanglement entropy and nonabelian gauge symmetry*, *Class. Quant. Grav.* **31** (2014), no. 21 214003, [arXiv:1406.7304].

[219] C. P. Herzog and T. Nishioka, *The Edge of Entanglement: Getting the Boundary Right for Non-Minimally Coupled Scalar Fields*, arXiv:1610.0226.

[220] K. Ohmori and Y. Tachikawa, *Physics at the entangling surface*, *J. Stat. Mech.* **1504** (2015) P04010, [arXiv:1406.4167].

[221] J. Cardy and E. Tonni, *Entanglement hamiltonians in two-dimensional conformal field theory*, arXiv:1608.0128.

[222] W. Donnelly, B. Michel, and A. Wall, *Electromagnetic Duality and Entanglement Anomalies*, *Phys. Rev.* **D96** (2017), no. 4 045008, [arXiv:1611.0592].

[223] C. A. Agon, M. Headrick, D. L. Jafferis, and S. Kasko, *Disk entanglement entropy for a Maxwell field*, *Phys. Rev.* **D89** (2014), no. 2 025018, [arXiv:1310.4886].

[224] D. Radicevic, *Entanglement in Weakly Coupled Lattice Gauge Theories*, *JHEP* **04** (2016) 163, [arXiv:1509.0847].

[225] K. Sakai and Y. Satoh, *Entanglement through conformal interfaces*, *JHEP* **12** (2008) 001, [arXiv:0809.4548].

[226] E. M. Brehm and I. Brunner, *Entanglement entropy through conformal interfaces in the 2D Ising model*, *JHEP* **09** (2015) 080, [arXiv:1505.0264].

[227] E. M. Brehm, I. Brunner, D. Jaud, and C. Schmidt-Colinet, *Entanglement and topological interfaces*, *Fortsch. Phys.* **64** (2016), no. 6-7 516–535, [arXiv:1512.0594].

[228] M. A. Metlitski and T. Grover, *Entanglement Entropy of Systems with Spontaneously Broken Continuous Symmetry*, arXiv:1112.5166.

[229] D. Radicevic, *Entanglement Entropy and Duality*, *JHEP* **11** (2016) 130, [arXiv:1605.0939].

[230] W.-Z. Guo and S. He, *Rnyi entropy of locally excited states with thermal and boundary effect in 2D CFTs*, *JHEP* **04** (2015) 099, [arXiv:1501.0075].

[231] R. Haag, *Local Quantum Physics*. Springer Nature, 1996.

[232] C. Holzhey, F. Larsen, and F. Wilczek, *Geometric and renormalized entropy in conformal field theory*, *Nucl. Phys.* **B424** (1994) 443–467, [hep-th/9403108].

[233] I. Affleck and A. W. W. Ludwig, *Universal noninteger 'ground state degeneracy' in critical quantum systems*, *Phys. Rev. Lett.* **67** (1991) 161–164.

[234] J. L. Cardy, *Boundary conformal field theory*, hep-th/0411189.

[235] S. Eggert and I. Affleck, *Magnetic impurities in half integer spin Heisenberg antiferromagnetic chains*, *Phys. Rev.* **B46** (1992) 10866–10883.

[236] D. V. Vassilevich, *Heat kernel expansion: User's manual*, *Phys. Rept.* **388** (2003) 279–360, [hep-th/0306138].

[237] M. A. Metlitski, *S-duality of $u(1)$ gauge theory with $\theta = \pi$ on non-orientable manifolds: Applications to topological insulators and superconductors*, arXiv:1510.0566.

[238] J. Polchinski, *Dirichlet Branes and Ramond-Ramond charges*, *Phys. Rev. Lett.* **75** (1995) 4724–4727, [hep-th/9510017].

[239] M. J. Duff and P. van Nieuwenhuizen, *Quantum Inequivalence of Different Field Representations*, *Phys. Lett.* **B94** (1980) 179–182.

[240] W. Siegel, *Quantum Equivalence of Different Field Representations*, *Phys. Lett.* **B103** (1981) 107–109.

[241] E. S. Fradkin and A. A. Tseytlin, *Quantum Equivalence of Dual Field Theories*, *Annals Phys.* **162** (1985) 31.

[242] M. T. Grisaru, N. K. Nielsen, W. Siegel, and D. Zanon, *Energy Momentum Tensors, Supercurrents, (Super)traces and Quantum Equivalence*, *Nucl. Phys.* **B247** (1984) 157.

[243] A. Schwarz and Y. Tyupkin, *Quantization of antisymmetric tensors and ray-singer torsion*, *Nuclear Physics B* **242** (1984), no. 2 436 – 446.

[244] E. Witten, *On S duality in Abelian gauge theory*, *Selecta Math.* **1** (1995) 383, [hep-th/9505186].

[245] F. Larsen and P. Lisbao, *Divergences and boundary modes in $\mathcal{N} = 8$ supergravity*, *JHEP* **01** (2016) 024, [arXiv:1508.0341].

[246] B. McLellan, *Localization in Abelian Chern-Simons Theory*, *J. Math. Phys.* **54** (2013) 023507, [arXiv:1208.1724].

[247] I. L. Buchbinder, E. N. Kirillova, and N. G. Pletnev, *Quantum Equivalence of Massive Antisymmetric Tensor Field Models in Curved Space*, *Phys. Rev.* **D78** (2008) 084024, [arXiv:0806.3505].

[248] J. Cheeger, *Analytic torsion and the heat equation*, *Annals of Mathematics* **109** (1979), no. 2 259–321.

[249] D. B. Ray and I. M. Singer, *Analytic torsion*, in *Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971)*, pp. 167–181. Amer. Math. Soc., Providence, R.I., 1973.

[250] K. Reidemeister, *Homotopieringe und linsenräume*, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* **11** (dec, 1935) 102–109.

[251] A. C. Wall, *Testing the Generalized Second Law in 1+1 dimensional Conformal Vacua: An Argument for the Causal Horizon*, *Phys. Rev.* **D85** (2012) 024015, [arXiv:1105.3520].

[252] A. Castro, S. Detournay, N. Iqbal, and E. Perlmutter, *Holographic entanglement entropy and gravitational anomalies*, *JHEP* **07** (2014) 114, [arXiv:1405.2792].

[253] T. Nishioka and A. Yarom, *Anomalies and Entanglement Entropy*, *JHEP* **03** (2016) 077, [arXiv:1509.0428].

[254] N. Iqbal and A. C. Wall, *Anomalies of the Entanglement Entropy in Chiral Theories*, *JHEP* **10** (2016) 111, [arXiv:1509.0432].

[255] T. Azeyanagi, R. Loganayagam, and G. S. Ng, *Holographic Entanglement for Chern-Simons Terms*, *JHEP* **02** (2017) 001, [arXiv:1507.0229].

[256] T. L. Hughes, R. G. Leigh, O. Parrikar, and S. T. Ramamurthy, *Entanglement entropy and anomaly inflow*, *Phys. Rev.* **D93** (2016), no. 6 065059, [arXiv:1509.0496].

[257] W. Donnelly and A. C. Wall, *Entanglement entropy of electromagnetic edge modes*, *Phys. Rev. Lett.* **114** (2015), no. 11 111603, [arXiv:1412.1895].

[258] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath, *Quasi-particle Statistics and Braiding from Ground State Entanglement*, *Phys. Rev.* **B85** (2012) 235151, [arXiv:1111.2342].

[259] I. H. Kim and B. J. Brown, *Ground-state entanglement constrains low-energy excitations*, *Phys. Rev. B* **92** (Sep, 2015) 115–139.

[260] X. Huang and C.-T. Ma, *Analysis of the Entanglement with Centers*, arXiv:1607.0675.

[261] C. Delcamp, B. Dittrich, and A. Riello, *On entanglement entropy in non-Abelian lattice gauge theory and 3D quantum gravity*, *JHEP* **11** (2016) 102, [arXiv:1609.0480].

[262] H. Raj, *A note on sphere free energy of p-form gauge theory and Hodge duality*, arXiv:1611.0250.

[263] W. Donnelly and A. C. Wall, *Unitarity of Maxwell theory on curved spacetimes in the covariant formalism*, *Phys. Rev.* **D87** (2013), no. 12 125033, [arXiv:1303.1885].

[264] P. Townsend, *Covariant quantization of antisymmetric tensor gauge fields*, *Physics Letters B* **88** (1979), no. 1 97 – 101.

[265] W. Siegel, *Hidden Ghosts*, *Phys. Lett.* **B93** (1980) 170–172.

[266] D. Ray and I. Singer, *R-torsion and the laplacian on riemannian manifolds*, *Advances in Mathematics* **7** (1971), no. 2 145 – 210.

[267] V. G. Turaev, *Reidemeister torsion in knot theory*, *Russian Mathematical Surveys* **41** (feb, 1986) 119–182.

[268] J. Cheeger, *Analytic torsion and reidemeister torsion*, *Proceedings of the National Academy of Sciences of the United States of America* **74** (1977), no. 7 2651–2654.

[269] W. Müller, *Analytic torsion and r-torsion of riemannian manifolds*, *Advances in Mathematics* **28** (1978), no. 3 233 – 305.

[270] H. McKean and I. M. Singer, *Curvature and the eigenvalues of the laplacian*, *J. Differential Geom.* **1** (1967), no. 1-2 43 – 69.

[271] W. Donnelly and A. C. Wall, *Do gauge fields really contribute negatively to black hole entropy?*, *Phys. Rev.* **D86** (2012) 064042, [arXiv:1206.5831].

[272] D. V. Fursaev and S. N. Solodukhin, *On the description of the Riemannian geometry in the presence of conical defects*, *Phys. Rev.* **D52** (1995) 2133–2143, [hep-th/9501127].

[273] C. T. C. Wall, *Non-additivity of the signature*, *Inventiones mathematicae* **7** (1969), no. 3 269–274.

[274] W. Donnelly, *Decomposition of entanglement entropy in lattice gauge theory*, *Phys. Rev.* **D85** (2012) 085004, [arXiv:1109.0036].

[275] S. N. Solodukhin, *Entanglement entropy, conformal invariance and extrinsic geometry*, *Phys. Lett.* **B665** (2008) 305–309, [arXiv:0802.3117].

[276] K.-W. Huang, *Central Charge and Entangled Gauge Fields*, *Phys. Rev.* **D92** (2015), no. 2 025010, [arXiv:1412.2730].

[277] S. Rosenberg, *The Laplacian on a Riemannian Manifold: An Introduction to Analysis on Manifolds*. London Mathematical Society Student Texts. Cambridge University Press, 1997.

[278] A. Hatcher, *Algebraic Topology*. Cambridge University Press, 1 ed., 2001.

[279] C. Montonen and D. I. Olive, *Magnetic Monopoles as Gauge Particles?*, *Phys. Lett.* **B72** (1977) 117–120.

[280] M. Alvarez and D. I. Olive, *Spin and abelian electromagnetic duality on four-manifolds*, *Communications in Mathematical Physics* **217** (mar, 2001) 331–356.

[281] E. Witten, *Dyons of charge $e\theta/2\pi$*, *Physics Letters B* **86** (1979), no. 3 283 – 287.

[282] W. G. Unruh, *Notes on black hole evaporation*, *Phys. Rev.* **D14** (1976) 870.

[283] S. A. Fulling, *Nonuniqueness of canonical field quantization in Riemannian space-time*, *Phys. Rev.* **D7** (1973) 2850–2862.

[284] W. Israel, *Thermo-field dynamics of black holes*, *Phys. Lett.* **A57** (1976) 107–110.

[285] J. B. Hartle and S. W. Hawking, *Path Integral Derivation of Black Hole Radiance*, *Phys. Rev.* **D13** (1976) 2188–2203.

[286] D. G. Boulware, *Quantum Field Theory in Schwarzschild and Rindler Spaces*, *Phys. Rev.* **D11** (1975) 1404.

[287] J. S. Dowker, *Quantum Field Theory on a Cone*, *J. Phys.* **A10** (1977) 115–124.

[288] J. S. Dowker, *Thermal properties of Green's functions in Rindler, de Sitter, and Schwarzschild spaces*, *Phys. Rev.* **D18** (1978), no. 6 1856.

[289] H. S. Carslaw, *Diffraction of waves by a wedge of any angle*, *Proc. London Math. Soc.* **18** (1919) 291–306.

[290] V. Moretti and L. Vanzo, *Thermal Wightman functions and renormalized stress tensors in the Rindler wedge*, *Phys. Lett.* **B375** (1996) 54–59, [hep-th/9507139].

[291] M. R. Brown and A. C. Ottewill, *Effective actions and conformal transformations*, *Phys. Rev.* **D31** (1985) 2514–2520.

[292] P. Candelas and D. Deutsch, *On the vacuum stress induced by uniform acceleration or supporting the ether*, *Proc. Roy. Soc. Lond.* **A354** (1977) 79–99.

[293] P. Candelas, *Vacuum polarization in schwarzschild spacetime*, *Physical Review D* **21** (apr, 1980) 2185–2202.

[294] A. Higuchi, G. E. A. Matsas, and D. Sudarsky, *Bremsstrahlung and fulling-davies-unruh thermal bath*, *Physical Review D* **46** (oct, 1992) 3450–3457.

[295] V. Moretti, *Canonical quantization of photons in a Rindler wedge*, *J. Math. Phys.* **38** (1997) 2922–2953, [gr-qc/9603057].

[296] V. Moretti, *Euclidean thermal Green functions of photons in generalized Euclidean Rindler spaces for any Feynman - like gauge*, *Int. J. Mod. Phys.* **A12** (1997) 3787–3798, [hep-th/9607178].

[297] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, 1984.

[298] R. Szmytkowski and S. Bielski, *Comment on the orthogonality of the macdonald functions of imaginary order*, *Journal of Mathematical Analysis and Applications* **365** (may, 2010) 195–197.

[299] J. S. Dowker, *Casimir Effect Around a Cone*, *Phys. Rev.* **D36** (1987) 3095.

[300] D. N. Kabat, *Black hole entropy and entropy of entanglement*, *Nucl. Phys.* **B453** (1995) 281–299, [hep-th/9503016].

[301] I. Affleck and A. W. W. Ludwig, *Universal noninteger "ground-state degeneracy" in critical quantum systems*, *Physical Review Letters* **67** (July, 1991) 161–164.

[302] T. M. Dunster, *Bessel functions of purely imaginary order, with an application to second-order linear differential equations having a large parameter*, *SIAM Journal on Mathematical Analysis* **21** (July, 1990) 995–1018.

[303] A. Almheiri, X. Dong, and D. Harlow, *Bulk Locality and Quantum Error Correction in AdS/CFT*, *JHEP* **04** (2015) 163, [arXiv:1411.7041].

[304] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, *Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence*, *JHEP* **06** (2015) 149, [arXiv:1503.0623].

[305] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, *Holographic duality from random tensor networks*, arXiv:1601.0169.

[306] D. Harlow, *The Ryu-Takayanagi Formula from Quantum Error Correction*, arXiv:1607.0390.

[307] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602, [hep-th/0603001].

[308] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045, [hep-th/0605073].

[309] T. Faulkner, A. Lewkowycz, and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074, [arXiv:1307.2892].

[310] P. V. Buividovich and M. I. Polikarpov, *Entanglement entropy in gauge theories and the holographic principle for electric strings*, *Phys. Lett.* **B670** (2008) 141–145, [arXiv:0806.3376].

[311] X.-L. Qi, *Exact holographic mapping and emergent space-time geometry*, arXiv:1309.6282.

[312] Z. Yang, P. Hayden, and X.-L. Qi, *Bidirectional holographic codes and sub-AdS locality*, *JHEP* **01** (2016) 175, [arXiv:1510.0378].

[313] W. Donnelly and L. Freidel, *Local subsystems in gauge theory and gravity*, *JHEP* **09** (2016) 102, [arXiv:1601.0474].

[314] B. Michel, E. Mintun, J. Polchinski, A. Puhm and P. Saad, *Remarks on brane and antibrane dynamics*, JHEP **1509**, 021 (2015) doi:10.1007/JHEP09(2015)021 [arXiv:1412.5702 [hep-th]].

[315] B. Michel, J. Polchinski, V. Rosenhaus and S. J. Suh, *Four-point function in the IOP matrix model*, JHEP **1605**, 048 (2016) doi:10.1007/JHEP05(2016)048 [arXiv:1602.06422 [hep-th]].

[316] A. Belin, J. de Boer, J. Kruthoff, B. Michel, E. Shaghoulian and M. Shyani, *Universality of sparse $d > 2$ conformal field theory at large N*, JHEP **1703**, 067 (2017) doi:10.1007/JHEP03(2017)067 [arXiv:1610.06186 [hep-th]].

[317] W. Donnelly, B. Michel, D. Marolf and J. Wien, *Living on the Edge: A Toy Model for Holographic Reconstruction of Algebras with Centers*, JHEP **1704**, 093 (2017) doi:10.1007/JHEP04(2017)093 [arXiv:1611.05841 [hep-th]].

[318] B. Michel, *Remarks on Rindler Quantization*, arXiv:1612.03158 [hep-th].

[319] D. Marolf, B. Michel and A. Puhm, *A rough end for smooth microstate geometries*, JHEP **1705**, 021 (2017) doi:10.1007/JHEP05(2017)021 [arXiv:1612.05235 [hep-th]].

[320] B. Michel and M. Srednicki, *Entanglement Entropy and Boundary Conditions in 1+1 Dimensions*, arXiv:1612.08682 [hep-th].

[321] M. C. Begelman, R. D. Blandford, and M. J. Rees, *Massive black hole binaries in active galactic nuclei*, *Nature* **287** (1980), no. 5780 307–309.

[322] R. Genzel, R. Schdel, T. Ott, F. Eisenhauer, R. Hofmann, M. Lehnert, A. Eckart, T. Alexander, A. Sternberg, R. Lenzen, Y. Clnet, F. Lacombe, D. Rouan, A. Renzini, and L. E. Tacconi-Garman, *The stellar cusp around the supermassive black hole in the galactic center*, *The Astrophysical Journal* **594** (2003), no. 2 812.

[323] B. A. et al, *Observation of gravitational waves from a binary black hole merger*, *Physical Review Letters* **116** (feb, 2016).

[324] R. J. McLure and M. J. Jarvis, *Measuring the black hole masses of high-redshift quasars*, *Monthly Notices of the Royal Astronomical Society* **337** (2002), no. 1 109–116.

[325] R. A. Remillard and J. E. McClintock, *X-ray properties of black-hole binaries*, *Annual Review of Astronomy and Astrophysics* **44** (sep, 2006) 49–92.

[326] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113–1133, [hep-th/9711200]. [Adv. Theor. Math. Phys.2,231(1998)].

[327] S. Sachdev, *Condensed Matter and AdS/CFT*, arXiv:1002.2947. [Lect. Notes Phys.828,273(2011)].

[328] D. Marolf and J. Polchinski, *Violations of the Born rule in cool state-dependent horizons*, *JHEP* **01** (2016) 008, [arXiv:1506.0133].

[329] J. M. Maldacena, *Black holes in string theory*. PhD thesis, Princeton U., 1996. hep-th/9607235.

[330] C. M. Hull and P. K. Townsend, *Unity of superstring dualities*, *Nucl. Phys.* **B438** (1995) 109–137, [hep-th/9410167].

[331] E. Witten, *String theory dynamics in various dimensions*, *Nucl. Phys.* **B443** (1995) 85–126, [hep-th/9503124].

[332] A. Hanany and E. Witten, *Type IIB superstrings, BPS monopoles, and three-dimensional gauge dynamics*, *Nucl. Phys.* **B492** (1997) 152–190, [hep-th/9611230].

[333] A. Marshakov, *Seiberg-Witten theory, integrable systems and D-branes*, in *Landau Summer University for Theoretical Physics Chernogolovka, Russia, June 23-August 20, 1997*, pp. 279–286, 1997. hep-th/9709001.

[334] F. Benini, S. Benvenuti, and Y. Tachikawa, *Webs of five-branes and N=2 superconformal field theories*, *JHEP* **09** (2009) 052, [arXiv:0906.0359].

[335] S. W. Hawking and D. N. Page, *Thermodynamics of Black Holes in anti-De Sitter Space*, *Commun. Math. Phys.* **87** (1983) 577.

[336] I. Bena, M. Graa, S. Kuperstein, and S. Massai, *Giant Tachyons in the Landscape*, *JHEP* **02** (2015) 146, [arXiv:1410.7776].

[337] J. Polchinski, *String Theory, Volumes 1 and 2*. Cambridge University Press, Cambridge, first ed., 2005.

[338] L. J. Dixon, V. Kaplunovsky, and C. Vafa, *On Four-Dimensional Gauge Theories from Type II Superstrings*, *Nucl. Phys.* **B294** (1987) 43–82.

[339] I. Antoniadis, J. R. Ellis, J. S. Hagelin, and D. V. Nanopoulos, *GUT Model Building with Fermionic Four-Dimensional Strings*, *Phys. Lett.* **B205** (1988) 459–465.

[340] M. Cvetic, T. Li, and T. Liu, *Standard-like models as type IIB flux vacua*, *Phys. Rev.* **D71** (2005) 106008, [hep-th/0501041].

[341] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370–3373, [hep-ph/9905221].

[342] S. Kachru, R. Kallosh, A. D. Linde, J. M. Maldacena, L. P. McAllister, and S. P. Trivedi, *Towards inflation in string theory*, *JCAP* **0310** (2003) 013, [hep-th/0308055].

[343] R. Bousso and J. Polchinski, *Quantization of four form fluxes and dynamical neutralization of the cosmological constant*, *JHEP* **06** (2000) 006, [hep-th/0004134].

[344] F. Denef and M. R. Douglas, *Computational complexity of the landscape. I.*, *Annals Phys.* **322** (2007) 1096–1142, [hep-th/0602072].

[345] F. Denef, M. R. Douglas, B. Greene, and C. Zukowski, *Computational complexity of the landscape II - Cosmological considerations*, arXiv:1706.0643.

[346] P. Candelas, G. T. Horowitz, A. Strominger, and E. Witten, *Vacuum Configurations for Superstrings*, *Nucl. Phys.* **B258** (1985) 46–74.

[347] S. B. Giddings, S. Kachru, and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev.* **D66** (2002) 106006, [hep-th/0105097].

[348] R. Blumenhagen, B. Kors, D. Lust, and S. Stieberger, *Four-dimensional String Compactifications with D-Branes, Orientifolds and Fluxes*, *Phys. Rept.* **445** (2007) 1–193, [hep-th/0610327].

[349] J. Maldacena, D. Simmons-Duffin, and A. Zhiboedov, *Looking for a bulk point*, *JHEP* **01** (2017) 013, [arXiv:1509.0361].

[350] M. A. Levin and X.-G. Wen, *String-net condensation: a physical mechanism for topological phases*, *Physical Review B* **71** (jan, 2005).

[351] A. Kitaev and J. Preskill, *Topological entanglement entropy*, *Phys. Rev. Lett.* **96** (2006) 110404, [hep-th/0510092].

[352] A. Yu. Kitaev, *Fault tolerant quantum computation by anyons*, *Annals Phys.* **303** (2003) 2–30, [quant-ph/9707021].

[353] D. Gottesman, *Stabilizer codes and quantum error correction*, arXiv:quant-ph/9705052.

[354] J. Preskill, *Fault tolerant quantum computation*, CALT-68-2150, QUIC-97-034, arXiv:quant-ph/9712048.

[355] M. Z. Hasan and C. L. Kane, *Topological Insulators*, *Rev. Mod. Phys.* **82** (2010) 3045, [arXiv:1002.3895].

[356] T. Grover, *Entanglement monotonicity and the stability of gauge theories in three spacetime dimensions*, *Physical Review Letters* **112** (apr, 2014).

[357] H. Liu and S. J. Suh, *Entanglement Tsunami: Universal Scaling in Holographic Thermalization*, *Phys. Rev. Lett.* **112** (2014) 011601, [arXiv:1305.7244].

[358] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, *Holographic Entanglement Entropy from 2d CFT: Heavy States and Local Quenches*, *JHEP* **02** (2015) 171, [arXiv:1410.1392].

[359] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, arXiv:0907.2939.

[360] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, *Gravitation from Entanglement in Holographic CFTs*, *JHEP* **03** (2014) 051, [arXiv:1312.7856].

[361] T. Faulkner, F. M. Haehl, E. Hijano, O. Parrikar, C. Rabideau, and M. Van Raamsdonk, *Nonlinear Gravity from Entanglement in Conformal Field Theories*, *JHEP* **08** (2017) 057, [arXiv:1705.0302].

[362] C. Montonen and D. Olive, *Magnetic monopoles as gauge particles?*, *Physics Letters B* **72** (dec, 1977) 117–120.

[363] N. Seiberg and E. Witten, *Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory*, *Nucl. Phys.* **B426** (1994) 19–52, [hep-th/9407087]. [Erratum: Nucl. Phys.B430,485(1994)].

[364] J. S. Dowker, *Entanglement entropy for even spheres*, arXiv:1009.3854.

[365] X. Dong, D. Harlow, and A. C. Wall, *Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality*, *Phys. Rev. Lett.* **117** (2016), no. 2 021601, [arXiv:1601.0541].

[366] B. Swingle, *Entanglement renormalization and holography*, *Physical Review D* **86** (sep, 2012).

[367] G. Vidal, *Entanglement renormalization*, *Physical Review Letters* **99** (nov, 2007).