The geometry and topology of quantum entanglement in holography

HENRY DAVID MAXFIELD

The geometry and topology of quantum entanglement in holography

Henry Maxfield

A Thesis presented for the degree of
Doctor of Philosophy

Department of Mathematical Sciences
Durham University

March 2015

The geometry and topology of quantum entanglement in holography

Henry Maxfield

Submitted for the degree of Doctor of Philosophy
March 2015

Abstract

In this thesis I explore the connection between geometry and quantum entanglement, in the context of holographic duality, where entanglement entropies in a quantum field theory are associated with the areas of surfaces in a dual gravitational theory.

The first chapter looks at a phase transition in such systems in finite size and at finite temperature, associated with the properties of minimal surfaces in a static black hole background. This is followed by the related problem of extremal surfaces in a spacetime describing the dynamical process of black hole formation, with a view towards understanding the connections between bulk locality and various field theory observables including entanglement entropy.

The third chapter looks at the simple case of pure gravity in three spacetime dimensions, where I show how evaluating the entanglement entropy can be reduced to a simple algebraic calculation, and apply it to some interesting examples.

Finally, the rôle played by topology of surfaces in a proposed derivation of a holographic entanglement entropy formula is investigated. This makes it clear what assumptions are required in order to reproduce the 'homology constraint', a topological condition necessary for consistency with field theory.

Declaration

The work in this thesis is based on research carried out at the Department of Mathematical Sciences at Durham University. The results are based on the following collaborative works:

- V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, *Holographic entanglement plateaux*, *JHEP* **1308** (2013) 092, [arXiv:1306.4004]
- V. E. Hubeny and H. Maxfield, *Holographic probes of collapsing black holes*, *JHEP* **1403** (2014) 097, [arXiv:1312.6887]
- H. Maxfield, *Entanglement entropy in three dimensional gravity*, arXiv:1412.0687, to appear in *JHEP*.
- F. M. Haehl, T. Hartman, D. Marolf, H. Maxfield, and M. Rangamani, *Topological aspects of generalized gravitational entropy*, arXiv:1412.7561, to appear in *JHEP*.

No part of this thesis has been submitted for a degree in this or any other institution.

Acknowledgements

This thesis owes its existence to many people, and I would like to take this opportunity to thank just a few of them.

Firstly, my parents, for their unwavering support and encouragement, and the value they placed on all aspects of education.

My supervisor Veronika Hubeny, for her knowledge, enthusiasm and patience.

Tadashi Tokieda, for teaching me how to think like a mathematician, and Mukund Rangamani, from whom I learned a great deal about thinking like a physicist too.

My housemates Andy and Alex, for their camaraderie, and for putting up with my idiosyncrasies.

Finally, to all the sailors and windsurfers who helped make my time at Durham (and Cambridge before that) such a great experience.

My studies were supported by an STFC studentship.

Contents

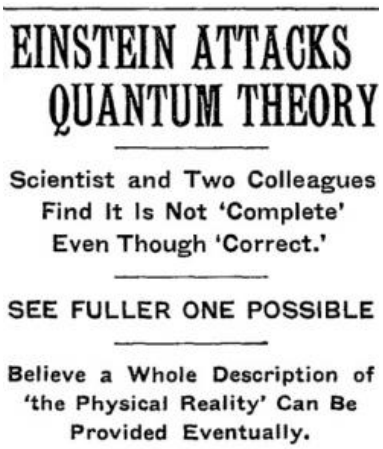
Chapter 1

Introduction

The overriding theme of this thesis is the exploration of a deep connection between two seemingly unrelated concepts in the fields of quantum information and geometry.

The first of these is the idea of quantum entanglement, the feature of quantum mechanics that most starkly contrasts it from classical theories. This distinction between the nature of quantum and classical information was demonstrated in the seminal five page paper of Bell [5], who showed that the way information is shared between even the simplest quantum mechanical systems is inherently nonlocal. This consequence of quantum mechanics is so counterintuitive that many refused to accept it, as illustrated by the 1935 *New York Times* headline reporting the paper of Einstein, Podolsky and Rosen that would later inspire Bell.

The consequences of the nature of quantum entanglement are far reaching, most obviously in applications such as quantum computation and cryptography, but also in fields such as condensed matter physics. It has recently even received some high-profile attention in a surprising biological context, with some evidence that robins may use entanglement directly for navigation [6], which exemplifies the possibilities for unexpected applications. Of central importance for this thesis, it also appears to have central significance in quantum mechanical theories of gravity, which brings us to the second of the themes.


EINSTEIN ATTACKS QUANTUM THEORY

Scientist and Two Colleagues Find It Is Not 'Complete' Even Though 'Correct.'

SEE FULLER ONE POSSIBLE

Believe a Whole Description of 'the Physical Reality' Can Be Provided Eventually.

Figure 1.1: New York Times headline from 1935 reporting the EPR paper

Einstein's major insight about gravity is that it is geometric, which explains the fact, famously (though apocryphally) demonstrated by Galileo dropping balls of different masses from the Leaning Tower of Pisa, that the path of an object falling freely in a gravitational field is independent of its mass. A theory of gravity is a theory of the shape of spacetime, of lengths, angles, and areas, and of how matter shapes this geometry and moves within it. This gives new life and importance to many classical problems in geometry. One example of such a problem, with a 350 year history beginning with Lagrange, is that of finding a

minimal surface: a surface of smallest area with a given boundary. This problem is often associated with the name of Joseph Plateau, who studied soap films in the 19^{th} century, which form minimal surfaces to minimise their energy due to surface tension.

In stark contrast to down-to-earth application to bubbles of soap, this thesis will revolve around applying Plateau's problem, appropriately generalised to the geometries of gravity, to study quantum entanglement. The connection is far from obvious, and most of the introduction will be dedicated to explaining it in detail, but the eventual upshot is that entanglement, in an appropriate context, can be measured by finding the area of a minimal surface.

The motivations for the study are at least twofold. Firstly, the geometrical picture gives us a new framework for quantum entanglement in which to calculate and build intuition, usually in regimes where calculations by more direct methods are prohibitively difficult, or impossible. The second, which is the primary motivation for the work contained in this thesis, is what we can learn from the connection itself about the nature of quantum gravity.

Much of what is known about quantum gravity is obtained from a semiclassical approximation, which is perhaps inevitable given the difficulties of defining a full theory and of calculations away from such an approximation. But this has led to several paradoxes, which continue to elude consensus on a satisfactory resolution. Most notably the information paradox and variants consider formation and evaporation of a black hole, and derive a contradiction assuming the equivalence principle, validity of local effective field theory, and unitarity of the process, which all seem very difficult to give up. Given such a profound and direct connection between geometry and inherently quantum mechanical ideas, by better understanding the connection one might hope to learn how the geometry emerges from some more fundamental description, and resolve the outstanding questions in quantum gravity.

1.1 Quantum entanglement

> I would not call [entanglement] *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.
>
> Erwin Schrödinger, 1935 [7]

Schrödinger's quote, in his response to the famous EPR gedanken experiment [8], remains truer than ever 80 years on. If we want to understand a quantum mechanical system fully, we have to abandon classical intuition, and accept that certain phenomena are inherently nonclassical due to quantum correlations.

Classically, if we have a system consisting of several subsystems, knowledge of the state of all subsystems constitutes knowledge of the whole: a point in phase space is no more than the collection of all such data for the subsystems. In quantum mechanics, this is no longer true. The state of a system is inherently nonlocal, and the information to describe

the physics for each part is insufficient to determine the physics for the whole.

Take the simplest example of two qubits (perhaps the spin degrees of freedom of a pair of spin-half particles), each described by a two-dimensional Hilbert space spanned by $|0\rangle$ and $|1\rangle$ (the eigenstate of spin along the z-axis, say). Prepare them in a state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right), \tag{1.1}$$

which can be arranged experimentally, for example, by the decay of a particle with zero angular momentum into a pair of spin-half particles, so that they are in a singlet state. Now send each qubit to a different experimenter, and allow them to perform any measurement. Repeating this many times, the experimenter will discover that she is equally likely to measure spin about *any* axis as up or down. If she tried to describe her qubit with some pure state, this would be a contradiction: a spin-half particle always has definite spin up in some direction. She might conclude that the qubits she is sent are not identically prepared, but rather in some ensemble of several different states. But if the experimenters compare data, they will find that their results are correlated: for example, in the event that for some pair of particles they both measure spin in the same direction, they will find that their results always give opposite answers. The apparent ensemble nature of the state of each qubit alone is due to the loss of the information about the state shared nonlocally between the two qubits, and available to neither individually. The amount of such information can be quantified by the *entanglement entropy*, which is a single qubit in this case.

More generally, suppose you have a system which splits into two separate parts, so the Hilbert space of the system is a tensor product of the subsystem Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Now prepare some pure state $|\psi\rangle \in \mathcal{H}$, with associated density matrix $\rho = |\psi\rangle\langle\psi|$. Now how do we describe the physics in one of the subsystems, supposing that we have no access to the other? We can throw away all the information in subsystem 2 by performing a partial trace over that part of the Hilbert space, to get the reduced density matrix $\rho_1 = \mathrm{Tr}_2(\rho)$. Now if $\mathcal{O}$ is any observable acting in subsystem 1 only, that is an operator on $\mathcal{H}_1$ (or $\mathcal{O} \otimes \mathbb{1}_2$ as an operator on $\mathcal{H}$), its expectation value in the state $|\psi\rangle$ is given by $\langle\mathcal{O}\rangle_\psi = \mathrm{Tr}_1(\rho_1\mathcal{O})$. In this sense, ρ_1 contains all the information about physics in subsystem 1 alone.

Now, crucially, this physics can not usually be described by a pure state, an element of $\mathcal{H}_1$, but requires the reduced density matrix to be a mixed state. The fact that this ensemble description is required is due to the correlations with the remainder of the system: the entanglement between 1 and 2. A natural way to measure this entanglement is therefore to count the number of different states that would be required to make an ensemble with density matrix ρ_1. This number of states (or, rather, its log) for a density matrix ρ is given by the von Neumann entropy

$$S(\rho) = -\,\mathrm{Tr}(\rho \log \rho), \tag{1.2}$$

which gives $\log n$ for an ensemble of n orthogonal states in equal proportions[1]. The *entanglement entropy* is defined as the von Neumann entropy of the reduced density matrix.

The same definition can be made when the total system itself is in a mixed state, and this is still called entanglement entropy, though it should be borne in mind that this measures not only the quantum correlations, but also a classical contribution.

There are many other useful measures of entanglement besides entanglement entropy. Some of these are defined in terms of what information-theoretic protocols one can perform on the system (such as entanglement cost, and distillable entanglement); some attempt to distinguish classical correlation from quantum correlation (such as negativity); some aim to measure multipartite entanglement between many subsystems. Many of these, particularly in pure states, reduce to entanglement entropy, or can be expressed in terms of it, showing the ubiquity of this quantity. In any case, mostly due to the relationship with geometry described later, it is the entanglement entropy that will be the main focus of this thesis.

1.1.1 Properties of entanglement entropy

There are many general properties of entanglement entropy following from the definition, which are useful and physical, and emphasise that this is a natural quantity to consider.

Perhaps the simplest of these is that, for a pure state, the entanglement entropy associated with one subsystem is equal to that of its complement. This means that we may freely talk about the entanglement 'between' the subsystems, and the apparent asymmetry between the two Hilbert spaces in the decomposition is illusory. This (along with many other properties) can be proven simply from the existence of a Schmidt basis for any given state. Given some $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, this a pair of orthonormal sets $\{|u_1\rangle, \cdots, |u_n\rangle\} \subseteq \mathcal{H}_1$ and $\{|v_1\rangle, \cdots, |v_n\rangle\} \subseteq \mathcal{H}_2$, with $n \leq \min\{\dim \mathcal{H}_1, \dim \mathcal{H}_2\}$, such that

$$|\psi\rangle = \sum_{i=1}^{n} \lambda_i^{1/2} |u_i\rangle \otimes |v_i\rangle \tag{1.3}$$

where the λ_i are real and positive, and sum to unity if the state is normalised. It follows that the reduced density matrices are given by

$$\rho_1 = \sum_{i=1}^{n} \lambda_i |u_i\rangle\langle u_i|, \quad \rho_2 = \sum_{i=1}^{n} \lambda_i |v_i\rangle\langle v_i| \tag{1.4}$$

From this it is manifest that not only are the von Neumann entropies of the reduced density matrices equal, but the much stronger statement that their spectra, given by the λ_i, are identical (up to a set of zero eigenvalues). This means that any measure of entanglement constructed from these eigenvalues, such as the Rényi entropies that we will introduce

[1] If this log is taken with base 2, this can be given the unit of bits (also 'shannons' in the context of information theory), or qubits to emphasise the quantum nature of the information. For natural logarithms, there exists the analogous term 'nat' (or sometimes nit, or nepit, named for John Napier, Latinized as Neperi, the discoverer of the logarithm), standing for 'natural unit of information'. This suggests the term qunat for a 'natural unit of quantum information', though I know of nowhere this has been used.

later, shares this property.

One immediate consequence of this is that the entanglement entropy can be at most the log of the dimension of the smaller subsystem. This means that maximal entropy can be used to count the number of degrees of freedom of a system, which will be useful when we turn to look at quantum gravity. In fact, in a generic state, when the subsystems are large, the entanglement entropy will be close to maximal [9].

Apart from this, the entanglement entropy obeys a number of inequalities. These require splitting the degrees of freedom into many sets $\mathcal{A}, \mathcal{B}, \ldots$, so the Hilbert space correspondingly decomposes as a tensor product with many factors, and considering the entanglement entropy associated with various associated bipartitionings. The most important of these is the *strong subadditivity*, or convexity, of entanglement. For this, consider dividing the Hilbert space into three subsystems $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and possibly a fourth comprising the remainder of the system:

$$S(\mathcal{AB}) + S(\mathcal{BC}) \geq S(\mathcal{B}) + S(\mathcal{ABC}) \tag{1.5}$$

This remains true if the total system is in a mixed state. In fact, this distinction is unimportant, because any system in a mixed state ρ can be adjoined to an auxiliary 'purifying' system, so that ρ is the reduced density matrix of some pure state when the auxiliary system is traced out [10]. Using this fact, along with the equality of entanglement entropy for complementary subsystems in a pure state, strong subadditivity can be specialised to get various other inequalities. Two important examples are the 'triangle inequalities'

$$|S(\mathcal{A}) - S(\mathcal{B})| \leq S(\mathcal{AB}) \leq S(\mathcal{A}) + S(\mathcal{B}), \tag{1.6}$$

respectively known as the Araki-Lieb inequality and subadditivity of entropy. Note that when $\rho_{\mathcal{AB}}$ is pure, the Araki-Lieb inequality reproduces the equality of entanglement entropy for a subsystem and its complement.

The second of these inequalities helps to motivate the definition of *mutual information*,

$$I(\mathcal{A} : \mathcal{B}) = S(\mathcal{A}) + S(\mathcal{B}) - S(\mathcal{AB}), \tag{1.7}$$

which is manifestly symmetric in $\mathcal{A}$ and $\mathcal{B}$, and positive from subadditivity. It measures the amount of correlation between the two subsystems (both classical and quantum), and can be used to bound the correlation functions of any operators $\mathcal{O}_{\mathcal{A}}$ and $\mathcal{O}_{\mathcal{B}}$ acting only within subsystems $\mathcal{A}$ and $\mathcal{B}$ respectively [11], by

$$I(\mathcal{A} : \mathcal{B}) \geq \frac{1}{2}\left(\frac{\langle \mathcal{O}_{\mathcal{A}}\mathcal{O}_{\mathcal{B}}\rangle - \langle \mathcal{O}_{\mathcal{A}}\rangle\langle \mathcal{O}_{\mathcal{B}}\rangle}{\|\mathcal{O}_{\mathcal{A}}\|\|\mathcal{O}_{\mathcal{B}}\|}\right)^2, \tag{1.8}$$

where $\|\cdot\|$ denotes the operator ∞-norm.

Saturation of these inequalities often has a direct equivalent interpretation in terms of the reduced density matrices. For example, the mutual information vanishes if and only if

the density matrix factorises:

$$I(\mathcal{A}:\mathcal{B}) = 0 \iff \rho_{\mathcal{AB}} = \rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}} \tag{1.9}$$

Similarly, the saturation of the Araki-Lieb inequality implies that one of the subsystems can be split into two parts, one purifying the other and carrying all the correlations, and one carrying the combined entropy:

$$S(\mathcal{A}) = S(\mathcal{B}) + S(\mathcal{AB}) \iff \mathcal{A} = \mathcal{A}_1\mathcal{A}_2 \text{ with } \rho_{\mathcal{AB}} = \rho_{\mathcal{A}_1} \otimes (|\psi\rangle\langle\psi|)_{\mathcal{A}_2\mathcal{B}} \tag{1.10}$$

1.1.2 Entanglement in field theory

In a quantum field theory, there is a very natural way to partition the system, by separating the degrees of freedom into different spatial regions. Entanglement entropy can be usefully applied to measure the entanglement between such spatial regions.

To make this more precise, one can consider formulating the theory on a lattice, so that the Hilbert space is a product of Hilbert spaces for each lattice point[2], $\mathcal{H} = \otimes_{i\in\text{lattice}}\mathcal{H}_i$. Then given a spatial region $\mathcal{A}$, let $\mathcal{H}_{\mathcal{A}}$ be the product of Hilbert spaces at lattice sites located inside $\mathcal{A}$, and $\mathcal{H}^c_{\mathcal{A}}$ the product over the remaining lattice sites so that $\mathcal{H} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}^c_{\mathcal{A}}$. Now the quantum mechanical definitions can be applied to give the entanglement entropy associated with a region in some state of the theory.

From this description it is far from clear that the quantity we obtain is sensible or useful. In particular, when we take the lattice spacing to zero to pass to the continuum field theory, it is not obvious that we find any information that is insensitive to the UV cutoff scale, or characteristic of the field theory independently of the choice of how it is implemented on the lattice.

In a generic state the entropy will be nearly maximal, so will scale with the number of lattice sites, so $S(\mathcal{A}) \sim (\frac{L}{\epsilon})^{d-1}$, where L is the size of $\mathcal{A}$, ϵ the lattice spacing, and d the number of spacetime dimensions. In this case the entanglement entropy is UV divergent. However, a generic state is very highly excited, and things are much improved for the low-lying states we are more usually interested in, such as the vacuum: the entanglement entropy is still UV divergent, but in many cases scales with the number of links between neighbouring lattice sites that are cut by the bipartitioning of the system. In the limit of small lattice spacing, this leads to an *area law*:

$$S(A) \sim \frac{\text{Area}(\partial\mathcal{A})}{\epsilon^{d-2}} \quad (\text{or } S(A) \sim -\log\epsilon \text{ for } d = 2) \tag{1.11}$$

This was first shown for a spherical region in massless free field [12], and has since been shown to be true in a wide variety of other circumstances. This has a natural physical interpretation: there is very little entanglement in these states, and it is mostly concentrated locally at the interface between $\mathcal{A}$ and $\mathcal{A}^c$.

[2]This is more subtle in gauge theories, due to link variables and gauge redundancies.

Beyond this nonuniversal divergent piece, and further subleading divergent terms constructed from geometric invariants of $\partial\mathcal{A}$, there remains a cutoff independent finite contribution (see [13], for example), which may be useful for characterising the theory, or the state of the system, as we will see in some examples below. Furthermore, in some combinations of entanglement entropies the divergent pieces cancel, so that the quantity is finite and cutoff independent without any additional prescription for extracting the universal piece. The simplest such example is the mutual information for two regions with no shared boundary: it is manifest that the divergent area contributions from $S(\mathcal{A}) + S(\mathcal{B})$ cancel exactly those from $S(\mathcal{AB})$.

So it would certainly appear that there are universal quantities that can be found from the entanglement entropy in field theory. There are several contexts in which these are very useful, a small sample of which follow.

One of the early examples is for critical models in two dimensions. These are described by conformal field theories, for which there is a universal answer for the entanglement entropy of an interval in the ground state, as we will review later. This depends only on one parameter, the central charge c of the field theory. Hence this universality of entanglement entropy can be used, for example, to find critical points in lattice simulations, and to compute the central charge of the model.

A second important example is the topological entanglement entropy [14, 15]. Certain systems in two spatial dimensions can undergo quantum phase transitions to a topologically ordered state, which is not characterised by any local order parameter. Cutoff-independent quantities constructed from the entanglement entropies have been proposed as novel order parameters for such phases.

Another application is to study the space of field theories in relation to the renormalisation group flow, by using entanglement entropy to generalise Zamalodchikov's famous c-theorem, which formalises the idea that integrating out high energies reduces the number of degrees of freedom. Zamalodchikov proved that in two dimensions there is a c-function associated to quantum field theories, that decreases monotonically under RG flow, and at fixed points equals the central charge of the corresponding conformal field theory. Much later, entanglement entropy was suggested as a means to define an 'entropic' c-function [16], which may be extended to higher dimensions. This is particularly useful in odd dimensions, where there is no conformal anomaly to characterise the number of degrees of freedom.

1.1.3 Computing entanglement entropy

Having given a collection of applications to motivate the study of entanglement entropy in quantum field theory, we must now face the unfortunate fact that it is not measurable experimentally, and very rarely calculable theoretically. Of the methods of calculation available, perhaps the most important is the replica trick.

The starting point for this is to notice that the entanglement entropy is in fact just one

member of a continuous family of measures of entanglement, the *Rényi entropies*[3] of the reduced density matrix, defined as

$$S_q(\rho_{\mathcal{A}}) := \frac{1}{1-q} \log \mathrm{Tr}(\rho_{\mathcal{A}}^q) \tag{1.12}$$

In the limit $q \to 1$, this becomes the entanglement entropy. The point of the replica trick is that $\mathrm{Tr}(\rho_{\mathcal{A}}^q)$ may be calculable when q is a positive integer. Having done this for all such values of q, it may be possible to fit this to a function on some larger complex domain, using analyticity[4] and growth properties to determine the correct continuation (or, more usually, simply performing the most natural continuation in the parameter q). Then the $q \to 1$ limit may be taken to recover the entanglement entropy.

It remains to describe how to compute $\mathrm{Tr}(\rho_{\mathcal{A}}^q)$ for an integer exponent q. This can be performed in principle by a certain path integral over q 'replicas' of a Euclidean spacetime, hence the name. An illustration of the method, for the thermal state of a two-dimensional field theory on a circle, is given in fig. 1.2, to show each stage described in the text.

The first ingredient required for this is a path integral definition of the state. The most important example is the vacuum, whose wavefunction evaluated on a particular field configuration is given by the path integral over semi-infinite Euclidean time, with boundary conditions specified by that configuration:

$$\langle \phi | \mathrm{VAC} \rangle = \int\limits_{\substack{\tau \in (-\infty, 0] \\ \Phi(\tau=0)=\phi}} \mathcal{D}\Phi \; e^{-S[\Phi]} \tag{1.13}$$

Here the $|\phi\rangle$ denote eigenvectors of some field operators, with eigenvalues given by the field configuration ϕ. The associated density matrix elements $\langle \phi' | \mathrm{VAC} \rangle \langle \mathrm{VAC} | \phi \rangle$ are then given by two copies of this, the second time-reversed with boundary condition ϕ' on the fields at $\tau = 0$.

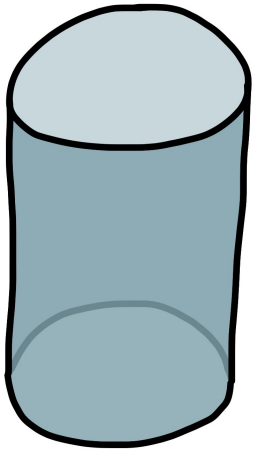
For another example, the thermal state, given by the Maxwell-Bolzmann distribution $\rho_{\mathrm{MB}} \propto e^{-\beta H}$ at temperature $T = 1/\beta$, has matrix elements computed by a similar path integral over an interval of length β in Euclidean time:

$$\langle \phi' | e^{-\beta H} | \phi \rangle = \int\limits_{\substack{\tau \in [0, \beta] \\ \Phi(\tau=0)=\phi \\ \Phi(\tau=\beta)=\phi'}} \mathcal{D}\Phi \; e^{-S[\Phi]} \tag{1.14}$$

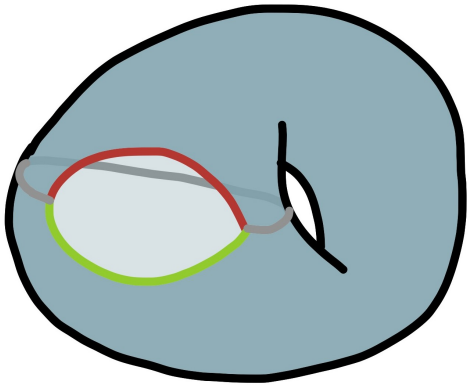
Other states can be built by altering the geometry, inserting operators, introducing some Lorentzian piece to the path integral to implement time evolution, and so forth.

[3] Alfréd Rényi is famous not only for his definition of these entropies, but also of mathematicians: 'A mathematician is a device for turning coffee into theorems'.

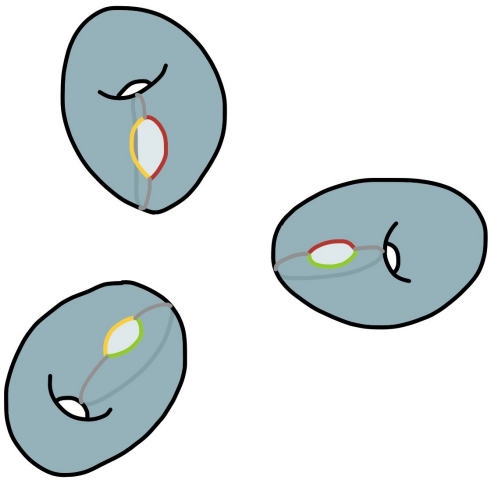
[4] It should be noted that this is not strictly analytic continuation, since analyticity alone is insufficient to uniquely specify a function known only at the positive integers. For example, $\Gamma(-n)^{-1}$ is an entire function with zeroes at all positive integral n. Demanding that the Rényi entropy is analytic at infinity would be sufficient for uniqueness, but it is not clear if this is the correct prescription.



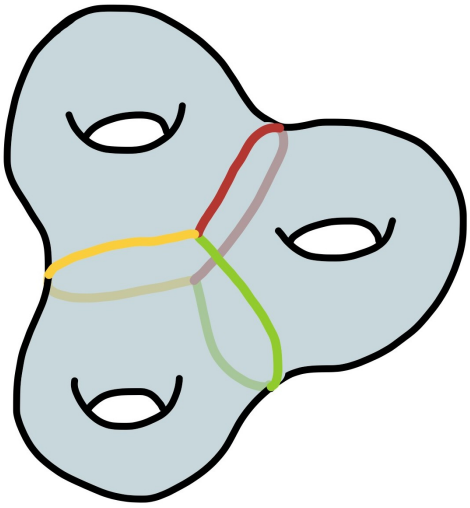
(a) The cylinder, of length β, represents the thermal density matrix $\rho = e^{-\beta H}$.



(b) The top and bottom of the cylinder are glued along $\mathcal{A}^c$. A path integral with specified boundary conditions on $\mathcal{A}$, shown by the red and green intervals, gives matrix elements of $\rho_{\mathcal{A}}$.



(c) Three copies of matrix elements of $\rho_{\mathcal{A}}$, with field configurations on the boundary identified cyclically by colours.



(d) Gluing the replicas along matching colours gives $\mathrm{Tr}(\rho_{\mathcal{A}}^3)$, by a path integral on this genus 3 surface.

Figure 1.2: Cartoons illustrating the replica construction for computing $\mathrm{Tr}(\rho_{\mathcal{A}}^3)$, for $\mathcal{A}$ a single interval, in a two-dimensional field theory on a circle, in the thermal state. The metric on the surfaces is not faithfully represented.

With this description in hand, we must now obtain the reduced density matrix $\rho_{\mathcal{A}}$ by tracing out the complementary region $\mathcal{A}^c$. This means identifying field configurations of ϕ and ϕ' in $\mathcal{A}$, and then integrating over all such configurations; this serves to essentially 'glue' the appropriate Euclidean spacetime along $\mathcal{A}^c$, leaving boundary conditions $\phi_{\mathcal{A}}, \phi'_{\mathcal{A}}$ to be specified only in $\mathcal{A}$. For example, in the vacuum state,

$$\langle \phi'_{\mathcal{A}} | \rho_{\mathcal{A}} | \phi_{\mathcal{A}} \rangle = \int\limits_{\substack{\tau \in \mathbb{R} \\ \Phi(\mathcal{A}, \tau = 0_-) = \phi'_{\mathcal{A}} \\ \Phi(\mathcal{A}, \tau = 0_+) = \phi_{\mathcal{A}}}} \mathcal{D}\Phi \, e^{-S[\Phi]}, \tag{1.15}$$

where $\tau = 0_\pm$ specify the different boundary conditions for Euclidean times just before and after the 'cut' at $\tau = 0$. Similarly, in the thermal state this glues the Euclidean time into a circle, but only in $\mathcal{A}^c$, leaving boundary conditions to be specified in $\mathcal{A}$.

The trace of the qth power of the reduced density matrix is then the product of q copies of these matrix elements $\langle \phi_{\mathcal{A}}^{(i-1)} | \rho_{\mathcal{A}} | \phi_{\mathcal{A}}^{(i)} \rangle$ for $i = 1, \ldots, q$, with $\phi_{\mathcal{A}}^{(q)} = \phi_{\mathcal{A}}^{(0)}$, integrated over all field configurations for the $\phi_{\mathcal{A}}^{(i)}$. This process of identifying and then integrating over field configurations has the effect of creating a path integral where the q sheets are 'glued', with the top of region $\mathcal{A}$ on one sheet glued to the bottom of the region on the next, cyclically.

This process leaves no boundary components remaining, so what remains is just a path integral over some geometry depending on the details of the state and region $\mathcal{A}$. In particular, the resulting geometry will be singular, having a conical excess at the boundary of $\mathcal{A}$.

1.1.4 Universal results

Even with the replica trick it is still usually intractable to compute entanglement entropies. But some theories offer enough symmetry that calculations can be performed explicitly.

One such class of theories are conformal field theories (CFTs) in two dimensions, where there are several universal results depending only on the central charge c of the field theory (see [17] for a fuller review of these and other results). These are the entanglement entropy for a single interval of length l in the vacuum state, on an infinite line

$$S_{\mathcal{A}} = \frac{c}{3} \log \frac{l}{\epsilon} + \text{non-universal constant} \tag{1.16}$$

or on a circle of length L

$$S_{\mathcal{A}} = \frac{c}{3} \log \left(\frac{L}{\pi \epsilon} \sin \frac{\pi l}{L} \right) + \text{non-universal constant}, \tag{1.17}$$

and on a line at finite temperature $T = 1/\beta$

$$S_{\mathcal{A}} = \frac{c}{3} \log \left(\frac{\beta}{\pi \epsilon} \sinh \frac{\pi l}{\beta} \right) + \text{non-universal constant}, \tag{1.18}$$

where ϵ is a UV cutoff scale. These results are obtainable because the relevant surfaces for computing the Rényi entropies can be conformally mapped to a sphere. The result can then be computed from the conformal anomaly of this transformation, which is proportional to c.

More generally, the replication of the surface over which the path integral is computed can be shifted from the spacetime into the target space, by replicating the fields instead. This gives q copies of the fields on a single copy of the spacetime, with boundary conditions on these fields to implement the gluing construction. These boundary conditions can be implemented in 2D CFTs by the insertion of local operators at the endpoints of intervals, which act by cyclically permuting the q copies of the fields. These *twist operators* are in fact primary operators, of dimension $\frac{c}{12}(q - \frac{1}{q})$, which can be found from their behaviour under the conformal maps taking the multi-sheeted surface, used above for the universal results, to a sphere.

The upshot is that the Rényi entropies for n disjoint intervals can be computed from the $2n$-point correlation function of twist operators in the appropriate state. The results are no longer universal in the thermal state at finite size, since the correlation functions on the torus that this requires depend on the details of the theory. This is also true for more than one interval in vacuum, since four and higher point functions are not universal; equivalently, the Rényi entropies come from Riemann surfaces of nontrivial topology (of genus $(n-1)(q-1)$). Even with the simplifications of 2D CFT, there are still only very few results known.

The universal result for a CFT in vacuum can be extended to other dimensions in the case where $\mathcal{A}$ is a ball of radius R, since the replica geometries can be conformally mapped to a sphere [18]. When the dimension d is even, the universal piece, being the coefficient of the logarithmic divergence, is then found from the conformal anomaly, specifically the 'A-type' anomaly ($A = \frac{c}{12}$ for $d = 2$) proportional to the Euler density:

$$S_{\mathcal{A}}^{(\text{universal})} = (-1)^{\frac{d}{2}+1} 4A \log \frac{R}{\epsilon} \tag{1.19}$$

We move now to the other main ingredient of this thesis, holographic duality. This will, remarkably, enable us to compute entanglement entropies with relative ease in hitherto seemingly inaccessible regimes.

1.2 Holography

1.2.1 Quantum gravity is holographic

One of the earliest and most important results in quantum gravity was the realisation that black holes are thermodynamical objects. The earliest progress in this direction was made by Bekenstein [19], who observed that black holes appear to allow for the second law of thermodynamics to be violated (or 'transcended', to use Bekenstein's word). This can be achieved by dropping a box of entropic matter into the black hole, or lowering a hot body

near the horizon, allowing it to radiate into the black hole and raising it back up (now lighter) to turn heat into work. Any entropy behind the horizon is invisible to an external observer, because in equilibrium a black hole is determined entirely from its conserved charges (such a result is a 'no-hair theorem'). Motivated by this, Bekenstein suggested that the black hole itself should be assigned an entropy, and proposed a 'generalised second law', that the sum of ordinary and black hole entropy is nondecreasing. The natural candidate for this black hole entropy is proportional to the area of the event horizon, since the Hawking area theorem [20] shows that this is nondecreasing.

This was lent extra significance after Hawking's semiclassical field theory calculation [21] showing that black holes in fact radiate like a black body. They have thermodynamic properties associated with simple geometrical quantities: the area for the entropy, and the surface gravity for the temperature.

The generalised second law can be used to produce a bound for the maximal entropy that can be contained in a gravitational system. Roughly speaking, a zero-entropy shell can be collapsed onto any system in order to form a black hole, without allowing any radiation to escape in the process. The final entropy is then given by the area of the resulting black hole; since this must be larger than the initial entropy, this is bounded by the area within which it is contained[5].

As remarked earlier, the maximal entropy of a system provides a count of the number of degrees of freedom. Most systems are expected to be extensive, so the number of degrees of freedom scales with the volume. But for a gravitational system, this would for sufficiently large regions outstrip the area bound on the entropy: the state appears to be encoded somehow on the boundary of the system! This is the *holographic principle*.

This is clearly very mysterious and deep, and seems very hard to understand from semiclassical considerations alone (see [22, 23] for reviews). But there is one concrete proposal that explicitly implements the holographic principle, to which we now turn.

1.2.2 The paradigmatic duality

Gauge/gravity duality posits a remarkable equivalence between two seemingly disparate theories: a theory of quantum gravity, and a field theory in a lower dimension.

This duality was first conjectured by Maldacena [24], by considering the low energy dynamics of D-branes in string theory.

Specifically, consider a stack of N D3-branes in type IIB string theory, at small string coupling g_s (meaning $Ng_s \ll 1$), so the perturbative description of weakly coupled strings propagating in flat spacetime is valid. At low energy, the closed strings have only massless excitations, described by type IIB supergravity, and the open strings attached to the D-branes give a theory of maximally supersymmetric Yang-Mills theory on the 4-dimensional worldvolume of the branes describing their fluctuations; furthermore, these two sectors

[5]This is formulated more carefully, in a covariant way, by bounding the entropy that can pass through non-expanding light-sheets shone from any spatial surface by the area of that surface, the so-called *Bousso bound* [22].

decouple in the low energy limit.

In the opposite regime, when $g_sN \gg 1$, the branes interact with the closed string sector, and backreact on the geometry, being described by an extremal black brane. The low energy dynamics far from the brane stack is the same type IIB supergravity in flat 10-dimensional spacetime. In the near horizon region of the geometry, the redshift relative to infinity means that the full content of the string theory remains in the low-energy limit. Once again, this near-horizon physics decouples from the flat space supergravity.

So we have two very different low-energy descriptions of the physics: apart from the decoupled flat space supergravity in both pictures, one is a quantum field theory in four dimensions, and the other a string theory on the background of the near-horizon geometry. The duality follows if the low energy limit commutes with dialling the string coupling, so these are in fact just two languages for the same thing: a string theory equivalent to a quantum field theory!

The near-horizon geometry is a five-dimensional anti-de Sitter space (AdS_5) in the worldvolume and radial directions, times a 5-sphere in the transverse directions, with N units of flux from the self-dual 5-form Ramond-Ramond field strength F_5 under which the branes are charged. This gives us the statement of the duality:

$$\mathcal{N}{=}4\ SU(N) \text{ Yang-Mills } = \text{ IIB string theory on } AdS_5 \times S^5 \tag{1.20}$$

As a first check, we can match the symmetries of the two theories. The isometry group of AdS_5 is $SO(4,2)$, which happens to be the group of conformal transformations in 4 dimensions, which extends the Poincaré group by adding symmetry under changes of scale and other angle-preserving maps. The Yang-Mills theory is a conformal field theory (CFT), so possesses this symmetry; in particular there is no mass scale, and the β-functions all vanish, as a consequence of the scale invariance. On top of this is the $SO(6)$ symmetry of rotations of the S^5, equivalent to the $SU(4)$ R-symmetry group of the CFT, as well as supersymmetries.

Furthermore, we can match the parameters on each side: the Yang-Mills coupling g_{YM} and the rank of the gauge group N on one side, and the number of units of F_5 flux and the string coupling g_s on the other. This can be expressed simply in terms of the 't Hooft coupling $\lambda = g_{YM}^2 N$, and the string and Planck lengths ℓ_s, ℓ_p relative to the curvature scale ℓ of the AdS and S^5:

$$\frac{\ell_s}{\ell} = \lambda^{-\frac{1}{4}},\ \frac{\ell_p}{\ell} = N^{-\frac{1}{4}} \tag{1.21}$$

In particular, this shows why a theory of gravity, and an extra dimension, was never previously noticed lurking in a gauge theory. In the regime in which classical supergravity is valid, the coupling is large, so we are well outside the validity of perturbation theory. This is also one reason for the great utility of the duality: we can use classical gravity to do calculations in strongly coupled field theories. However, if we would like to translate results between the two pictures, we require more than just a statement of equivalence of the theories, but also a detailed matching of the observables.

So, what are the observables? In a field theory, everything is encoded in the partition function, with sources for each local operator turned on. But in a gravitational theory, turning on sources at given points in spacetime is not obviously a sensible thing to do, since the spacetime itself is dynamical, so will be altered by the presence of the sources. On the other hand, we do have the freedom to set the boundary conditions for the quantum gravity path integral. These boundary conditions are imposed at the timelike boundary of AdS, which is located at infinite proper distance, but reachable in finite time by null geodesics.

Given this, the natural proposal [25, 26] is to equate the partition functions of the two theories, with the field theory sources associated with the gravitational boundary conditions. Schematically,

$$\mathcal{Z}_{\text{gravity}}\left[\phi \rightsquigarrow \phi_0\right] = \langle e^{\mathcal{O}\cdot\phi_0}\rangle_{\text{CFT}} \tag{1.22}$$

where ϕ denotes the set of all bulk fields, with boundary values ϕ_0 in the appropriate well-defined sense, and $\mathcal{O}$ are the field theory operators. With this picture, the CFT can be understood as living at the boundary of the AdS spacetime, leading to the convenient terminology referring to the field theory as the 'boundary' theory, and the gravitational description as the 'bulk' theory.

This proposal requires a detailed matching between the operators of the field theory and the fields propagating in the gravitational theory, to associate the boundary conditions as the source of the correct operator. Most importantly, the defining feature of the gravitational theory, the dynamical metric, is associated with the energy-momentum operator; this is natural if one recalls that the energy-momentum tensor expectation value is nothing but the derivative of the partition function with respect to the background *boundary* metric, which will be the boundary condition of the *bulk* dynamical metric.

Of course, this seems unhelpful because we have no definition of the quantum gravity partition function[6]. But we may take the limit where the string theory is described by classical supergravity which, from eq. (1.21), corresponds to the planar limit[7] of the field theory at strong coupling. The upshot is remarkable, and worth repeating: we may do calculations in field theory at strong coupling, a regime to which we usually have no access, with just classical gravity.

1.2.3 How to interpret the extra dimension?

Having conjectured a precise matching between a field theory and gravity, it is still mysterious how this can be true. In particular, how does an extra dimension emerge from

[6]Indeed, one could regard this duality as defining a quantum gravity: type IIB string theory on asymptotically $AdS_5 \times S^5$ spacetimes is defined to be maximally supersymmetric Yang-Mills theory in four dimensions!

[7]This is the limit where N becomes large at fixed 't Hooft coupling λ, in which the perturbation theory in $1/N$ is controlled by the genus of the Feynman diagrams [27] (very reminiscently of string perturbation theory), being dominated by those which can be drawn on the plane without crossing propagators.

a field theory?

Notice that we associated the boundary conditions of fields with sources for operators in the field theory, or couplings. So it seems that the extra dimension is telling us that the couplings of the field theory depend not just on spacetime, but also on some extra parameter. But this is already a familiar situation from Wilsonian renormalisation: couplings of a theory depend on the energy scale at which we are making measurements. So it is natural to associate the additional bulk direction – the 'depth' in which we are inside AdS – with an energy scale in the field theory. The correspondence in effect is a geometrical realisation of the renormalisation group flow. What is striking is that the energy scale mixes in an entirely covariant way with the spacetime directions.

This picture works very nicely physically: if there is some high energy process deep within the bulk, the redshift of the spacetime will make it appear to be much lower energy out at the boundary. Similarly, very local processes on short length scales will appear to be much more diffuse and spread out from the perspective of an observer near the boundary.

Given this picture, in retrospect it is possible to imagine conjecturing gauge/gravity duality without string theory arguments. Since the renormalisation group flow is local in energy scale, it is natural to attempt to model this energy scale along with the spacetime directions. Somehow this extra dimension must manage to avoid contributing in the usual way to counting of states, so should have maximal entropy scaling with areas rather than volumes. But this is just the holographic principle for quantum gravity, so a gravitational theory is a natural candidate. For an approach using these sorts of arguments to give a convincing motivation for duality, see [28, 29].

Though we have focussed on a single concrete example, gauge/gravity duality should be much more general than this. Certain ingredients are essential: for example, a classical gravitational description requires a large number of degrees of freedom in a field theory to encode the extra dimension, and the field theory should be strongly coupled (there is certainly no extra dimension in free field theory). Others are helpful but may not be vital, like conformality and the associated AdS asymptotics, which are useful since it makes for the simplest beta functions, and supersymmetry to help keep a strongly coupled theory under control since the bounded below Hamiltonian guarantees stability. With this point of view, a natural approach is to try to avoid the complications of 'top-down' brane constructions and so forth, and to look for more universal features. The 'bottom-up' philosophy is to pick the minimal field content in a gravitational theory to give the aspects and behaviour of a field theory that one would like to model, for example a bulk electromagnetic field to give a conserved charge, and calculate in a simple model (with probably no field theory dual) in the hope that the results describe universal features.

1.3 Holographic entanglement entropy

Though gauge-gravity duality provides a powerful way to compute many quantities in otherwise inaccessible regimes, it is far from obvious that it should be helpful for studying

entanglement. But it turns out that not only is entanglement entropy computable using the duality, but it is associated with one of the most geometrically natural objects one could think of: an extremal surface.

Let us begin with a holographic field theory in some static state, like the vacuum, and suppose we would like to calculate the entanglement entropy associated with some spatial region $\mathcal{A}$ at some fixed time[8], say $t = 0$. Now assume that this state is described by a classical static geometry in the dual gravitational description, which, in particular, has a well defined extension of the $t = 0$ hypersurface of constant time into the bulk. Then the proposal of Ryu and Takayanagi [30, 31] states that the entanglement entropy is given by the area of a minimal surface $\mathcal{E}$ on this constant time slice, whose boundary coincides with that of $\mathcal{A}$:

$$S(\mathcal{A}) = \frac{1}{4G_N}\text{Area}(\mathcal{E}),\ \text{minimising Area}(\mathcal{E})\ \text{with}\ \partial\mathcal{E} = \partial\mathcal{A} \tag{1.23}$$

This is very reminiscent of the Bekenstein formula for black hole entropy, which is given by $\frac{1}{4G_N}$ times the area of the bifurcation surface of the event horizon.

1.3.1 Consistency checks

As a first motivation, we can check that this is consistent with the various known properties of entanglement entropy, for example those in section 1.1.1, and that it reproduces the expected results when there are other available methods for computing it. In simple cases, in this geometric description the required properties are often geometrically obvious or tautological, in contrast to the direct quantum mechanical proofs, which can be quite involved.

For example, the equality of entanglement entropy for a subsystem $\mathcal{A}$ and its complement $\mathcal{A}^c$ is immediate, because the boundary of a region is the same as the boundary of its complement ($\partial\mathcal{A} = \partial\mathcal{A}^c$), so the minimal surfaces associated with $\mathcal{A}$ and $\mathcal{A}^c$ are identical. This should be modified in a mixed state, such as a thermal ensemble, which indicates that the prescription needs some alteration. A natural guess for the correct statement is to implement a restriction on the topology of the allowed surfaces, specifying that $\mathcal{E}$ must be homologous to $\mathcal{A}$ in an appropriate sense. This has an effect in the thermal state when there is a black hole present in the geometry, because the topology of the $t = 0$ slice is then nontrivial. Some aspects of this are the subject of chapters 2 and 5, so more in depth discussion is deferred to those chapters.

There is a similar simple geometric proof of the strong subadditivity of entropy eq. (1.5), which specialises to similar proofs of many of the various other entropy inequalities. The combination $S(\mathcal{AB}) + S(\mathcal{BC})$ is computed by the sum of the areas of one surface bounded by $\partial(\mathcal{A}\cup\mathcal{B})$, and one bounded by $\partial(\mathcal{B}\cup\mathcal{C})$. But this pair of surfaces can be rearranged, as in fig. 1.3, to another pair of surfaces (usually with some kinks), bounded by $\partial(\mathcal{A}\cup\mathcal{B}\cup\mathcal{C})$

[8]More generally, one can take a time-reversal invariant state at the moment of time-reflection symmetry, having a geometric bulk description which is also time-reversal invariant, with some hypersurface fixed by the reflection.

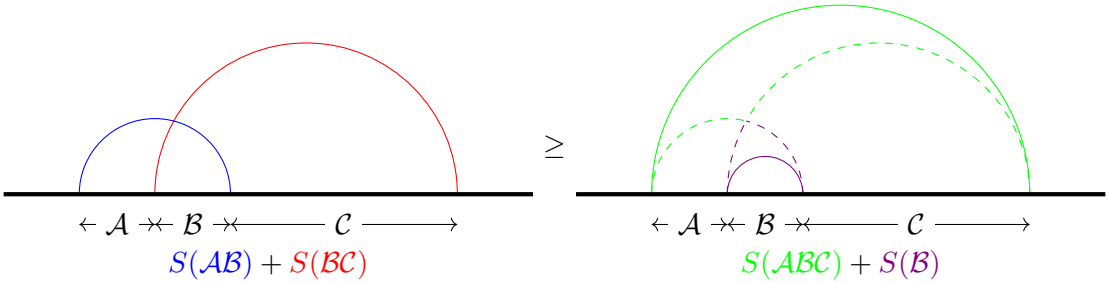


Figure 1.3: Graphical proof of strong subadditivity in a simple case. The thick black line is the boundary, and the bulk lies above it. The blue and red curves giving $S(\mathcal{AB}) + S(\mathcal{BC})$ can be adjusted to the dashed curves in the right figure, which are longer than the green and purple curves giving $S(\mathcal{ABC}) + S(\mathcal{B})$.

and by $\partial\mathcal{B}$. These will, by definition, have greater area than the minimal surfaces with the same boundaries that compute $S(\mathcal{ABC}) + S(\mathcal{B})$, which proves the desired inequality. This can be formulated more carefully, paying attention to the condition on the topology of the surfaces, with relative ease [32]. There are also quite generic circumstances where the inequalities can be saturated, with consequences for the structure of entanglement in the field theory, at least of the colour degrees of freedom which give the leading order N^2 contribution.

The RT prescription in fact implies additional inequalities for entanglement entropies that do not hold in general, but are special for holographic systems. Chief amongst these is the 'monogamy of mutual information' [33], which can be stated as the negativity of a certain measure of multipartite entanglement, the tripartite information:

$$I_3(\mathcal{A},\mathcal{B},\mathcal{C}) := S(\mathcal{A}) + S(\mathcal{B}) + S(\mathcal{C}) - S(\mathcal{BC}) - S(\mathcal{CA}) - S(\mathcal{AB}) + S(\mathcal{ABC}) \leq 0 \quad (1.24)$$

Such inequalities are not contradictory, but limit the sort of correlations that can be present in a holographic theory.

The properties specific to field theory are also relatively straightforward to understand from the prescription. Most important among these is the structure of the divergence of the entanglement entropy, and specifically the area law. From the geometry, the fact that the boundary is in fact at infinite proper distance from any point in the bulk leads to an infrared divergence of the area. The area is regulated by cutting off the geometry at some large but finite distance, and systematically taking that distance to infinity; this is dual to taking a UV cutoff to zero in the field theory. The IR divergence of the area and the UV divergence of the entanglement entropy match, and in particular the leading divergent contribution to $S(\mathcal{A})$ is proportional to the area of $\partial\mathcal{A}$; this is clear since, near to the boundary, the surface $\mathcal{E}$ looks like $\partial\mathcal{A}$ times an interval in the holographic radial direction.

In the cases that can be explicitly checked by computations on both sides of the duality, there is exact agreement. The simplest of these are the universal results eqs. (1.16) to (1.18) for two-dimensional CFTs, in which case the extremal surfaces $\mathcal{E}$ are just geodesics, and the central charge of the CFT is related to the bulk curvature scale ℓ_{AdS} and Planck length

G_N by $c = \frac{3\ell_{AdS}}{2G_N}$. The results are reproduced by the regularised lengths of geodesics in the relevant geometries, being planar and global AdS (whose $t = 0$ slices are both hyperbolic space, with different regularisations, natural in half-plane and disc models respectively), and the planar BTZ black brane.

1.3.2 Further evidence

Beyond these various consistency checks, there has more recently been progress towards a derivation of the Ryu-Takayanagi formula directly from the usual AdS/CFT rules. This was first done in 2D CFTs in the vacuum state, by a close matching between calculations of the entanglement entropy for multiple intervals in field theory [34] and gravity [35], using the replica trick. The CFT calculation proceeds by expressing the Rényi entropies as correlation functions of twist operators, as briefly described in section 1.1.4, and expanding these in some channel using the OPEs. The crucial new ingredient is the assumption that the CFT has a large central charge, and few light operators in a precise sense, conditions which are expected of a theory with a holographic dual. The consequence of this assumption is that the sum over operators in the OPE is dominated, when the operators in question are sufficiently close, by the identity operator and its descendants. The result can then be computed by a procedure identical to a computation coming from 3D gravity. The different possible choices of geodesics then correspond to the different possible OPE channels.

The strongest argument comes from a more general calculation [36] which extends the replica trick into the bulk, for arbitrary time-reflection symmetric geometries in any dimension, under certain assumptions. This is reviewed in detail in chapter 5, in which its relationship with the topological restriction on surfaces is discussed. It also provides the clearest way to motivate appropriate extensions of the original proposal, for example generalising the prescription to field theories with higher derivative gravitational duals [37, 38], in which case some more general functional of the surface $\mathcal{E}$ than the area must be minimised.

1.3.3 A covariant generalisation

The Ryu-Takayanagi proposal has one obvious drawback: it applies only to a very limited set of states, being static or at least invariant under time reversal, so we would like to have a generalisation that circumvents this limitation. In a Lorentzian geometry, the area of a surface is not bounded below, since making the surface nearly lightlike will make the area arbitrarily small, so minimisation is not useful. The correct prescription to generalise a surface of minimal area, the Hubeny-Rangamani-Takayanagi (HRT) proposal [39], is to find a surface of *extremal* area, so the area functional is stationary under all variations of the surface.

More explicitly, take some boundary Cauchy surface, on which the state of the field theory is described by some density matrix ρ, and suppose we would like to compute the entanglement entropy associated with some spatial subregion $\mathcal{A}$ of the Cauchy surface. The

HRT prescription then states that the entanglement entropy is given by (in Planck units) the area of a codimension two extremal surface $\mathcal{E}$ in the bulk, ending on the boundary of the spatial region $\mathcal{A}$, so $\partial\mathcal{E} = \partial\mathcal{A}$. If there are several such surfaces, then take the smallest area (again, with a restriction on topology of $\mathcal{E}$).

There are a number of equivalent reformulations of the extremality condition, useful in different circumstances. The most important equivalent definitions of $\mathcal{E}$ are the following:

Extremum $\mathcal{E}$ is a stationary point of the area functional

Vanishing null expansions Sheets of light shone out from $\mathcal{E}$ in the perpendicular directions initially neither expand nor contract

Maximin To define $\mathcal{E}$, first find a minimal surface on every possible *AdS*-Cauchy surface[9] Σ, and then maximise the area of this minimal surface varying over the set of all Σ.

The fact that the result is no longer the minimum of a functional makes many of the required properties for an entanglement entropy far less clear. Even existence of a surface $\mathcal{E}$ is not guaranteed *a priori*, and the simple proofs of inequalities such as given in fig. 1.3 are no longer applicable. Nonetheless, it is still possible to show, under certain assumptions about the spacetime in question, that these properties hold [40], using in particular the maximin construction. A new ingredient in the proofs is a physical restriction on the spacetime for the properties to hold, specifically the null energy condition.

The time dependence brings additional checks that must be made for consistency with known field theory properties of entanglement entropy. Perhaps the most obvious of these is the causality of evolution; in particular, in a relativistic field theory, disturbances can propagate no faster than the speed of light so sources at some point should not effect the results of measurements outside their future lightcone. Causality of HRT in this sense has been shown to be upheld as long as the null energy condition is satisfied [41].

The topological constraint on surfaces in time-dependent situations is more subtle than the static case, in particular requiring some causal aspect for consistency. The most natural form this should take is the requirement that the extremal surface $\mathcal{E}$ and the region in question $\mathcal{A}$ must bound some codimension one *spacelike* surface. Some of these issues are briefly addressed in chapter 2.

A derivation of the HRT proposal paralleling that of RT remains elusive, so there is rather less evidence for this more general conjecture. The consistency checks described here, and others, are nontrivial and should be regarded as strong support, along with myriad applications that have always yielded sensible results. An idea for extending the Euclidean quantum gravity calculation to apply to HRT in some very limited circumstances are discussed in chapter 4, but a more general derivation is an important open problem that may help to clarify the outstanding issues.

[9]Here, by an *AdS*-Cauchy surface I mean a set intersected exactly once by all inextensible causal curves in the conformal compactification of the spacetime, differing from a true Cauchy surface since causal curves may reach the boundary, and so escape from the spacetime.

1.3.4 Beyond the classical limit

The geometric proposals of RT and HRT apply only in the classical limit of the gravitational theory, when G_N is much larger than the other length scales in the problem. Away from this limit, quantum corrections will alter the result, which can be described perturbatively as a loop expansion in powers of G_N, the classical result being of order G_N^{-1}. This is important, for example, when certain inequalities are saturated in the classical limit, since this saturation will generically be lifted by order one corrections.

There are circumstances in which the quantum corrections can be explicitly calculated, for example by performing the one loop calculation in three dimensional pure gravity [42]. A more general picture for the one loop correction [43] is that it is given essentially by the bulk entanglement entropy between the inside and outside of $\mathcal{E}$, and it is conjectured that the higher order corrections can be characterised as an entropy of a 'quantum extremal surface' [44].

Apart from the perturbative corrections, there should also be nonperturbative 'instanton' corrections coming from subdominant classical solutions. This raises a particularly tricky problem when considering the replica trick, since the contributions of competing saddles in the partition function become comparable when the replica number is parametrically close to one. The upshot is that the classical limit and the replica limit used to obtain the von Neumann entropy seem to fail to commute, which leads to a contradictory result. A similar phenomenon occurs in the CFT computation of [34]. The resolution of this apparent paradox is currently an open problem.

Chapter 2

Holographic entanglement in systems of finite size

This chapter is a reproduction of the paper[1] *Holographic entanglement plateaux* [1], written in collaboration with Veronika Hubeny, Mukund Rangamani and Erik Tonni.

The focus of the paper is the entanglement entropy in holographic systems of finite size in a thermal state, when the temperature is above the Hawking-Page transition so the bulk geometry contains a black hole. In such systems, there is a phase transition in the entanglement entropy as the size of the region $\mathcal{A}$ is altered, where the minimal surface $\mathcal{E}$ computing $S(\mathcal{A})$ becomes disconnected, with one piece being the event horizon. As a consequence, the entanglement entropy satisfies $S(\mathcal{A}) = S(\mathcal{A}^c) + S$, where S is the thermal entropy, given by the horizon area, giving a plateau in the quantity $S(\mathcal{A}^c) - S(\mathcal{A})$. This is a saturation of the Araki-Lieb inequality, as described in section 1.1.1. We compared the result to a simple 2D CFT, where no phase transition, and indeed no crossover or other feature was observed analogous to the plateau.

The picture we gave provided a geometric interpretation and proof in holography for this inequality. Recall that the Araki-Lieb inequality follows from strong subadditivity, but only if one assumes the existence of a purifying system. For the holographic proof of strong subadditivity to be applicable to prove Araki-Lieb, we therefore further require that the purification can be described *geometrically*, for example by the thermofield double by considering a maximally extended eternal black hole. It would be interesting to know if there are relevant mixed states with a geometric description, but no 'geometric purification' in an appropriate sense.

Part of the discussion is a more general question about the smoothness of $S(\mathcal{A})$ as the region $\mathcal{A}$ is smoothly varied. First order phase transitions, like the one discussed in the paper, lead to a kink in this quantity. In systems of finite size, it is usually impossible to have a sharp phase transition, but the classical gravity description means that there

[1]The use of the word 'plateau' in the title relates to flat portions of the plots of the main quantity of interest. The association between minimal surfaces and the name of Joseph Plateau was unknown to me at this point, and, as far as I know, the other authors, so the overlap in language is merely a coincidence

is a thermodynamic limit, implemented by the number of degrees of freedom in the field theory (central charge c, or gauge group rank N, for example) going to infinity. Quantum corrections should smooth out such kinks. From Ryu-Takayanagi, we gave a simple argument that this is the worst that can happen; the entanglement entropy may not jump discontinuously. There is still no generalisation of this argument to dynamical situations and HRT, though no discontinuity has, as far as I am aware, ever been observed.

A possibly surprising finding was that in more than three bulk dimensions, the phase transition is inevitable since there is in fact no connected minimal surface for certain boundary regions $\mathcal{A}$. This is closely related to a simpler example of minimal surfaces between two rings in ordinary three-dimensional space, forming hyperboloids, which for large enough separations do not exist. We gave one argument for nonexistence of connected surfaces, which illustrates that this property is generic and not a result of symmetry or similar, by looking at the properties of the causal wedge. The causal wedge of $\mathcal{A}$ firstly excludes minimal surfaces, and secondly develops holes around black holes, which forces extremal surfaces to become disconnected to satisfy the homology constraint.

Related to this, we also found that there are an infinite number of minimal surfaces ending on a given region $\mathcal{A}$, wrapping closely round the event horizon a number of times before heading out to the boundary. These have an intricate self-similar structure, analysed from an analytic computation in the near horizon. We subsequently realised that these are in fact all unstable, in the sense that there are small perturbations that decrease the area (quadratically in the size of the perturbation). This can be argued from Morse theory ideas[2]. This instability is even present in the simple hyperboloid problem: in essence, a hyperboloid with a sufficiently thin neck reduces its area by contracting and pinching off, or by receding and becoming wider.

We also introduced a new (in the gravitational context) tool for computing minimal surfaces, using the idea of mean curvature flow. This is a geometrically very natural construction, that has been important in the mathematical theory of minimal surfaces, by which a general surface will flow with decreasing area, to ultimately settle at a minimal surface. The algorithm is possible to implement in any dimension for codimension-one surfaces, and one particular virtue, that is not shared by finite element methods, for example, is that it allows automatically for the sort of changes in topology that are important in the present setting.

We also discussed some examples where one has to be careful with the homology constraint on extremal surfaces, which show that the HRT prescription requires an additional constraint related to causality.

[2]I would like to thank Matt Headrick for illuminating discussions of this issue.

2.1 Introduction

Entanglement is one of the most non-classical features of quantum mechanics and much effort has been expended in trying to figure out a quantitative measure for the amount of entanglement inherent in a quantum state. In local quantum field theories one such measure is provided by the entanglement entropy $S_{\mathcal{A}}$ associated with a specified (spacelike) region $\mathcal{A}$ located on a Cauchy slice Σ. This quantity is defined to be the von Neumann entropy of the reduced density matrix $\rho_{\mathcal{A}}$ obtained by integrating out the degrees of freedom in the complementary region $\mathcal{A}^c$. Famously the entanglement entropy is UV divergent, with the leading divergence being given by the area of the entangling surface $\partial\mathcal{A}$ (the boundary of $\mathcal{A}$). However, the finite part of the entanglement entropy contains non-trivial information about the quantum state and in certain cases serves as a novel order parameter. As usual, it is this finite part which we shall consider.

Our interest is in understanding the behaviour of entanglement entropy for field theories defined on compact spatial geometries; so we take Σ to be a compact Riemannian manifold (typically $\mathbf{S}^{d-1}$). One of the questions which we wish to address is whether the entanglement entropy for a fixed state of the field on Σ is a smooth function of the (geometrical attributes of the) region $\mathcal{A}$. We envisage considering a family of *smooth* regions $\mathcal{A}_\alpha$ characterized by some parameter α, which is a proxy for the relative size of $\mathcal{A} \subset \Sigma$. The main question we want to ask is whether $S_{\mathcal{A}}(\alpha)$ is smooth under changing α.

For finite systems, analogy with statistical mechanics suggests that this would indeed be true. There should be no room for any discontinuity in $S_{\mathcal{A}}(\alpha)$ since the reduced density matrix will change analytically with α. The place where this is expected to break down is when we attain some analog of the thermodynamic limit, i.e., when the number of degrees of freedom involved gets large. So a natural place to look for non-smooth behaviour is in the dynamics of large N field theories on compact spacetimes in the planar limit. This naturally motivates the study of field theories which can be captured holographically by gravitational dynamics using the gauge/gravity correspondence. We will refer to these holographic field theories as large c (central charge) theories; for conventional non-abelian gauge theories $c \sim N^2$.

In the past few years much progress has been made in understanding entanglement entropy in large N field theories (at strong coupling) thanks to the seminal work of Ryu & Takayanagi [RT] [30, 31], who gave a very simple geometric prescription for computing $S_{\mathcal{A}}$ for static states in terms of the area of a bulk minimal surface anchored on $\partial\mathcal{A}$. This prescription was extended by Hubeny, Rangamani, and Takayanagi [HRT] to arbitrary time-dependent states in [39] where one considers extremal surfaces, which can also be related to light-sheets discussed in covariant entropy bound context [22]. There have been many attempts to derive the prescription from first principles; the first was made in [45] which was critically examined in [46]. More recently, [18] gave a nice argument deriving the prescription for a special class of states (conformally invariant vacuum) and spherically symmetric regions, by converting the reduced density matrix $\rho_{\mathcal{A}}$ to a thermal density

matrix. A local version of this argument has been made recently in [36] and together with the results of [34, 35] goes quite a way in establishing the holographic prescription of [30] for static states. Given these developments, it is apposite to take stock of the implications of the holographic entanglement entropy prescription for the question we have in mind.

Let us first record some basic facts about $S_{\mathcal{A}}$ which we will use to investigate the behaviour of $S_{\mathcal{A}}(\alpha)$. It is a well known fact that when the total state of the field theory is pure, the entanglement entropy of a given region and its complement are the same: $S_{\mathcal{A}} = S_{\mathcal{A}^c}$. This however ceases to be true when the entire system is itself in a density matrix ρ_Σ. To measure the deviation from purity of the system we could monitor the difference $\delta S_{\mathcal{A}} = S_{\mathcal{A}} - S_{\mathcal{A}^c}$. From our perspective this quantity has some useful advantages. Firstly, it is finite since the divergent contributions which are given in terms of intrinsic and extrinsic geometry of the entangling surface $\partial\mathcal{A}$ cancel. Secondly, it is bounded from above by the von Neumann entropy of the entire density matrix ρ_Σ, by the Araki-Lieb inequality [10]:

$$| \, \delta S_{\mathcal{A}} \, | \equiv | \, S_{\mathcal{A}} - S_{\mathcal{A}^c} \, | \leq \; S_{\mathcal{A}\cup\mathcal{A}^c} = S_{\rho_\Sigma} \, . \tag{2.1}$$

So one way to phrase our original question is to ask whether $\delta S_{\mathcal{A}}(\alpha)$ is a smooth function of $\alpha \in [0,1]$ which we take to be a suitable function of the ratio of $\mathrm{Vol}(\mathcal{A})/\mathrm{Vol}(\Sigma)$ such that $\delta S_{\mathcal{A}}$ is an odd-function around $\alpha = \frac{1}{2}$ and becomes the total entropy $\pm S_{\rho_\Sigma}$ at $\alpha = 0, 1$. The issue we want to focus on is whether there is any discontinuity either in $S_{\mathcal{A}}(\alpha)$ or its derivatives as we vary α. We will argue here that for large c field theories $S_{\mathcal{A}}(\alpha)$ (given by the RT prescription) has to be continuous for time-independent (static) density matrices ρ_Σ. However, there can be non-trivial behaviour in $\partial_\alpha S_{\mathcal{A}}(\alpha)$: we will exhibit explicit examples where the function $S_{\mathcal{A}}(\alpha)$ is continuous but fails to be differentiable. [3] It is important to distinguish this from situations where the total density matrix itself varies; one can certainly have entanglement entropy of a fixed-size region which is discontinuous e.g. as a function of temperature [49, 50], as is quite familiar from the simpler example of Hawking-Page transition. Here we fix the total density matrix (which has the bulk equivalent of fixing the spacetime) and consider the entanglement entropy as function of the region.

To understand the potential issues, we need to explain one key feature of the holographic prescription of [30, 39]. To compute $S_{\mathcal{A}}$, we find extremal surfaces $\mathcal{E}_{\mathcal{A}}$ in the bulk spacetime $\mathcal{M}$ which are anchored on $\partial\mathcal{A}$; in the asymptotically AdS spacetimes we consider, we demand that $\partial\mathcal{E}_{\mathcal{A}} = \partial\mathcal{A}$. However, there can be multiple such surfaces in a given spacetime. We are instructed to restrict attention to extremal surfaces $\mathcal{E}_{\mathcal{A}}$ which are homologous to the boundary region $\mathcal{A}$ under consideration [32] and from the set of such surfaces pick the

[3] Earlier studies of holographic entanglement entropy have indeed revealed examples where $S(\alpha)$ undergoes a first order phase transition, i.e., $\partial_\alpha S(\alpha)$ is discontinuous due to change in the nature of the minimal surfaces [47, 48]. For example the recent analysis of [13] exhibits this phenomenon arising due to the minimal surface changing topology (even in causally trivial spacetimes) as we vary α.

one with smallest area. To wit,

$$S_{\mathcal{A}} = \min_{X} \frac{\text{Area}(\mathcal{E}_{\mathcal{A}})}{4\,G_N}\,, \qquad X = \mathcal{E} : \begin{cases} \partial\mathcal{E}_{\mathcal{A}} \equiv \mathcal{E}_{\mathcal{A}}\big|_{\partial\mathcal{M}} = \partial\mathcal{A} \\ \exists\, \mathcal{R} \subset \mathcal{M} : \partial\mathcal{R} = \mathcal{E}_{\mathcal{A}} \cup \mathcal{A} \end{cases} \tag{2.2}$$

where the region $\mathcal{R}$ is a bulk co-dimension one smooth surface (in the $d+1$ dimensional spacetime $\mathcal{M}$) which is bounded by the extremal surface $\mathcal{E}_{\mathcal{A}}$ and the region $\mathcal{A}$ on the boundary.

It was appreciated already in [32] that the homology constraint is crucial for the Araki-Lieb inequality to be satisfied. The issue was further elaborated in [51] where 1+1 dimensional CFTs on a torus were considered (see also [52] for a recent discussion). In particular, its effects are most acutely felt when ρ_Σ is a density matrix, for then we anticipate the bulk spacetime to have a horizon [53].[4] In static spacetimes this implies a non-trivial topology in the bulk when restricted to a constant time slice, which can be easily intuited by considering the Euclidean section of the geometry.

The homology constraint, being non-local from the boundary perspective, allows non-trivial behaviour in the nature of the extremal surfaces that are admissible for the problem. Indeed as we will show explicitly, there are many simple examples where one has multiple extremal surfaces and only some of them are homologous to the boundary region in question. In fact, the most bizarre aspect of our analysis is that for certain choices of boundary regions there are no connected minimal surfaces satisfying the homology constraint: one is forced into considering disconnected surfaces.[5] Multiply connected extremal surfaces in turn imply that one can have novel behaviour in $S_{\mathcal{A}}$ or equivalently in $\delta S_{\mathcal{A}}$; we go on to show that in static spacetimes these indicate that $\delta S_{\mathcal{A}}(\alpha)$ is continuous but not differentiable, exhibiting explicit examples involving global AdS black hole geometries. We argue that the lack of differentiability is the worst it gets for $\delta S_{\mathcal{A}}$ in static spacetimes, proving that $\delta S_{\mathcal{A}}$ has to be a continuous function of α. Furthermore, when we are forced onto the branch of disconnected extremal surfaces, it is easy to establish that $\delta S_{\mathcal{A}} = S_{\rho_\Sigma}$, i.e., the Araki-Lieb inequality is saturated; we refer to this phenomenon as *entanglement plateau*.[6]

While in simple examples one can establish entanglement plateaux by explicit construction, it is interesting to examine when it should happen on general grounds. Curiously, it is easy to provide a bound, though in a somewhat roundabout manner using machinery outside the extremal surface technology. The necessary concept is geometric and has to do with the behaviour of the causal wedge associated with $\mathcal{A}$. These objects were studied in

[4] This statement presumes that the von Neumann entropy of ρ_Σ scales like the central charge of the field theory.

[5] The exchange of dominance between connected and disconnected surfaces in confining backgrounds for field theories on non-compact geometries $\mathbb{R}^{1,d-2} \times \mathbf{S}^1$ has been well studied in the context of holographic entanglement entropy, cf., [47, 48] for initial work. We will be however be considering field theories on compact spatial volumes.

[6] This phenomenon like many others encountered in holographic duals is a feature of large c theories. At finite central charge we cannot have any sharp plateaux; we thank Hong Liu for discussions on this issue.

the context of 'causal holographic information' in [54] from a very different perspective (the motivation being to characterize the minimal amount of holographic information in the reduced density matrix). Using the topology of the causal wedge one can establish criteria for when the extremal surfaces $\mathcal{E}_{\mathcal{A}}$ become disconnected. The precise statement and its proof will appear elsewhere [55], but we will flesh out the physical aspects of the argument in what follows. Suffice to say for now that we find the interplay between causality and extremal surfaces extremely intriguing and believe that it points to some yet to be fathomed facet of holography.

2.2 Generic behaviour of holographic $S_{\mathcal{A}}(\alpha)$

We begin our discussion by explaining the continuity of $S_{\mathcal{A}}(\alpha)$. Importantly, we focus on the RT prescription for static configurations, where the entire problem can be formulated on a d dimensional Riemannian bulk geometry with $(d-1)$-dimensional boundary (in the conformal class of) Σ. Consider a family of boundary regions $\mathcal{A}_\alpha \subseteq \Sigma$ specified by a real number α. For example, we can fix the shape of $\mathcal{A}$ and let α denote the overall size; in the simplest case of $(d-2)$-spherically symmetric $\mathcal{A}$ on a spatial slice of ESU_d boundary spacetime, we take $\alpha \in [0,1]$ to be the fractional volume of the system.[7]

Now, consider the function $S(\alpha)$ defined by a *minimization* of area over all smooth bulk surfaces homologous to $\mathcal{A}_\alpha$. Starting from any surface homologous to $\mathcal{A}_\alpha$, we may allow it to relax, continuously decreasing the area, such as by the mean curvature flow described in Appendix 2.A. Since the area is bounded from below, it must tend to a limit, of which one quarter in Planck units is given by $S_k(\alpha)$. This may take several values, labeled by k, depending on the initial surface chosen, corresponding to different local minima of the area functional. Generically, the surface itself will also converge to an endpoint[8], which will be a corresponding minimal surface $\mathfrak{M}^k_\alpha$. As α is varied, we expect these minima to come in a discrete set of families depending smoothly on some parameter (since $\mathcal{A}_\alpha$ and the spacetime are smooth), giving a set of curves in the plane of α and area. This means we have a set of functions $S_k(\alpha)$, defined on some interval of α, continuous, and smooth away from critical points of α, with $S(\alpha) = \min_k S_k(\alpha)$.

We now wish to consider the behaviour of $S(\alpha) \equiv \min_k S_k(\alpha)$. In simple spacetimes, such as pure AdS and small deformations thereof, there is only a single family $k = 1$ which covers the full range of α, so in such cases $S(\alpha)$ is manifestly smooth. However, it may happen at some point that two families, $k = 1, 2$, exchange dominance such that $S(\alpha) = S_1(\alpha) < S_2(\alpha)$ for α smaller than some critical value $\alpha_{\mathcal{X}}$, and $S(\alpha) = S_2(\alpha) < S_1(\alpha)$ for $\alpha > \alpha_{\mathcal{X}}$. At $\alpha = \alpha_{\mathcal{X}}$, $S_1 = S_2 = S$, so $S(\alpha)$ is necessarily continuous. However, there will generically be a discontinuity in $\partial_\alpha S(\alpha)$, so $S(\alpha)$ has a kink.

[7] With the standard $SO(d-1)$ symmetric metric on spatial sections of ESU_d, i.e., $ds^2_\Sigma = d\theta^2 + \sin^2\theta\, d\Omega^2_{d-2}$, we have $\alpha = \int_0^{\theta_\infty} (\sin\theta)^{d-2} d\theta / \int_0^{\pi} (\sin\theta)^{d-2} d\theta = \frac{1}{2} - \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\,\Gamma(\frac{d-1}{2})} \cos\theta_\infty \, {}_2F_1\left(\frac{1}{2}, \frac{3-d}{2}, \frac{3}{2}, \cos^2\theta_\infty\right)$ for a polar cap characterized by co-latitude θ_∞.

[8] For some special spacetimes, such as extremal black holes with an infinite throat, this need not be strictly true, though it does not materially affect the argument.

The above argument indicates that within the RT prescription, the entanglement entropy of a given region $\mathcal{A}$ should vary continuously with the parameters specifying the geometrical attributes of $\mathcal{A}$. In the following two sections, we will see this behaviour realized manifestly, even in situations where new families of minimal surfaces get nucleated at some intermediate α. By examining these families more closely, we will identify examples with large multiplicities of minimal surfaces.

2.3 Entanglement entropy in $1+1$ dimensional CFTs

In the previous section we have argued that for the holographic entanglement entropy given by a minimization procedure, $S_{\mathcal{A}}(\alpha)$ must be continuous but need not be differentiable. The lack of differentiability is typically associated with two families of minimal surfaces exchanging dominance. Such an occurrence is not new, and good examples already exist in the literature. The simplest one occurs for the bulk spacetime being the BTZ black hole, where this point was appreciated already in the early days [30, 32, 51] and fleshed out a bit more explicitly in [52]. We quickly review this story to illustrate the contrast with our other examples in higher dimensions.

The metric for the BTZ black hole is given by[9]

$$ds^2 = -f(r)\,dt^2 + \frac{dr^2}{f(r)} + r^2\,d\theta^2\,, \qquad f(r) = r^2 - r_+^2\ . \tag{2.3}$$

It is a simple matter to find the minimal surfaces for regions $\mathcal{A} = \{\theta : |\theta| \leq \theta_\infty\}$ since these are given by spacelike geodesics. The result is best described by writing down the spatial projection of the surfaces [54]

$$\mathfrak{M}_1(\theta_\infty): \qquad \left\{(r,\theta):\ r = \gamma(\theta,\theta_\infty,r_+) \equiv r_+\left(1 - \frac{\cosh^2(r_+\,\theta)}{\cosh^2(r_+\,\theta_\infty)}\right)^{-\frac{1}{2}}\right\} \tag{2.4}$$

and is plotted in fig. 2.1 for large and small black holes, for a set of $\theta_\infty \in [0,\pi]$.

As can be easily seen in fig. 2.1, connected spacelike geodesics (satisfying the homology constraint) always exist for any θ_∞ and r_+. This makes sense, since there is no reason for the geodesics to break up (in fact, spacelike geodesics can orbit the black hole arbitrarily many times before returning to the boundary, albeit at the expense of greater length). As we will see in section 2.4, this is in stark contrast to the behaviour of co-dimension 2 surfaces in higher dimensional black hole spacetimes. Also note that in the BTZ case, another effect of the low dimensionality is that arbitrarily small black hole ($r_+ \to 0$ in eq. (2.3)) always looks effectively large in terms of the effect it has on geodesics.

Let us now consider the proper length along these geodesics and compute the entanglement entropy. As is well known, this computation reproduces the CFT computation of Cardy-Calabrese [56] for thermal CFT on the infinite line. Naive application of the RT

[9] We work in units where the AdS length $\ell_{\rm AdS} = 1$ and also set the radius of the boundary circle R parameterized by $\theta \in [0, 2\pi]$ to unity. It is easy to restore dimensions when necessary as we illustrate later.

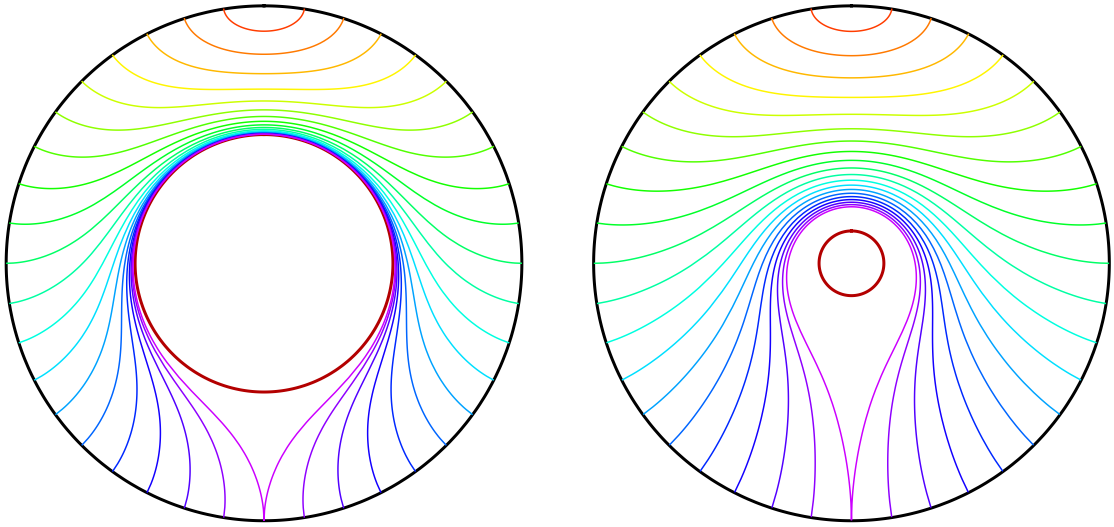


Figure 2.1: Minimal surfaces (geodesics) in BTZ geometry with black hole size $r_+ = 1$ (left) and $r_+ = 0.2$ (right). In each panel, the thick black circle represents the AdS boundary, and the thick red circle the black hole horizon. The minimal surfaces are depicted by the thin curves, color-coded by θ_∞, with θ_∞ varying between 0 (red) and π (purple) in increments of 0.05π. For ease of visualization we use a compactified radial coordinate $\rho = \tan^{-1} r$. The boundary region $\mathcal{A}$ is centered at $\theta = 0$ which is plotted on top.

result for the CFT on a cylinder leads to[10]

$$(S_{\mathcal{A}})_{\text{naive}} = \frac{c}{3} \log\left(\frac{2\,r_\infty}{r_+}\,\sinh(r_+\,\theta_\infty)\right)\,, \qquad \forall\ \theta_\infty \in [0,\pi] \tag{2.5}$$

where r_∞ is a radial cut-off to regulate the proper length of the geodesic and is our proxy for the boundary UV cut-off scale. It is then a simple matter to check that the naive result for $\delta S_{\mathcal{A}}$

$$(\delta S_{\mathcal{A}})_{\text{naive}} = \frac{c}{3} \log\Big[\sinh(r_+\,\theta_\infty)\,\text{csch}(r_+\,(\pi-\theta_\infty)\Big] \tag{2.6}$$

violates the Araki-Lieb inequality. For the BTZ black hole $S_{\rho_\Sigma} = \frac{\pi\,r_+}{2\,G_N^{(3)}} = \frac{2\pi^2}{3}\,c\,T\,R$, which is clearly bounded, while (2.6) diverges as $\theta_\infty \to 0$.

The basic point is simply the following: for regions $\mathcal{A}$ which are sufficiently large $\theta_\infty \geq \theta_\infty^{\mathcal{X}} > \pi/2$ the holographic entanglement entropy is not computed from the connected minimal surface $\mathfrak{M}_1$ but rather from a shorter two-component (disconnected) one $\mathfrak{M}_2$. Note that in any static black hole spacetime, the bifurcation surface is minimal by virtue of the null generator of the horizon vanishing there quadratically, forcing its extrinsic curvature to vanish. While for any $\mathcal{A}$ there exists a connected minimal surface which satisfies the homology condition, there also exists a disconnected minimal surface, given by the union of the (connected) minimal surface for $\mathcal{A}^c$ and the bifurcation surface of the event horizon,

[10] We use CFT central charge determined by the Brown-Henneaux analysis $c = \frac{3\,\ell_{\text{AdS}}}{2\,G_N^{(3)}}$ and note that $T = \frac{r_+}{2\pi\,\ell_{\text{AdS}}\,R}$ is the temperature of the thermal density matrix ρ_Σ.

which likewise does the job. Thus the second family of minimal surfaces, parameterized by θ_∞, is simply

$$\mathfrak{M}_2(\theta_\infty) : \left\{ (r,\theta) :\ r = r_+ \ \ \cup \ \ r = \gamma(\theta, \pi - \theta_\infty, r_+) \right\}, \tag{2.7}$$

where γ is given by eq. (2.4) (an example of both families of surfaces is plotted in fig. 2.2 explained below). Taking this fact into account we learn that the holographic entanglement entropy is given by

$$S_{\mathcal{A}}(\theta_\infty) = \min \left\{ \frac{\text{Area}(\mathfrak{M}_1)}{4\, G_N^{(3)}} \ , \ \frac{\text{Area}(\mathfrak{M}_2)}{4\, G_N^{(3)}} \right\}, \tag{2.8}$$

which then implies upon evaluating the lengths of the curves explicitly that

$$S_{\mathcal{A}} = \begin{cases} \frac{c}{3} \log\left(\frac{2\, r_\infty}{r_+} \sinh(r_+\, \theta_\infty) \right) , & \theta_\infty < \theta_\infty^{\mathcal{X}} \\ \frac{c}{3}\, \pi\, r_+ + \frac{c}{3} \log\left(\frac{2\, r_\infty}{r_+} \sinh(r_+\, (\pi - \theta_\infty)) \right) , & \theta_\infty \geq \theta_\infty^{\mathcal{X}} \end{cases} \tag{2.9}$$

where we introduce the entanglement plateau scale $\theta_\infty^{\mathcal{X}}$ to distinguish the dominant saddle point of the area functional. This can be explicitly evaluated as a function of the black hole size (or temperature) to be

$$\theta_\infty^{\mathcal{X}}(r_+) = \frac{1}{r_+} \coth^{-1}\left(2 \coth(\pi\, r_+) - 1\right). \tag{2.10}$$

For orientation, we plot the two sets of minimal surfaces $\mathfrak{M}_1(\alpha_{\mathcal{X}})$ and $\mathfrak{M}_2(\alpha_{\mathcal{X}})$ at the transition value $\alpha_{\mathcal{X}} = \theta_\infty^{\mathcal{X}}/\pi$ in fig. 2.2 (again for two black hole sizes for ease of comparison with fig. 2.1). At this value of α, $\mathfrak{M}_1$ (solid curve) and $\mathfrak{M}_2$ (dashed curves) have equal proper lengths. Note that in the high temperature (large black hole) limit, the entanglement plateau scale approaches the size of the entire system:

$$\lim_{r_+ \to \infty} \theta_\infty^{\mathcal{X}}(r_+) = \pi\,. \tag{2.11}$$

Translating this into field theory quantities and using $\alpha = \frac{\theta_\infty}{\pi}$ to parameterize the fraction of the system we consider, we have

$$\alpha_{\mathcal{X}} = \frac{1}{2\pi^2\, T\, R} \coth^{-1}\left(2 \coth\left(2\pi^2\, T\, R\right) - 1\right) \tag{2.12}$$

where R is the radius of the boundary CFT cylinder. The main feature we want to illustrate is that for $\alpha > \alpha_{\mathcal{X}}$

$$\delta S_{\mathcal{A}} = S_{\rho_\Sigma} = \frac{2\pi^2}{3}\, c\, T\, R \qquad \Longrightarrow \qquad S_{\mathcal{A}} = S_{\mathcal{A}^c} + S_{\rho_\Sigma} \tag{2.13}$$

as anticipated. Essentially for large enough boundary regions we can read off the thermal entropy by comparing directly the entanglement entropy of a region and its complement.

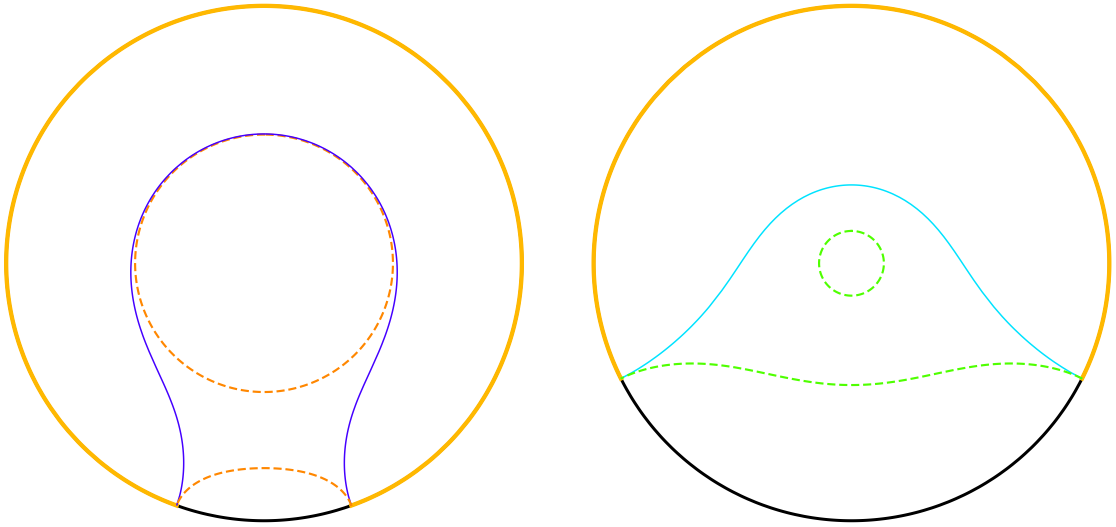


Figure 2.2: Comparison of the two families $\mathfrak{M}_1$ and $\mathfrak{M}_2$ of the minimal surfaces (geodesics) in BTZ, with black hole sizes $r_+ = 1$ (left) and $r_+ = 0.2$ (right) as in fig. 2.1, plotted at the critical value of θ_∞ where they exchange dominance, i.e. at $\theta_\infty = \theta_\infty^{\mathcal{X}}$. The connected family $\mathfrak{M}_1$ is represented by the solid curve (color-coded by θ_∞ as in fig. 2.1), while the disconnected family $\mathfrak{M}_2$ is given by the two dashed curves (color-coded by $\pi - \theta_\infty$). The thick orange arc on the boundary represents the region $\mathcal{A}$.

The behaviour of $\delta S_{\mathcal{A}}$ for $1+1$ dimensional CFTs is shown explicitly later in fig. 2.5 (where we also demonstrate a similar feature in higher dimensional holographic field theories).

In deriving the relations eq. (2.13), we have implicitly assumed $r_+ \geq 1$ (equivalently $TR \geq \frac{1}{2\pi}$), which is where the BTZ black hole is dual to the thermal density matrix for the CFT. For lower temperatures the density matrix is dual to thermal AdS geometry and the holographic computation described above should be modified. Geodesics in thermal AdS will actually give a result which says $S_{\mathcal{A}} = S_{\mathcal{A}^c}$ at leading order in the $c \to \infty$ limit.[11] The thermal result will only be recovered by considering $1/c$ corrections. This makes sense since for $TR < \frac{1}{2\pi}$ one is in the 'confined' phase of the CFT (the terminology is inherited from higher dimensions where one has an honest confinement/deconfinement transition in the large N planar gauge theory).

In $1+1$ dimensions we have also the ability to compare our large c result for $S_{\mathcal{A}}$ or $\delta S_{\mathcal{A}}$ with the behaviour encountered in small central charge systems, which is known for a couple of cases. The result for the entanglement entropy of a free neutral Dirac fermion $c = 1$ was derived in [51] and was extended to a grand canonical density matrix including a $U(1)$ chemical potential in [57].[12] In either case we can use the resulting expressions (which being rather long we refrain from reproducing here) to understand the behaviour of the entanglement in the system. Using the explicit expressions one can check that the resulting

[11] This is the statement that $\mathcal{O}(1)$ corrections cannot be recovered from the classical gravity approximation.

[12] We note that angular momentum chemical potential was considered in [39], which can be captured holographically using a rotating BTZ black hole in the large c limit. However the ensemble is stationary (and not static) so one needs to work with extremal surfaces as emphasized there.

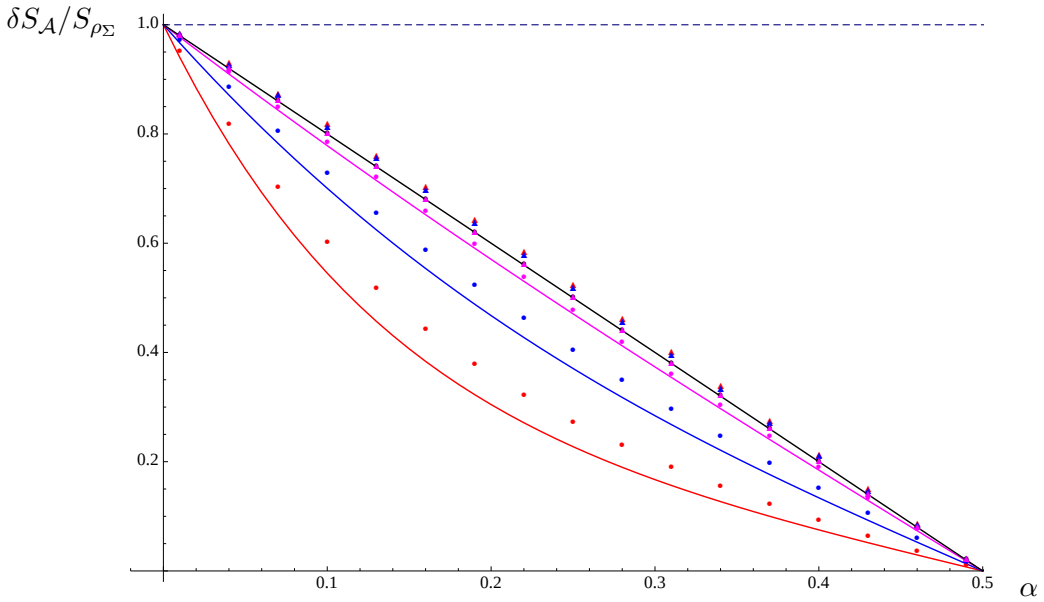


Figure 2.3: Plot of the curves $\delta S_{\mathcal{A}}/S_{\rho_\Sigma}$ for a Dirac fermion in $1+1$ dimensions in the canonical ($T \neq 0$) [51] and grand canonical ($T,\ \mu \neq 0$) [57]. We examine the behaviour for a range of temperatures and chemical potential. The solid curves from bottom are $\beta = 4$ (red), $\beta = 2$ (blue), $\beta = 1$ (magenta) and $\beta = 0.1$ (black). The situation with the chemical potential turned on is indicated with the same colour coding with $\mu = 0.1$ represented by circles and $\mu = 0.5$ by triangles. The symmetry $\mu \leftrightarrow 1-\mu$ is used to restrict $\mu \in [0, 1/2]$ and we see that for $\mu = 0.5$ the normalized $\delta S_{\mathcal{A}}$ is essentially the same at any temperature. This behaviour should be contrasted against the large c holographic result displayed in fig. 2.5 for $d \geq 2$ thermal CFTs.

answer exhibits no sharp feature. $\delta S_{\mathcal{A}}/S_{\rho_\Sigma}$ is a smooth monotone function of α, see fig. 2.3. This is also consistent with the recent discussion of [58] who compute analogous quantities in a harmonic chain which is a gapped theory and note that sharp features can only occur in the $c \to \infty$ limit (which as we mention earlier implements the thermodynamic limit).

Note that the Dirac fermion example is so far the only case for which the entanglement entropy of a thermal CFT in finite volume is explicitly known. In general one expects that the details of the spectrum of the CFT play a crucial role in the entanglement, i.e., the answer is not a simple universal function of the central charge as the holographic result (2.9) suggests. It would be interesting to see if one can use the technology of [34] to argue directly from a CFT analysis for an entanglement plateau relation like (2.13) in the asymptotic large c limit.

The behaviour of $S_{\mathcal{A}}(\alpha)$, not surprisingly, is in fact rather similar to the behaviour of the entropy or free energy in the thermal ensemble (a point we will revisit in §2.7). In both cases there are multiple saddle points, which exchange dominance. For the thermal density matrix the saddles are the thermal AdS geometry and the BTZ black hole both of which exist for the entire range of the dimensionless parameter TR. For the holographic entanglement entropy $S_{\mathcal{A}}$ we have again two distinct saddles available for the entire range of parameters: the connected and disconnected surfaces $\mathfrak{M}_1$ and $\mathfrak{M}_2$ exist for all values of

$\alpha \in [0,1]$. For a sufficiently large region $\alpha \geq \alpha_{\mathcal{X}}$, however, it is the disconnected surface that dominates and results in the entanglement plateau. We will soon see that existence of both families over the entire range of α is peculiar to three dimensions and the situation is much more intricate in higher dimensions.

2.4 Entanglement entropy in higher dimensional field theories

In the previous section we saw that already in the 3-dimensional bulk spacetime corresponding to a thermal state in the dual CFT, we have multiple families of minimal surfaces $\mathfrak{M}_i$ anchored at the boundary entangling surface $\partial\mathcal{A}$, for arbitrary $\mathcal{A}$. Let us now examine the analogous situation in higher dimensions. We will consider the Schwarzschild-AdS_{d+1} bulk spacetime ($\ell_{\text{AdS}} = 1$),

$$ds^2 = -f(r)\,dt^2 + \frac{dr^2}{f(r)} + r^2\,\left(d\theta^2 + \sin^2\theta\, d\Omega_{d-2}^2\right)\,, \qquad f(r) = r^2 + 1 - \frac{r_+^{d-2}\,(r_+^2+1)}{r^{d-2}}\,. \tag{2.14}$$

Large black holes $r_+ \geq 1$ describe the thermal state on ESU (e.g. of $\mathcal{N}=4$ SYM on $\mathbf{S}^3 \times \mathbb{R}$ in the best-understood case of $d=4$), but we will consider black holes of any size for generality. For simplicity, we will take the boundary region $\mathcal{A}$ to be a disk centered at the 'north pole' $\theta = 0$ with radius θ_∞, and consider only surfaces which maintain the residual $SO(d-1)$ spherical symmetry and remain at constant t. This effectively reduces the problem to a 2-dimensional one: we can specify any such minimal surface as a curve in the (r,θ) plane.[13]

The equations of motion for the minimal surface are obtained from the Nambu-Goto action for minimizing the area of the surface. Using an auxiliary parameter s we have the Lagrangian for the system

$$\mathcal{L} = (r\,\sin\theta)^{d-2}\,\sqrt{\frac{1}{f(r)}\,\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\theta}{ds}\right)^2}\,. \tag{2.15}$$

The Euler-Lagrange equations for $r(s)$ and $\theta(s)$ are equivalent due to the reparameterization invariance, giving a second order ODE for $r(s)$ and $\theta(s)$. The second equation comes from choice of parameter.[14]

The topology of the problem constrains any surface to pass through a pole of the sphere, so to classify all connected surfaces we may start integration (w.l.o.g.) at the North pole. Requiring smoothness of $\mathfrak{M}$ there, we obtain a one-parameter family of minimal surfaces,

[13] For plotting purposes, we will consider a compactified radial coordinate $\rho = \tan^{-1} r$ as in §2.3, and double up the $\theta \in [0,\pi]$ coordinate to $\theta \in [-\pi,\pi]$ so that all curves will be reflection-symmetric. The black hole will then be represented by a disk of radius $\tan^{-1} r_+$ centered at the origin.

[14] One might be tempted to use θ itself as a parameter, but it turns out that θ is not monotonic in higher dimensions. We choose s such that the evaluation of the area integral on-shell reduces to $\int ds\ (r\,\sin\theta)^{d-2}$, for good behaviour numerically. In particular, the results for the regularized area of the surface are less error-prone since r increases exponentially in s as the boundary is approached.

specified by the 'initial value' for the radial coordinate at the north pole, $r_0 \equiv r(\theta = 0)$.[15] Given r_0, we find the minimal surface and from this read off the latitude θ_∞ which it is anchored on. From this we will classify the boundary region size by α, the proportion of the area contained in the boundary region homologous to the surface: $\alpha = \frac{\mathrm{Vol}(\mathcal{A})}{\mathrm{Vol}(\mathbf{S^{d-1}})}$ which tends to $0, 1$ as $\theta_\infty \to 0, \pi$ respectively.

2.4.1 Entanglement plateaux in $d > 2$

Considering α as a function of initial radius r_0, the situation is rather different from the $d = 2$ case. In $d > 2$, α reaches a maximum value of $\alpha_m < 1$: for sufficiently large regions, there are no connected minimal surfaces obeying the homology constraint! This can be intuited from analogous behaviour in the classic 'soap bubble' problem in flat space, of finding a minimal surface between two circular rings. To reduce the area, there is a tendency to shrink the radius of the tube, counteracted by the constraint of ending on the rings. But when the ratio of ring radius to separation is sufficiently small, the rings do not hold the surface up enough to prevent the tube radius from shrinking to zero, and the surface separates into two disconnected parts. This reasoning carries over directly to our set-up, and the process of surfaces splitting into disconnected pieces can be seen explicitly in the animations of our numerical simulations of mean curvature flow, to be found on the arXiv as http://arxiv.org/src/1306.4004/anc (ancillary files) for the submission. The area cost of having a wide tube is greater in higher dimensions, which leads one to expect that α_m should be smaller, reflecting the greater tendency of the surfaces to split up.[16] The physical point is that this leaves no option but to consider the disconnected surfaces, to which we now turn. In the next section, we offer a very different geometrical justification of why connected extremal surfaces homologous to $\mathcal{A}$ cannot exist beyond a certain θ_∞, which is based on causality in the full Lorentzian geometry.

So far, we have a one-parameter family of connected surfaces $\mathfrak{M}_1$. In addition to this, we must consider the 'disconnected' family $\mathfrak{M}_2$ of surfaces with two connected components: the first, anchored to the boundary, is a reflection of a connected surface already considered, passing through the South pole. Due to the nontrivial $\mathbf{S}^{d-1} \times \mathbb{R}$ topology of the static slice of the geometry, this alone fails to satisfy the homology constraint, so it is supplemented by a second piece, the bifurcation sphere on the event horizon of the black hole. To identify the correct entangling surface determining $S_{\mathcal{A}}(\alpha)$, we must compare the areas of these two families; they exchange dominance at $\alpha_{\mathcal{X}} < \alpha_m$, as is inevitable from continuity. This means that at this value $\alpha_{\mathcal{X}}$, there are two surfaces of equal area, and as we vary α through $\alpha_{\mathcal{X}}$, the surfaces jump and the entanglement entropy as a function of region size has a kink. In fig. 2.4 we show the surfaces which determine $S_{\mathcal{A}}(\alpha)$ for various black hole sizes

[15] Apart from ODE methods, the numerical construction of the surfaces is also done using a mean curvature flow. The reader interested in the details is encouraged to consult Appendix 2.A where we outline the necessary mathematical technology and the algorithm used for obtaining the surfaces.

[16] This behaviour is in fact reminiscent of similar observation in gravitational context of the Gregory-Laflamme instability of higher-dimensional black strings (or branes), which are likewise more prone to fragmenting with increasing dimension [59, 60].

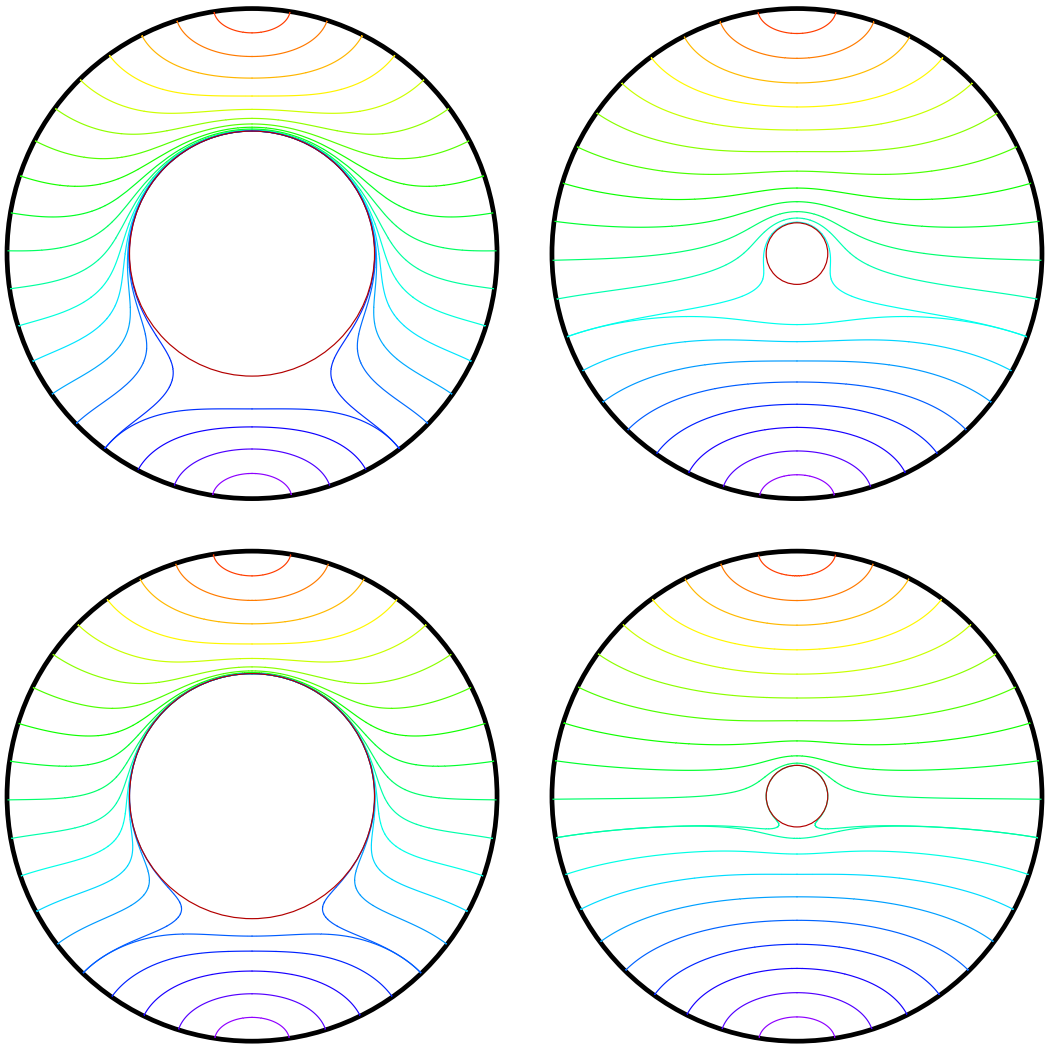


Figure 2.4: Minimal surfaces for $r_+ = 1$ (left) and $r_+ = 0.2$ (right) black holes in Schwarzschild-AdS in $3+1$ (top) and $4+1$ (bottom) dimensions, analogous to the plots of fig. 2.1. In each panel, the thick black circle represents the AdS boundary, and the thin red circle the black hole horizon. The boundary region $\mathcal{A}$ is centered at $\theta = 0$ (North Pole) which is plotted on top. Further plots with other black hole sizes and higher dimensions can be found with the submission.

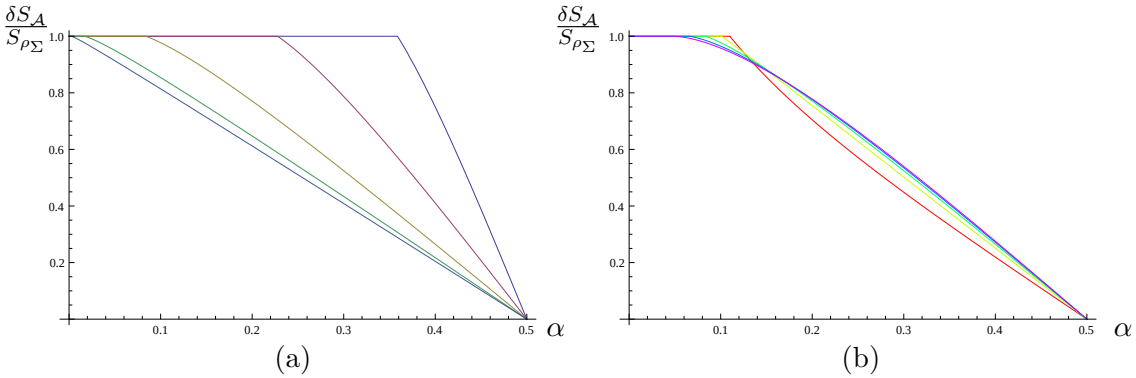


Figure 2.5: Entanglement entropy plateaux for Schwarzschild-AdS black holes dual to thermal field theories. (a). Behaviour in Schwarzschild-AdS_5 for varying r_+. From right to left we have $r_+ = \frac{1}{4}$ (blue), $r_+ = \frac{1}{2}$ (purple), $r_+ = 1$ (yellow), $r_+ = 2$ (green) and $r_+ = 4$ (blue). (b). Dimension dependence of the plateaux, at the Hawking-Page transition point $RT = \frac{d-1}{2\pi}$ for $d = 2, 3, 4, 5, 6$ (red, yellow, green, blue, purple respectively). Note that the horizon size is related to the temperature via $r_+ = \frac{2\pi}{d}\left(RT + \sqrt{(RT)^2 - \frac{d(d-2)}{4\pi^2}}\right)$.

and dimensions. Those passing 'above' the black hole belong to the family $\mathfrak{M}_1$, and those passing 'below' should be supplemented by the horizon, and belong to $\mathfrak{M}_2$. Comparing the panels horizontally, we see that the smaller the black hole, the smaller $\alpha_{\mathcal{X}}$ gets; indeed, as $r_+ \to 0$, $\alpha_{\mathcal{X}} \to 1/2$, whereas as $r_+ \to \infty$, $\alpha_{\mathcal{X}} \to 1$; we plot the actual curve $\alpha_{\mathcal{X}}(r_+)$ in section 2.5. On the other hand, comparing the panels vertically, we see that as d increases, the surfaces get less affected by the black hole until close to the horizon. This is easy to understand from the simple fact that gravity falls off faster in higher dimensions, and chimes perfectly with the intuition recently explained in [60]. Turning to common characteristics, one universal feature of all cases is that none of the minimal surfaces (anchored on the boundary) can penetrate past the event horizon. This is in fact true for any static geometry with a horizon [61].[17]

It is evident from the symmetry of our setup that in the regime $\alpha \geq \alpha_{\mathcal{X}}$, we have the identity

$$S_{\mathcal{A}} = S_{\mathcal{A}^c} + S_{\rho_\Sigma} \tag{2.16}$$

so that we manifestly saturate the Araki-Lieb inequality. This is identical to the situation described in §2.3. Indeed, a plot of $\delta S_{\mathcal{A}}$ as a function of α for various dimensions explicitly reveals the entanglement plateaux as illustrated in fig. 2.5 where we include the result for $d = 2$ CFTs as well for completeness. In fig. 2.5(a) we show the behaviour for a fixed dimension $d = 4$ as we vary the temperature by changing the black hole size, whereas in

[17] More specifically, [61] showed that for asymptotically AdS spacetimes with planar symmetry the deepest reach (i.e. the turning point, furthest away from the boundary) of any extremal surface which is fully anchored on (a single) AdS boundary cannot occur inside a black hole. On the other hand, in time-dependent spacetimes, there do exist extremal surfaces anchored on the boundary which penetrate past the horizon [2, 61, 62, 63, 64, 65]. For this reason, entanglement entropy is cognizant of some physics inside the horizon.

fig. 2.5(b) we display the behaviour for different dimensions at fixed $TR = \frac{d-1}{2\pi}$. We see that the entanglement plateau ends by a kink (e.g. in $d = 2$ the behaviour of $\delta S_{\mathcal{A}}/S_{\rho_\Sigma}$ can be obtained from (2.9)), the extent of which gets smaller with increasing dimension. To be sure, the precise details depend on the nature of the comparison; we have found it reasonable to fix the temperature (in units of the boundary sphere) and compare the behaviour across dimensions. In general the plateau width decreases as a function of T for a given dimension (the behaviour as a function of d depends on the precise value of RT; see ancillary files for $RT = \frac{2}{3}$).

2.4.2 Sub-dominant saddles: folds in minimal surfaces

Although we now have the curve $S_{\mathcal{A}}(\alpha)$ for the full range of α, it is interesting to ask what happens to the connected minimal surfaces (and hence the disconnected ones as well, by the reflection symmetry) in their sub-dominant regime. We know that a continuation of these families must exist, since the deepest radius r_0 did not cover the full range down to r_+.

As we decrease r_0 from the critical value $r_0(\alpha_{\mathcal{X}})$, the endpoint of the surface α increases beyond $\alpha_{\mathcal{X}}$ as expected, but as previously noted, reaches a maximum at $\alpha = \alpha_m$, beyond which there simply are no connected surfaces at all. Instead, as one decreases r_0 further, the endpoints α start to recede to lower values again, even though the 'neck' of the surface $\mathfrak{M}_1$ near $\theta = \pi$ keeps closing off. This behaviour is illustrated in fig. 2.6. Bringing r_0 still closer to the horizon reveals another surprise: the pattern repeats itself. Eventually α reaches a minimal value (which happens to be very close to $1 - \alpha_m$) and turns around again. In other words, now the corresponding minimal surface has two necks, one near $\theta = \pi$ and the other near $\theta = 0$.

Careful examination reveals that the closer the minimal radius r_0 is to the horizon, the more intricate the surface. The endpoints θ_∞ are restricted to lie sufficiently close to the equator $\theta = \frac{\pi}{2}$, in particular, $1 - \alpha_m \lesssim \alpha \lesssim \alpha_m$, as manifested in fig. 2.7, where we plot α as a function of rescaled coordinate $x_0 = \frac{1}{2}\log(r_0 - r_+)$ indicating the proximity of r_0 to the horizon (cf Appendix 2.B).

In fact, the pattern of minimal surfaces exhibits a discrete self-similarity. The minimal surfaces wrap the horizon multiple times, the distance from the horizon of each wrapping parametrically separated from the next. We denote the solution with n-folds around the horizon as $\mathfrak{M}_{1,(n)}$. The connected surfaces described earlier in fig. 2.4 are $\mathfrak{M}_{1,(0)}$ with this refinement in the notation. The existence of these solutions and their properties can be analytically understood using a local analysis around the turning points, and near the horizon away from the turning points, as we illustrate in Appendix 2.B. Further, this analysis is generic for any spherically symmetric non-extremal horizon, so there is nothing particularly special about Schwarzschild-AdS in this regard.

Of course, such surfaces which fold around the black hole multiple times have larger area (approximately by the black hole area times the number of foldings) than their simpler cousins, so these surfaces are not directly relevant for the entanglement entropy. This

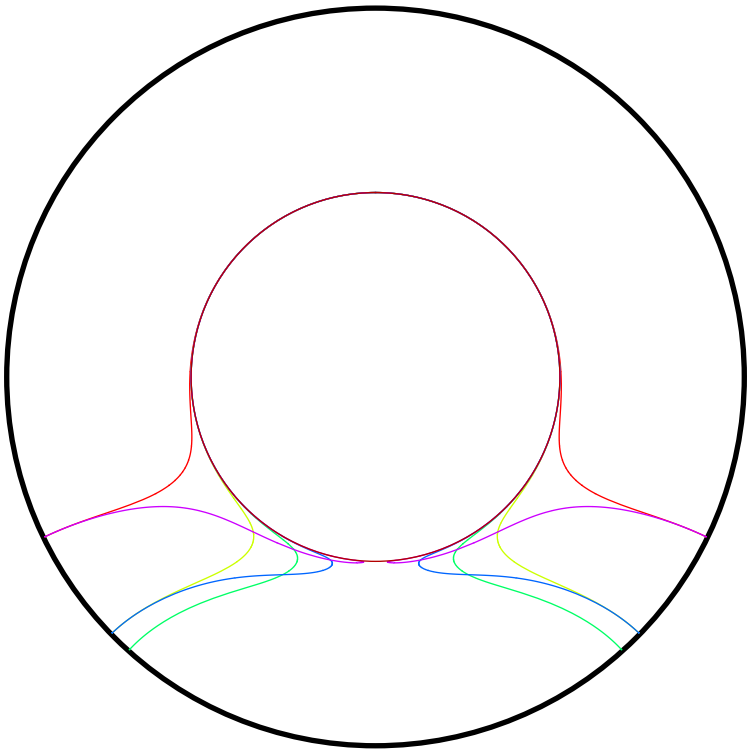


Figure 2.6: Continuation of family $\mathfrak{M}_1$ of minimal surfaces in Schwarzschild-AdS_5 for $r_+ = 1$. As the turning point r_0 approaches the horizon, we see an intricate pattern of the minimal surfaces folding back onto themselves; solutions with n-folds are denoted $\mathfrak{M}_{1,(n)}$. We illustrate $\mathfrak{M}_{1,(0)}$ (red, yellow) and $\mathfrak{M}_{1,(1)}$ (blue, purple) surfaces for two specific choices of α; these two families terminate at $\alpha = \alpha_m$ (green).

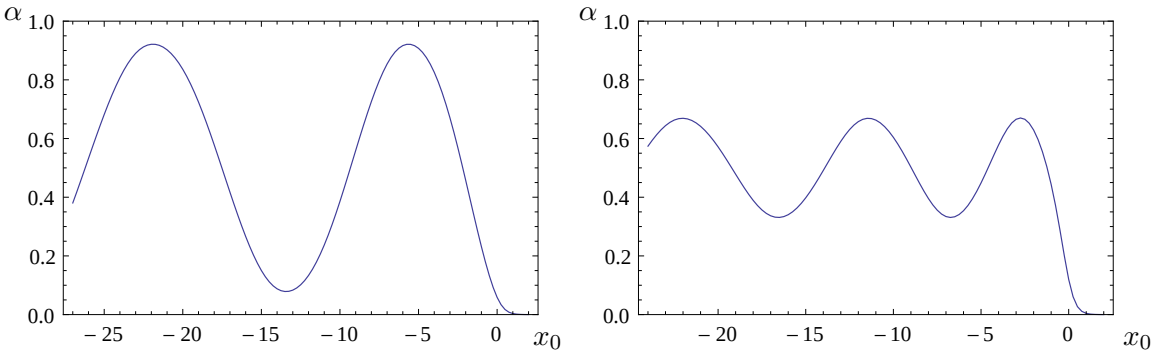


Figure 2.7: Near-horizon analysis of folds in the minimal surfaces for $r_+ = 1$ (left) and $r_+ = 0.2$ (right) black holes in Schwarzschild-AdS_4. Here $x_0 = \frac{1}{2}\log(r_0 - r_+)$ measures the closest approach of the surface to the horizon.

is already visible explicitly in fig. 2.8, where we plot the actual areas of the minimal surfaces, as a function of α. For $\alpha > \alpha_{\mathcal{X}}$ the disconnected family $\mathfrak{M}_2$ takes over, a fact we have already used to demonstrate the entanglement plateau in fig. 2.5. In the regime $\alpha > \alpha_{\mathcal{X}}$ the connected surfaces have greater area, and indeed the area continues to grow as the surface folds over about the horizon. It is intriguing to note that the difference $\text{Area}(\mathfrak{M}_{1,(n)}) - \text{Area}(\mathfrak{M}_{1,(n-2)}) \approx 2\, S_{\rho_\Sigma}$ increasingly more accurately as n gets large. While the sub-dominant connected surfaces $\mathfrak{M}_{1,(n)}$ for discussed above are irrelevant for the boundary entanglement for $\alpha > \alpha_{\mathcal{X}}$ it is nevertheless curious that such a simple geometry as Schwarzschild-AdS allows for such a rich structure of minimal surfaces!

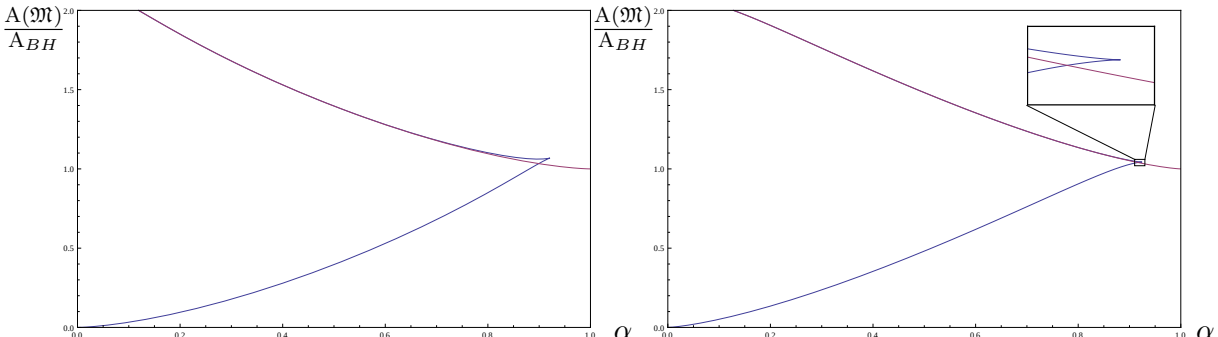


Figure 2.8: Regulated areas of minimal surfaces (normalized in units of black hole area), for (a) Schwarzschild-AdS_4 ($r_+ = 1$), and (b) Schwarzschild-AdS_5 ($r_+ = 1$). The blue curve is the family $\mathfrak{M}_1$, and the purple curve the disconnected family $\mathfrak{M}_2$. The area of the connected minimal surfaces grows without bound, since we encounter multiply-folded surfaces $\mathfrak{M}_{1,(n)}$ which for $n > 1$ which wrap the horizon multiple times. The swallow-tail behaviour characteristic of first order phase transitions is clearly visible in the plots.

2.5 Causality and holographic entanglement

In section 2.4 we have examined the families of minimal surfaces in static thermal density matrices for holographic theories which compute the entanglement entropy. The main surprise was that beyond a certain size of $\mathcal{A}$, namely for $\alpha > \alpha_m$, there are no connected surfaces homologous to $\mathcal{A}$, which automatically forces the holographic entanglement entropy to plateau in the sense described earlier. Above, we demonstrated this result by explicitly constructing minimal surfaces in the bulk Schwarzschild-AdS_{d+1} spacetime; the robustness of numerical construction rests on explicitly allowing surfaces to relax to minimal ones using a mean curvature flow as explained in Appendix 2.A. However, the absence of connected minimal surfaces is in fact necessitated by a certain relational property of minimal surfaces and causal wedges. The latter is a construct that crucially involves the Lorentzian structure since it has to do with causal relations between points. We now proceed to explain the basic physics behind this connection, leaving the technical details for a separate discussion [55].

Given a boundary region $\mathcal{A} \subset \Sigma$, one can construct a causal wedge (which, following [54], we denote by $\blacklozenge_{\mathcal{A}}$) associated with it in the bulk. The construction involves first identifying the boundary domain of dependence $\lozenge_{\mathcal{A}}$ of $\mathcal{A}$, which for the rigid boundary geometry is simply the region where one can reliably Cauchy-evolve initial data laid down on $\mathcal{A}$. If one has full knowledge of $\rho_{\mathcal{A}}$ then one can compute correlation functions of all local operators inserted in $\lozenge_{\mathcal{A}}$; this is just a basic statement of causality in relativistic quantum field theories. We then construct the causal wedge $\blacklozenge_{\mathcal{A}}$ in the bulk, defined as the set of points in $\mathcal{M}$ which can communicate with and simultaneously be communicated from $\lozenge_{\mathcal{A}}$; in other words, through which there exists a causal curve which starts and ends in $\lozenge_{\mathcal{A}}$. The reader interested in the precise definitions is invited to consult [54].

Since the causal wedge is constructed purely based on where causal curves reach into the bulk, one might naively imagine that for a simply-connected boundary region $\mathcal{A}$ the causal wedge has trivial topology. This turns out not to be the case [55]. While the causal wedge for a simply connected region has to be simply connected (a consequence of topological censorship), it can have 'holes'. This implies that $\blacklozenge_{\mathcal{A}}$ possesses non-contractible $d-2$ spheres on its boundary. Intuitively, this can happen because of geodesic trapping, as we now explain. The boundary of the causal wedge is generated by null geodesics. Bulk spacetimes which possess null circular orbits (arising from a sufficiently strong gravitational potential) cut off some of these generators, while allowing some others to fly-by and thereby create a hole in the wedge. Intuitively, the play-off is always between gravitational attraction and centrifugal repulsion and for a wide family of geometries one is able to engineer these so as to obtain the desired effect. Happily for us, the simplest class of examples where such a feature can be exhibited explicitly are the Schwarzschild-AdS_{d+1} geometries, wherein for large enough region $\mathcal{A}$ (which we denote by $\alpha > \alpha_{\blacklozenge}$), the causal wedge $\blacklozenge_{\mathcal{A}}$ exhibits a hole [55].

The presence of non-trivial topology has an important implication for the minimal

surfaces $\mathfrak{M}$. This is at first sight surprising, as one might have naively thought that since minimal surfaces are defined in the Riemannian section (constant-time slice), they ought to be utterly ignorant of any causal argument which is intrinsically Lorentzian. This naive reasoning is however too quick: One can show that all extremal surfaces $\mathcal{E}_{\mathcal{A}}$ (of which the minimal surfaces $\mathfrak{M}_{\mathcal{A}}$ considered herein are a special case) must lie outside (or at best on the boundary of) the causal wedge [40, 54, 55]. Here by 'outside' we mean 'not within' – in other words an extremal surface $\mathcal{E}_{\mathcal{A}}$ must penetrate deeper into the bulk than the corresponding causal wedge $\blacklozenge_{\mathcal{A}}$. The result can be intuitively understood by realizing that a minimal surface wants to minimize its area and thus wants to spend much of its existence deep in the interior where it can minimize the red-shift factor of AdS (which augments the area element of the directions along the boundary). This effect is less dominant for lower-dimensional surfaces, and null geodesics can additionally offset the large spatial distances by large temporal ones.[18] Hence generically the causal wedge boundary is nestled between the extremal surface and the AdS boundary. As an aside, this statement forms a crucial ingredient in the arguments of [66] who argue that the holographic dual of the reduced density matrix $\rho_{\mathcal{A}}$ must comprise of a bulk region that is larger than the casual wedge $\blacklozenge_{\mathcal{A}}$.[19]

We are now in a position to explain the behaviour found in §2.4. For $\alpha > \alpha_{\blacklozenge}$ the ability of the causal wedge to develop holes implies that we should anticipate a corresponding change in the minimal surface. Indeed $\mathfrak{M}_{\mathcal{A}}$ cannot pass into the hole in $\blacklozenge_{\mathcal{A}}$ while remaining connected to the boundary, because of the nesting property: in order to do so, it would have to pass through the causal wedge. This means that it either lies completely outside the causal wedge or is contained entirely within the hole! One fact which is somewhat obvious by causality is that the hole in the causal wedge must lie outside the black hole event horizon. For static black hole spacetimes we have argued that the bifurcation surface is a candidate extremal surface in §2.3. So in the presence of the hole in the wedge we can have a minimal surface component on the far side of the black hole (hence outside $\blacklozenge_{\mathcal{A}}$) and another component being simply the bifurcation surface. In fact, as in the BTZ discussion, the homology constraint imposed upon the RT prescription by the Araki-Lieb inequality necessitates both components. Hence the only surfaces that are both minimal and homologous to $\mathcal{A}$ are the disconnected surfaces $\mathfrak{M}_2$ described in §2.4.

We demonstrate explicitly that $\alpha_{\mathcal{X}} \leq \alpha_m \leq \alpha_{\blacklozenge}$ as one anticipates for first order transitions for Schwarzschild-AdS_5 in fig. 2.9. As the figure indicates, the causal wedge pinch off happens at $\alpha = \alpha_{\blacklozenge}$ which is significantly larger than the bound α_m on existence of corresponding extremal surfaces. In other words, the connected minimal surface ceases

[18] Note that in comparison to pure AdS geometry, both null geodesics and extremal surfaces are nevertheless pushed towards the boundary by the gravitational potential well of a deformed bulk spacetime. This in effect means that both entanglement entropy and causal holographic information defined in [54] grow with positive mass deformations of the spacetime.

[19] *Caveat:* The statements made above should be viewed with suitable caution, as they require some work to be established rigorously. For further discussion and a proof of the nesting property of causal wedges and extremal surfaces (modulo some technical assumptions), we refer the reader to [55] (see also [40] for a related discussion in causally trivial spacetimes).

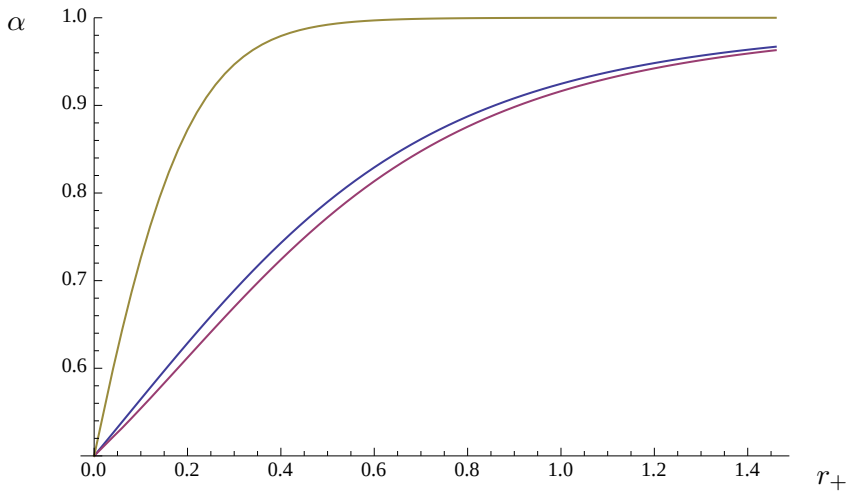


Figure 2.9: The critical values of α encountered in our discussion as a function of r_+. We plot the largest fraction of the boundary region characterized by α for which (i) a connected minimal surface exists α_m (blue), (ii) for which disconnected/connected exchange dominance $\alpha_{\mathcal{X}}$ (purple) and, (iii) the critical α for which the causal wedge develops holes $\alpha_{\blacklozenge}$ (olive) in the Schwarzschild-AdS_5 spacetime.

to exist long before this is necessitated by the topology of the causal wedge. While this may seem to somewhat weaken our arguments regarding the utility of the causal wedge, we should note that the pinch-off value $\alpha = \alpha_{\blacklozenge}$ is a rather weak bound from the causal wedge standpoint as well: already before the causal wedge pinches off, its geometry precludes the requisite minimal surface. This is because prior to the pinch-off the opening becomes too long and narrow to admit any extremal surface. To support the latter, the neck would have to accommodate a sufficient 'flare-out' shape, as explained in section 2.4.1, which occurs for θ_∞ smaller by $\sim \mathcal{O}(\rho_0 - \rho_+)$, where ρ_0 is the ρ-radius of null circular orbit and ρ_+ gives the horizon size. Hence a more sophisticated analysis which takes into account the shape of the causal wedge apart from its topology would provide a much stronger bound on α_m. It would interesting to examine in detail precisely how close we can get to α_m, but we leave this for future exploration.

2.6 Minimal versus extremal surfaces and the homology constraint

So far, we have examined the properties of the entanglement entropy as given by the RT proposal involving minimal surfaces (which is a well-defined problem if we can look for our surfaces in a Riemannian geometry). This allowed us to argue in section 2.2 that the entanglement entropy for a certain region $\mathcal{A}$, measured in a fixed total density matrix ρ_Σ, is continuous as a function of parameters α describing the region $\mathcal{A}$. On the other hand, the general prescription for holographic entanglement entropy must be formulated in a fully covariant fashion. In other words, in absence of a geometrically-preferred bulk spacelike

slice (which for static cases we could take to be a constant-t slice where $\frac{\partial}{\partial t}$ is the timelike Killing field), we need to have a prescription which does not rely on a specific coordinate choice. This motivated HRT [39] to propose the extremal surface prescription eq. (2.2), which is fully covariant and well-defined for arbitrary time-dependent asymptotically-AdS bulk geometries. It has henceforth been assumed that the two prescriptions, RT and HRT, are equivalent for static geometries. In this section we wish to revisit this assumption, in conjunction with the closely-related issue of the homology requirement.

Let us start by specifying the prescriptions more explicitly, for completeness also adding a recent 'maximin' reformulation of HRT by Wall [40]. In each formulation, the surface is required to be anchored on $\partial\mathcal{A}$, satisfy the homology constraint, and in case of multiple surfaces satisfying these criteria, be the minimal-area one. The differences in the prescriptions enter at the level of constructing the requisite surfaces, and can be summarized as follows:

- RT [30, 31]: minimal surface on constant-t spacelike slice Σ_t^{r-bulk}
- HRT [39]: extremal surface in the bulk $\mathcal{M}$
- Wall [40]: minimal surface on any achronal bulk slice $\tilde{\Sigma}$, maximized over all possible $\tilde{\Sigma}$.

While RT is restricted to static spacetimes, Wall's prescription is formulated in causally trivial bulk geometries. In this context, [40] proves equivalence between the maximin and HRT prescriptions assuming the null energy condition. Although the maximin construction is useful for some purposes (for example, it allows [40] to argue the existence of such surfaces and prove strong subadditivity in the time-dependent context), it is conceptually more complicated since the requisite surface is obtained by a two-step procedure of first minimizing the area on some achronal slice $\tilde{\Sigma}$, and then maximizing the area with respect to varying $\tilde{\Sigma}$. Moreover, here we wish to consider causally non-trivial spacetimes, so we will henceforth restrict attention to the RT and HRT proposals.

In a globally static geometry, RT and HRT proposals are indeed equivalent, since any extremal surface anchored at constant t on the boundary must coincide with the minimal surface on Σ_t^{r-bulk}, cf. [39].[20] However, in a static but not globally static geometry (i.e. when there is a global Killing field which is timelike near the AdS boundary but does not necessarily remain timelike everywhere in the bulk), the situation can appear more subtle.

To illustrate the point, let us consider the eternal non-extremal Reissner-Nordstrom-AdS geometry, corresponding to a static charged black hole. It has a metric of the form eq. (2.14), but with $f(r)$ having two positive real roots at $r = r_\pm$ with r_+ corresponding to

[20] It is worth commenting that the proof presented in [39] shows that the HRT proposal reduces to the minimal surface RT proposal in static geometries. The question of whether there are extremal surfaces not localized on a constant time slice and having smaller area than the minimal surface is, as far as we are aware, open. One might argue based on Euclidean continuation of static geometries that such surfaces do not play any role in determining $S_{\mathcal{A}}$; we thank Tadashi Takayanagi for useful discussions on this issue.

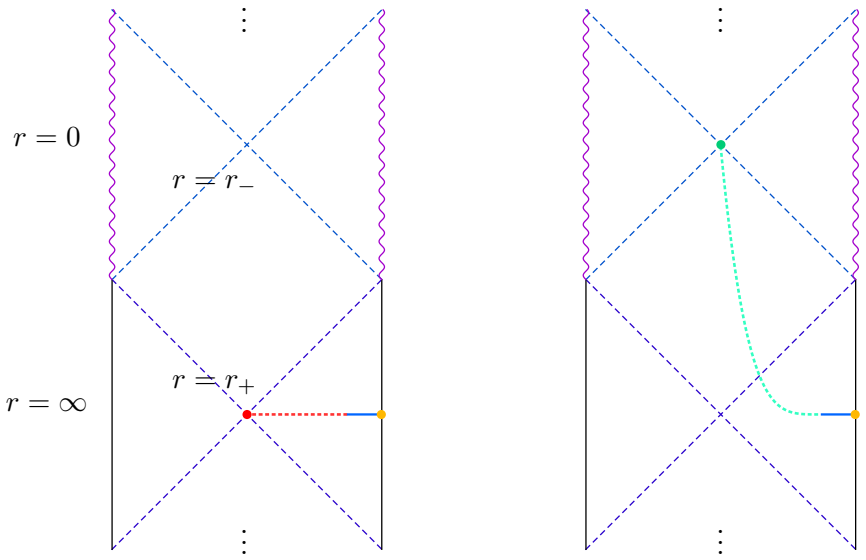


Figure 2.10: A sketch of the Penrose diagram for non-extremal Reissner-Nordstrom-AdS black hole (the figure repeats in the vertical direction, indicated by the ellipsis). The AdS boundaries are indicated by the black vertical lines, the curvature singularities by purple vertical wiggly curves, the horizons by blue diagonal dashed line, the projection of the boundary region $\mathcal{A}$ by the orange dot, and the projection of the extremal surface by the blue horizontal line. The two panels distinguish the RT (left) and naive HRT (right) prescriptions in case of disconnected surfaces: in the former the disconnected surface lies at the event horizon (red dot) with the homology region $\mathcal{R}$ indicated by the red dotted line. In the HRT case, the disconnected surface is at the Cauchy horizon (green dot), and corresponding $\mathcal{R}$ then interpolates between this surface and its other boundaries as indicated by the green dotted curve.

the outer (or event) horizon and r_- the inner (or Cauchy) horizon with $0 < r_- < r_+$.[21] The causal structure is indicated in the Penrose diagram in fig. 2.10. Apart from the bifurcation surface of the event horizon at $r = r_+$ which has zero extrinsic curvature (and therefore is a compact extremal surface), there is an analogous compact extremal surface at the inner horizon $r = r_-$. At both surfaces the $\frac{\partial}{\partial t}$ Killing field vanishes (and its norm decreases in spacelike directions and increases in timelike directions), but whereas the area of the surface is minimized to spatial deformations and maximized to temporal ones in the case of event horizon, it is the other way around for the inner horizon.

Outside the black hole, the geometry is qualitatively similar to that of Schwarzschild-AdS, in the sense that e.g. causal wedges for large enough boundary regions $\mathcal{A}$ will have holes. This will in turn preclude the existence of connected extremal surfaces anchored on $\partial\mathcal{A}$ and homologous to $\mathcal{A}$ in this regime, by the argument of section 2.5. In such a regime, the RT prescription then instructs us to take the surface homologous to $\mathcal{A}^c$, along with the compact surface wrapping the event horizon. On the other hand, in the HRT prescription, there would a-priori be nothing to prevent us from taking the latter component to wrap

[21] For example, in 4+1 bulk dimensions, $f(r) = r^2+1-\frac{M}{r^2}+\frac{Q^2}{r^4} = \left(1-\frac{r_+^2}{r^2}\right)\left(1-\frac{r_-^2}{r^2}\right)(r^2+1+r_+^2+r_-^2)$.

the inner horizon instead: its area is manifestly smaller than that of the event horizon, and it still satisfies the (naive) homology constraint, as indicated pictorially by the green curve in right panel of fig. 2.10. In particular, there exists a smooth co-dimension one surface whose only boundaries are the inner horizon, the extremal surface homologous to $\mathcal{A}^c$, and the region $\mathcal{A}$ on the boundary.

Note that if this were indeed the correct prescription, then we would find that, instead of saturating the Araki-Lieb inequality as in eq. (2.16), we would only satisfy it: $\delta S_{\mathcal{A}}$ would still plateau as a function of θ_∞, but at a value which is lower[22] than the expected value S_{ρ_Σ}:

$$\delta S_{\mathcal{A}}(\theta_\infty \geq \theta_\infty^{\mathcal{X}}) = \frac{\Omega_{d-1}\, r_-^{d-1}}{4\, G_N^{(d+1)}} = \left(\frac{r_-}{r_+}\right)^{d-1} S_{\rho_\Sigma}\,. \tag{2.17}$$

Although this result is consistent with the Araki-Lieb inequality, it is nevertheless at odds with the CFT expectation: for nearly-neutral black holes where $r_- \ll r_+$ we should be close to the thermal value rather than parametrically separated from it!

This observation suggests that we need some modification to the homology constraint specification in eq. (2.2): the mere presence of *some* smooth surface $\mathcal{R}$ with the requisite boundaries does not seem to suffice. One natural restriction, which indeed has already been employed in [40], is to require that $\mathcal{R}$ be everywhere spacelike.[23] With this additional restriction, the above example would be safely invalidated, since $\mathcal{R}$ cannot reach from the Cauchy horizon bifurcation surface $r = r_-$ to the AdS boundary $r = \infty$ while remaining spacelike everywhere, as evident from fig. 2.10. The only other compact extremal surface which is spacelike-separated from the boundary region $\mathcal{A}$ and the extremal surface homologous to $\mathcal{A}^c$ is the event horizon bifurcation surface, which recovers the thermal answer eq. (2.16), consistently with our expectations.

However, while the spacelike restriction on the homology constraint recovers the thermal answer for the global eternal charged black hole, there are other geometries where this does not suffice. As our second exhibit, consider a Schwarzschild-*AdS* 'bag of gold' geometry, discussed in e.g. [53, 67]. This has causal structure and a spatial embedding geometry as sketched in fig. 2.11. The right asymptotic region, as well as interior of the black hole and white hole, are the same as in the eternal Schwarzschild-*AdS* geometry, but the left asymptotic region is modified by a presence of a shell whose interior has a smooth origin. Moreover, one can fine-tune the shell's trajectory such that it remains static – so the entire spacetime admits a Killing field $\frac{\partial}{\partial t}$. In the CFT dual, such static geometry describes some equilibrium mixed (though not precisely thermal) density matrix.

Let us once again examine the entanglement entropy of a sufficiently large region $\mathcal{A}$, for which the causal wedge has a hole, so that no connected extremal surface anchored

[22] The special case of extremal RN-AdS black hole with $r_- = r_+$ is somewhat more subtle, since there is no bifurcation surface (instead the spacial geometry exhibits an infinite throat). We think this is a feature rather than a bug, indicative of being at strictly zero temperature; however the HRT prescription can be applied in the same limiting fashion as is commonly done with e.g. the Wald entropy of extremal black holes.

[23] Alternately, we could require that no component of the extremal surface is allowed to lie in the causal future of $\mathcal{A}$. We thank Matt Headrick and Don Marolf for useful discussions on this point.

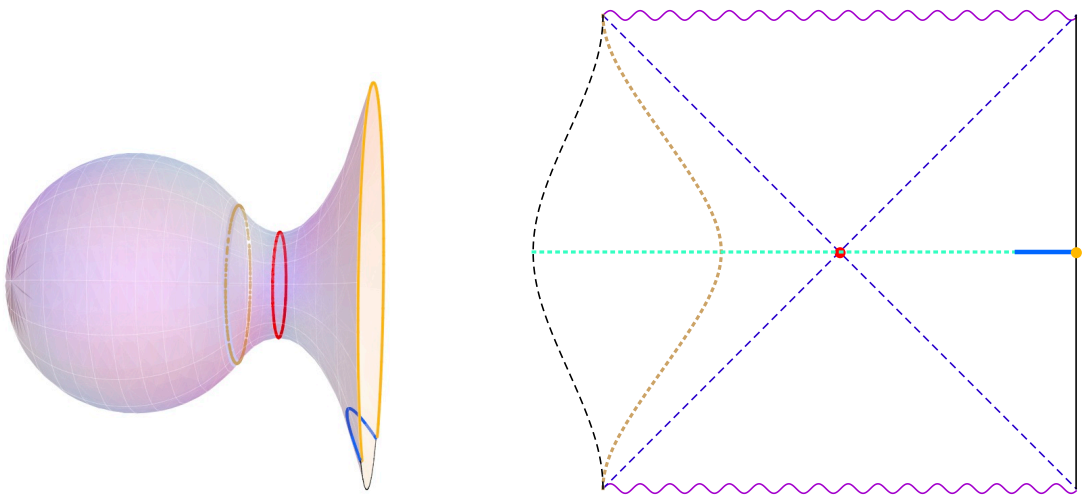


Figure 2.11: Sketch of the Schwarzschild-AdS 'bag of gold' embedding diagram (left) and Penrose diagram (right). Here the right asymptotic boundary of eternal Schwarzschild-AdS geometry is cut-off by a static shell (brown dotted curve), beyond which the spacetime caps off through a smooth origin as indicated. The embedding diagram depicts event horizon bifurcation surface (red curve) and the boundary region $\mathcal{A}$ (orange curve), at the endpoints of which is anchored the connected part of the extremal surface (blue curve). The Penrose diagram has the same conventions as in fig. 2.10, but now the homology surface $\mathcal{R}$ (dotted green line) goes all the way around the tip.

on $\partial\mathcal{A}$ can pass on $\mathcal{A}$'s side of the black hole. In the global eternal Schwarzschild-AdS spacetime, the homology constraint would then have forced us to take the connected surface on $\mathcal{A}^c$'s side of the black hole, along with the bifurcation surface on the event horizon. In the present case however, while the bifurcation surface is still an extremal surface (cf. the red curve in fig. 2.11), its inclusion in the entanglement entropy computation is no longer required by the homology constraint (even including the spacelike restriction): there exists a smooth spacelike co-dimension 1 region $\mathcal{R}$ (indicated by the dotted green line in the Penrose diagram of fig. 2.11), wrapping the bag of gold at the given instant in time, whose only boundaries are $\mathcal{A}$ and the extremal surface for $\mathcal{A}^c$ (denoted by the blue curve in fig. 2.11). In other words, in this geometry,

$$S_{\mathcal{A}} = S_{\mathcal{A}^c} \implies \delta S_{\mathcal{A}} = 0 \qquad \forall\, \theta_\infty\,. \tag{2.18}$$

We stress that while the CFT is not in the precisely thermal state, it is certainly not in a pure state either [53], so a relation of the form eq. (2.18) is wholly unexpected.[24]

In comparing the RT and HRT prescriptions for the Schwarzschild-AdS bag of gold case, we encounter a slight ambiguity in the RT prescription, namely in what is really meant by a 'constant t slice': strictly-speaking, the static coordinate patch ends at the horizon (which also follows from thinking about the Euclidean section), so the RT prescription

[24] However, as indicated in section 2.7, similar effect takes place in time-dependent situations involving black hole collapse.

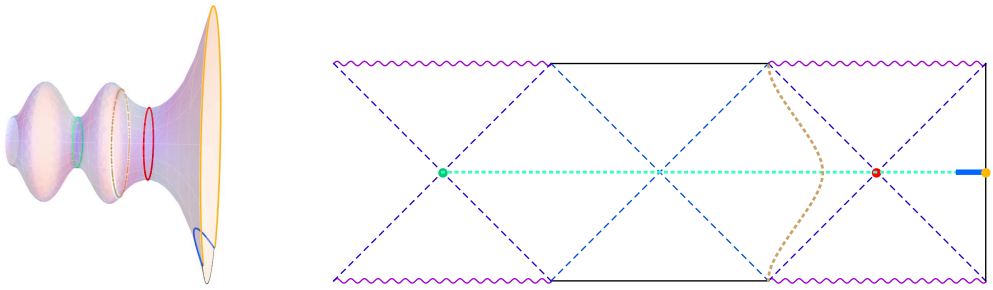


Figure 2.12: Sketch of the Schwarzschild-AdS joined to Schwarzschild-dS geometry across a shell: embedding diagram (left) and Penrose diagram (right), with same conventions as in fig. 2.11. In contrast to the capped off geometry, however, here we do have a smaller (but non-zero) area extremal surface (green circle/dot) which is required by the homology constraint.

should not see any difference between this and the eternal Schwarzschild-AdS geometry – hence predicting that $\delta S_{\mathcal{A}} = S_{\rho_\Sigma}$ for large regions. On the other hand, if one were instructed to take a full (geodesically complete) time-symmetric slice through the global geometry, one would reproduce the HRT result eq. (2.18). We also note that one could consider intermediate cases, with any $\delta S_{\mathcal{A}} \leq S_{\rho_\Sigma}$, by putting a Schwarzschild-dS geometry with smaller event horizon $\tilde{r}_+ < r_+$ on the left of the shell, as indicated in fig. 2.12. The corresponding embedding diagram for a time-reflection-symmetric slice would not cap off, necessitating the inclusion of a bifurcation surface for this Schwarzschild-dS geometry in the entanglement entropy computation, in order to satisfy the homology constraint.

We have presented several examples of static (though not globally static) asymptotically AdS geometries for which the RT and HRT proposals might differ, depending on the precise formulation of the homology constraint and of RT's constant time slice. We have not offered any definitive resolution; this is an interesting and important area which we leave for future exploration. The specific simple examples we have presented above all have some potentially dubious features: in case of Reissner-Nordstrom-AdS, the Cauchy horizon is unstable, while in the other two cases, the shell is unstable. Moreover, in all cases, the difference between the prescriptions occurs due to a part of the geometry which is beyond the horizon and thus causally inaccessible to an asymptotic observer. While this feature might therefore seem rather unappealing, we stress that in general we would be forced into such situation in any case, as long as the entanglement entropy is related to some locally-defined geometric construct, due to the teleological nature of the event horizon. The examples mentioned above merely illustrate the issues we have yet to confront to fully understand the holographic entanglement entropy prescription.

2.7 Discussion

We have focused on exploring the behaviour of entanglement entropy $S_{\mathcal{A}}$ under smooth deformations of the entangling region $\mathcal{A}$ in finite systems. Of particular interest to us is

the distinction between the behaviour of holographic field theories, which in the large c (planar limit) and strong coupling limit can be mapped onto the dynamics of classical gravity, versus field theories away from the planar limit. Sharp features in observables are possible in the latter since the planar limit allows one to enter a 'thermodynamic regime', as evidenced for example by the thermal phase transitions in finite volume [26, 68, 69].

We focused in particular on the Araki-Lieb inequality which gives a useful measure of the relative entanglement of a region $\mathcal{A}$ and its complement $\mathcal{A}^c$. While this would vanish if the entire system were in a pure state, it carries non-trivial information about the total density matrix in general. We indeed encounter an interesting phenomenon of entanglement plateaux: the Araki-Lieb inequality is saturated for finite system sizes, owing to some non-trivial features of minimal surfaces which compute the holographic entanglement entropy.

Focusing on thermal density matrices in CFTs we find that in $1+1$ dimensional field theory the Araki-Lieb inequality forces us to modify the expression of entanglement entropy for a large enough region (a point previously noted in [51, 52]) and provide an analytic prediction of when the plateau is attained. We also contrast this behaviour of large c theories against low central charge theories. There being very few exact results on the entanglement entropy of thermal CFTs in finite volume, we focused on the available expressions for Dirac fermion (with and without chemical potential). In the $c = 1$ case we note the absence of any plateau and the Araki-Lieb inequality is only saturated when the region under consideration (or its complement) is maximal. In higher dimensional holographic examples this no longer is the case, $\delta S_{\mathcal{A}}$ is forced by virtue of the features of the holographic construction to plateau (see below).

For the main part of this work, we have considered the RT prescription for calculating the holographic entanglement entropy, which is valid for static equilibrium situations. In this context, one is instructed to work at a given instant in time, and the entanglement entropy computation then involves finding the area of a requisite *minimal* surface in the corresponding Riemannian geometry. Though innocuously simple-sounding, there is a rich set of features associated with what precisely is meant by 'requisite'. For a specified boundary region $\mathcal{A}$, the boundary of the relevant surface $\mathfrak{M}_{\mathcal{A}}$ must coincide with the entangling surface $\partial\mathcal{A}$, it must be homologous to $\mathcal{A}$, and in case of multiple such surfaces, it must be the one with smallest area.

It has already been observed previously that the last restriction can cause the entanglement entropy to have a kink (i.e., its first derivative with respect to the parameter α characterizing the region $\mathcal{A}$ can be discontinuous). This is because there can exist multiple families of minimal surfaces which can exchange dominance. Here we have explored this multiplicity further, and discovered that even in the most simple case of global eternal Schwarzschild-AdS_{d+1} geometry with $d \geq 3$, in a certain regime of α there is actually an infinite tower of minimal surfaces anchored on the same entangling surface (though only the lowest two have a regime of dominance), seemingly approaching self-similar behaviour. This curious feature can arise thanks to the dimensionality of the surfaces and compactness

of the horizon. On the other hand, in other regimes of α (namely for sufficiently large region $\mathcal{A}$), there is only a single, disconnected minimal surface satisfying the homology requirement – unlike in the $2+1$ dimensional case, a connected minimal surface homologous to $\mathcal{A}$ simply does not exist. This novel feature may be understood as a consequence of certain properties of causal wedges discussed in [55].

We note in passing that the distinction between the AdS_3 and higher dimensional examples is quite reminiscent of the Hagedorn behaviour encountered in the dual field theories. In AdS_3 the BTZ black hole saddle point exists for all values of temperature (as does the thermal AdS one), but it becomes sub-dominant at low temperatures. In higher dimensions, the Schwarzschild-AdS saddles only exist above a minimum temperature T_{min}; they however take over from the thermal AdS saddle at a slightly higher temperature $T_H > T_{min}$ (when the horizon size is comparable to AdS length scale). So at low enough temperatures we are always forced into the confining state. In the context of entanglement entropy the analogous observation is about the existence of connected minimal surfaces in the black hole geometry (we should emphasize that we are always in the deconfined phase to be able to use the black hole saddle). In three dimensions, connected surfaces always exist but fail to be dominant past some critical region size; in higher dimensions they cease to exist past a critical region size given the homology constraint.

Our discussion was primarily focussed on entanglement entropy and in particular on $\delta S_{\mathcal{A}}$. One could equally have focussed on the mutual information, $I(\mathcal{A},\mathcal{B}) = S_{\mathcal{A}} + S_{\mathcal{B}} - S_{\mathcal{A}\cup\mathcal{B}}$, for two disjoint regions $\mathcal{A}$ and $\mathcal{B}$. Holographic studies of mutual information also reveal an interesting behaviour: for sufficiently separated regions $I(\mathcal{A};\mathcal{B}) = 0$. This is due to the fact that the extremal surface $\mathcal{E}\mathcal{A}\cup\mathcal{B}$ breaks up into a disconnected surface anchored on $\mathcal{A}$ and $\mathcal{B}$ to avoid thin necks. This has been extensively discussed in the literature cf., [46] and [63, 65, 70] for implications of this construct in the context of probing black hole interiors. For the regions we have considered $\mathcal{A}$ and $\mathcal{A}^c$ are not strictly disjoint, so we have to be a bit more careful; naively the mutual information involves the sum of the entropies of two regions (which we take to be $\mathcal{A}$ and $\mathcal{A}^c$) while the Araki-Lieb inequality constrains their difference, giving $I(\mathcal{A};\mathcal{A}^c) = 2\,S_{\mathcal{A}}$ which is UV divergent.[25] One can however constrain the mutual information between the smaller of $\mathcal{A}$ or $\mathcal{A}^c$ and the auxiliary system $\hat{\rho}_\Sigma$ which purifies the global density matrix ρ_Σ (thus $S_{\rho_\Sigma} = S_{\mathcal{A}\cup\mathcal{A}^c} = S_{\hat{\rho}_\Sigma}$) to vanish when Araki-Lieb inequality is saturated. For example one has

$$I(\mathcal{A};\hat{\rho}_\Sigma) = S_{\mathcal{A}} + S_{\hat{\rho}_\Sigma} - S_{\mathcal{A}\cup\hat{\rho}_\Sigma} = S_{\mathcal{A}} + S_{\rho_\Sigma} - S_{\mathcal{A}^c} = \begin{cases} 0, & \text{for } \alpha < \frac{1}{2} \\ 2\,S_{\rho_\Sigma}, & \text{for } \alpha > \frac{1}{2} \end{cases} \tag{2.19}$$

In deriving this we used the fact that $S_{\mathcal{A}\cup\mathcal{A}^c\cup\hat{\rho}_\Sigma} = 0$ by definition of the purifying degrees of freedom. In the case of the global thermal density matrix $\hat{\rho}_\Sigma$ is the thermofield double which lives on the second asymptotic region of the eternal black hole Penrose diagram.[26]

[25] We thank the referee for spotting an error in our original statement and for suggesting use of the auxiliary purifying system.

[26] It is worth mentioning that the saturation of the Araki-Lieb inequality in large c theories implies that

While these causal wedge considerations already by themselves guarantee that the entanglement entropy $S_{\mathcal{A}}(\alpha)$ cannot be a smooth function for thermal states in higher dimensions, we have argued that, if defined by a minimization procedure as in the RT prescription, the entanglement entropy must nevertheless be continuous. Our argument crucially used minimality. In particular, in the HRT prescription of computing entanglement entropy via a smallest area *extremal* surface, the continuity argument given in §2.2 does not apply. It would be interesting to explore whether sufficient time-dependence can provide counter-examples to continuity or whether the one can generalize the proof of continuity to the Lorentzian context; we leave further investigation of this issue for the future.

Note that in this collapsed black hole context, the homology requirement is satisfied automatically, though as explained in [72], that itself leads to very curious feature of entanglement entropy: it can distinguish between an eternal black hole and a collapsed one, arbitrarily long after the collapse had taken place [73]. In this sense, while in the AdS/CFT context we are used to classical bulk surfaces providing at best only some coarse-grained CFT information, the homology requirement induces a fine-grained aspect to the entanglement entropy observable.

While in the time-dependent setting mentioned above, the RT prescription is meaningless and therefore cannot be compared to the extremal surface generalization of HRT, we have seen that there are subtle differences between the RT and HRT proposals even is static situations. Indeed, a similar 'fine-grained' quality as mentioned above is manifested in the entanglement entropy for e.g. the Schwarzschild-AdS bag of gold geometry discussed in section 2.6: there the entanglement entropy (as given by the HRT prescription) can easily distinguish the differences in the geometry behind the horizon, on the other side of the Einstein-Rosen bridge. Other known diagnostics of the CFT density matrix not being precisely thermal typically require rather sophisticated machinery, such as detecting lack of periodicity on the Euclidean time circle of the analytically continued solution. We find it remarkable that entanglement entropy does the job so easily. This observation of course crucially hinges on the precise formulation of the homology constraint, which deserves to be understood better.

2.A Mean Curvature Flow

In this appendix we summarize the algorithm we used to construct the minimal surfaces in Schwarzschild-AdS_{d+1} for $d \geq 3$. While one can directly numerically solve the Euler-Lagrange equations of (2.15) directly, it is especially useful to consider alternate strategies to ensure that we have obtained all the surfaces of interest.

For conceptual and computational purposes, it is useful to view minimal surfaces as the endpoint of a flow of surfaces of decreasing area. This is analogous to a gradient

the degrees of freedom in $\mathcal{A}$ admit a canonical split into two uncorrelated parts: one that carries all the entanglement with $\mathcal{A}^c$ and another that carries the macroscopic entropy of ρ_Σ, cf., [71]. We thank Matt Headrick for suggesting this interpretation and other comments regarding Araki-Lieb inequality saturation.

descent algorithm for finding a local minimum of a function, for which one considers a point moving with velocity equal to minus the gradient of the function. This gives a curve along which the function monotonically decreases, and exponentially decays to the minimum as a function of flow time $t_{flow} \equiv \mathtt{t} \to \infty$.

In the analogous process for minimal submanifolds, the function to be minimized becomes the area functional, and the flow velocity vector becomes a vector field (which may be taken to be normal) defined on the submanifold. The appropriate vector field is given by the *mean curvature* of the submanifold, and the evolutionary process is known as mean curvature flow.

We review the required differential geometric technology, and outline an algorithm for computing the process in the case of codimension-one minimal surfaces.[27]

2.A.1 Some geometry

Let $\mathfrak{n}$ be a smooth n-dimensional Riemannian manifold (e.g., $\mathfrak{n}$ for situations of interest is a constant time surface in the bulk spacetime $\mathcal{M}$), with metric g and Levi-Civita connection ∇. Further, let $\mathfrak{m} \subset \mathfrak{n}$ be a smooth m-dimensional submanifold (with $m < n$), defined by embedding $\psi : \mathfrak{m} \to \mathfrak{n}$. The induced metric on $\mathfrak{m}$ is inherited from $\mathfrak{n}$ as $\gamma = \psi^* g$ given this embedding. Vectors at a point on the submanifold can be decomposed into components tangent and normal to $\mathfrak{m}$: $U = U^\top + U^\perp$.

The natural structures associated with $\mathfrak{m} \subset \mathfrak{n}$ are in terms of the intrinsic and extrinsic geometry of the embedding, characterized by the induced metric γ (the *first fundamental form*) and the extrinsic curvature K (the *second fundamental form*). Formally, given vector fields, say U and V, tangent to $\mathfrak{m}$, one can decompose the covariant derivative:

$$\nabla_U V = (\nabla_U V)^\top + (\nabla_U V)^\perp = \mathcal{D}_U V + K(U, V)\,, \tag{2.20}$$

which effectively defines $\mathcal{D}$ and K. The former is the Levi-Civita connection of the induced metric on $\mathfrak{m}$, while the latter is a normal-valued symmetric bi-linear form. It is worth emphasizing that the second fundamental form depends only on the vectors at the point, so it is a tensor on $\mathfrak{m}$. A particularly important object is obtained by taking its trace $\mathtt{k}$, giving a normal vector field on $\mathfrak{m}$: this is the *mean curvature field*.[28] In components, using coordinates x^a intrinsic to $\mathfrak{m}$, and X^α on $\mathfrak{n}$, we have $\gamma_{ab} = g_{\alpha\beta}\, \partial_a X^\alpha\, \partial_b X^\beta$, and

$$h^{\alpha\beta} = \gamma^{ab}\, \partial_a X^\alpha\, \partial_b X^\beta\,, \qquad K_{\alpha\beta}{}^{\gamma} = h_\alpha{}^\mu\, h_\beta{}^\nu\, \nabla_\nu\, h_\mu{}^\gamma\,, \qquad \mathtt{k}^\alpha = h^{\mu\nu}\, K_{\mu\nu}{}^{\alpha}.$$

Here h is the projection onto vectors tangent to $\mathfrak{m}$. Note that the extrinsic curvature tensor $K_{\alpha\beta}{}^{\gamma}$ is symmetric in its lower indices which lie tangent to $\mathfrak{m}$, while the upper index is

[27] The minimal surfaces relevant for computing entanglement entropy using the prescription of [30] are co-dimension two in the bulk. In static spacetimes however we can localize on a constant time bulk hypersurface, leading thus to a search of co-dimension one surfaces on the slice.

[28] This definition of the mean curvature is not entirely universal: in some conventions the sign differs, while in others the trace is divided by a factor of m, so that $\mathtt{k}$ is genuinely the mean of the eigenvalues of K (which incidentally define the principal curvatures of the submanifold).

transverse.

We are now in a position to make precise the geometric interpretation of the mean curvature field alluded to above, that it is minus the gradient of the area functional on the space of submanifolds.

Let us consider a smooth family of embeddings $\Psi : I \times \mathfrak{m} \to \mathfrak{n}$, where I is some interval parameterized by a flow time $\mathtt{t}$. For each $\mathtt{t} \in I$, $\psi_\mathtt{t} = \Psi(\mathtt{t}, \cdot) : \mathfrak{m} \to \mathfrak{n}$ is an embedding. Let A be the area of the embedded submanifold:

$$A(\mathtt{t}) = \int_\mathfrak{m} \omega_\mathtt{t}, \tag{2.21}$$

where $\omega_\mathtt{t}$ is the induced volume form on $\mathfrak{m}$ at time $\mathtt{t}$.

Let V be the vector field given by the local velocity of the surface in this deformation, so $V = \Psi_* \frac{\partial}{\partial \mathtt{t}}$. One should think of Ψ, or $\psi_\mathtt{t}$, as defining a curve through the space of embeddings, an infinite-dimensional manifold whose tangent space at ψ is given by the set of $T\mathfrak{n}$-valued vector fields on ψ. Then V is naturally the tangent vector to the curve.

We now have enough state our key result:

$$\frac{\mathrm{d}A}{\mathrm{d}\mathtt{t}} = -\int_\mathfrak{m} g(\mathtt{k}, V)\, \omega_\mathtt{t}. \tag{2.22}$$

In the space of embeddings, this is the natural inner product between V and $-\mathtt{k}$, which makes precise the analogy with a gradient. Note that, as is to be expected, only the normal component of V contributes, since the tangent component is 'pure gauge', describing how parameterization of the embedding evolves, and not the shape of the submanifold itself.

Proof. The method of proof is to compute $\frac{\partial}{\partial \mathtt{t}}$ of the induced volume form on $\mathfrak{m}$, and we will show it is equal to $(-g(\mathtt{k}, V) + \mathcal{D} \cdot V^\top)\omega_\mathtt{t}$. Since the last term is a total derivative, the result will follow, provided that the variation is constrained to a compact region, or we restrict the variation to be normal (so $V^\top = 0$). For ease of computation, we work in a chart of normal coordinates x^i on $\mathfrak{m}$ at the point p, with respect to the induced metric at time $\mathtt{t} = 0$: $\gamma(\partial_i, \partial_j) = \delta_{ij}$, and $\mathcal{D}_{\partial_i}\partial_j = 0$. Let $e_i = \Psi_* \partial_i$. For later use, we note that since Lie brackets commute with push-forward, $[e_i, V] = 0$, so $\nabla_V e_i = \nabla_{e_i} V$.

The induced volume form is given by $\omega_\mathtt{t} = \sqrt{\det(g(e_i, e_j))}\ \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m$, so its derivative at the point p and $\mathtt{t} = 0$ is

$$\frac{1}{2}\frac{\partial}{\partial \mathtt{t}}(g(e_i, e_i))\ \omega_0, \tag{2.23}$$

where we have used the fact that the derivative of $\det A$ is $\det A \operatorname{tr}(A^{-1}\dot{A})$, and that $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m = \omega_0$ at p in our normal coordinates. Now, again computing at the point p

and $\mathtt{t} = 0$,

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial \mathtt{t}} g(e_i, e_i) &= \frac{1}{2} V(g(e_i, e_i)) \\ &= g(\nabla_V e_i, e_i) \\ &= g(\nabla_{e_i} V, e_i) \\ &= e_i(g(V, e_i)) - g(V, \nabla_{e_i} e_i) \\ &= \partial_i V_i^\top - g(V, \mathcal{D}_{e_i} e_i + K(e_i, e_i)) \\ &= \mathcal{D} \cdot V^\top - g(V, \mathtt{k}).\end{aligned}$$

In the final line we have used the fact that we have normal coordinates, so $\partial_i = \mathcal{D}_i$, and the sum over i reduces to the trace. □

2.A.2 The algorithm

With the necessary mathematical formalism in hand, we now describe an algorithm for implementing the mean curvature flow, specializing for simplicity to the case of codimension-one surfaces. In this case, having chosen a direction for a normal, the mean curvature vector field reduces to a scalar field, since the normal space is one-dimensional.

The method used is based on a level-set approach: the surfaces are described by the sets

$$\mathfrak{m} = \{x \in \mathfrak{n} |\ \ \phi(x) = c\} \tag{2.24}$$

for some function Φ and some constant c, and the evolution of the function Φ is what we shall model. This has several advantages:

- There is no requirement to pick a parameterization of the surface, so no problem with 'gauge fixing'.
- A whole family of surfaces can be modeled at once, by computing the evolution for Φ and picking different level sets.
- The algorithm elegantly allows for changes in topology: under mean curvature flow, surfaces can split into several parts, and this causes no trouble for the level set method.

The definition of the flow gives the normal speed of a point on the surface as it evolves: $n(\frac{\mathrm{d}x}{\mathrm{d}\mathtt{t}}) = \mathtt{k}$, with n being the normal to $\mathfrak{m}$. Provided that $\mathrm{d}\Phi \neq 0$, we also have a simple way of computing a unit normal, $n = \frac{\mathrm{d}\Phi}{\|\mathrm{d}\Phi\|}$. To deduce the evolution of Φ, differentiate the definition (2.24) of a surface $\mathfrak{m}$:

$$0 = \frac{\partial \Phi}{\partial \mathtt{t}} + \mathrm{d}\Phi\big(\tfrac{\mathrm{d}x}{\mathrm{d}\mathtt{t}}\big) = \frac{\partial \Phi}{\partial \mathtt{t}} + \|\mathrm{d}\Phi\|\, n\big(\tfrac{\mathrm{d}x}{\mathrm{d}\mathtt{t}}\big) = \frac{\partial \Phi}{\partial \mathtt{t}} + \|\mathrm{d}\Phi\|\, \mathtt{k}\,.$$

The only ingredient which remains is to compute $\mathtt{k}$ using only information from Φ. Luckily, this is given by a simple formula in terms of the normal: $\mathtt{k} = -\nabla \cdot n$. To see this, take n

to be a unit one-form, normal to the surface, but also defined, and unit, off it. Extend to an orthonormal basis[29] $\{e_i, n^\#\}$. Then:

$$\begin{aligned}\mathtt{k} &= n(\nabla_{e_i} e_i) = e_i(n(e_i)) - (\nabla_{e_i} n)(e_i) \\ &= -\nabla\cdot n + (\nabla_{n^\#} n)(n^\#) \\ &= -\nabla\cdot n + g(n, \nabla_{n^\#} n) \\ &= -\nabla\cdot n + \frac{1}{2} n^\#(g(n,n)) = -\nabla\cdot n.\end{aligned}$$

The term $e_i(n(e_i))$ in the first line vanishes because $n(e_i) = 0$ everywhere, and the last equality follows because $g(n,n) = 1$ everywhere.

This gives us our final evolution equation for Φ:

$$\frac{\partial\Phi}{\partial \mathtt{t}} = \|\mathrm{d}\Phi\|\, \nabla\cdot\left(\frac{\mathrm{d}\Phi}{\|\mathrm{d}\Phi\|}\right) = \sqrt{g^{\gamma\delta}\,\partial_\gamma\Phi\,\partial_\delta\Phi}\;\frac{1}{\sqrt{g}}\;\partial_\alpha\left(\frac{\sqrt{g}\,g^{\alpha\beta}\,\partial_\beta\Phi}{\sqrt{g^{\gamma\delta}\,\partial_\gamma\Phi\partial_\delta\Phi}}\right). \tag{2.25}$$

This is a quasilinear parabolic PDE in Φ, which can be solved numerically by standard methods.

For the computations in this paper, only spherically symmetric surfaces were considered, so Φ was a function of two spatial coordinates θ and r (or another suitably redefined radial coordinate). The PDE was solved with difference methods, with Neumann boundary conditions at the poles of the sphere ($\theta = 0, \pi$) to keep the surfaces regular there, and mixed boundary conditions at radial boundaries set by the desired asymptotics.

2.B Near-horizon behaviour of minimal surfaces

An analytic solution of the equations of motion in a near-horizon approximation explains the existence of an infinite self-similar family of minimal surfaces, wrapping the horizon multiple times. We find that this is generic near any minimum of the radius of the $\mathbf{S}^{d-1}$, such as the Einstein-Rosen bridge of a non-extremal black hole, and is characterized entirely by the number of dimensions,[30], and a single parameter $\xi \equiv r_+\,\kappa$, where κ is the surface gravity (which is related to the black hole temperature as $T = \frac{1}{4\pi}\,\kappa$).

For the analysis we use a modified radial coordinate z, regular at the event horizon, defined by $r = r_+ + \frac{1}{2}\,r_+\,\xi\,z^2$, and parametrize the surface by θ throughout. Dots will denote differentiation with respect to θ.

We invoke two approximation schemes. The naive linearization holds away from the poles, but breaks down close the $\theta = 0, \pi$ where $\dot{z}$ grows without bound. Near the poles, we use a different approximation, reducing the problem locally to flat Euclidean space. We will find that these two approximation schemes have overlapping domains of validity, so the complete behaviour can be described by matching the asymptotics in each regime. This

[29] We denote the 'raised index' normal vector by $n^\#$, so that $g(n^\#, U) = n(U)$

[30] Here we will keep to spatial codimension-one surfaces, but the argument carries over essentially unchanged for surfaces of any dimension m, wrapping a $\mathbf{S}^{m-1}$ in an $SO(d)$ symmetric space

predicts a tower of arbitrarily many branches of the surface wrapping the horizon, with a known relationship between the distances from the horizon of successive surfaces.

Schematically: we start at the North pole $\theta = 0$, and solve in the linear regime towards the South pole $\theta = \pi$. There we change to the near-pole regime, in which there will be a turning point at a minimum angle Δ away from the pole, at $z = z_0$. These constants are both computed from matching to the linearized solution. We then match up to another branch moving back towards the North pole, where there will be a turning point given by new, parametrically larger, Δ and z_0. The matching process will allow us to produce a recurrence relation for the sequences Δ_n, z_n of turning points. We can obtain an entirely analytic expression for these values, which, sufficiently close to the horizon, match precisely with the values obtained from numerically integrated surfaces.

2.B.1 The linear regime

We start with the linear approximation, valid for $z, \dot{z} \ll 1$. In this case, the minimal surface equation obtained from (2.15) reduces to

$$\ddot{z} + (d-2)\cot\theta\,\dot{z} - (d-1)\,\xi\,z = 0\,, \qquad \xi \equiv r_+\,\kappa. \tag{2.26}$$

This equation has regular singular points at $\theta = 0, \pi$, where the solutions have asymptotics $z \sim \theta^\sigma$ for $\sigma = 0, 3-d$ (or constant and $z \sim \log\theta$ asymptotics in $d = 3$). For $\xi > 0$ it is easy to show that a solution regular at $\theta = 0$ cannot be regular at $\theta = \pi$, so the general solution can be written

$$z(\theta) = A\; z_r(\theta) + B\; z_r(\pi - \theta) \tag{2.27}$$

using the symmetry under $\theta \mapsto \pi - \theta$. Here z_r is chosen by demanding $z_r(\theta) \to 1$ as $\theta \to 0$, and the asymptotics are characterized by a parameter λ, by $z_r(\theta) \sim \lambda\,(\pi - \theta)^{3-d}$ as $\theta \to \pi$. In fact, the solutions can be found explicitly in terms of associated Legendre functions, and $\lambda(d, \xi)$ can be found analytically.

2.B.2 Near-pole flat space regime

For the second approximation scheme, we assume $z \ll 1, z \ll \dot{z}, \theta \ll 1$, in which case the problem reduces to that of a spherically symmetric "soap bubble" in flat d-dimensional Euclidean space, where (θ, z) are the radial and longitudinal coordinates respectively in a cylindrical polar system. This has the solution

$$z = z_0 \pm \Delta \int_1^{\frac{\theta}{\Delta}} \frac{\mathrm{d}x}{\sqrt{x^{2d-4} - 1}}. \tag{2.28}$$

The constants z_0 and Δ will be small values, describing the distance from horizon and pole respectively of the turning point of the surface. Note that taking the limit $\theta \to \infty$, the integral converges, so z tends to a fixed value when $d > 3$: the upper and lower branches

of the surface remain at bounded separation, characterized by the value of

$$I_d = \int_1^\infty \frac{\mathrm{d}x}{\sqrt{x^{2d-4}-1}} = -\sqrt{\pi}\,\frac{\Gamma\left(\frac{d-3}{2(d-2)}\right)}{\Gamma\left(\frac{-1}{2(d-2)}\right)}. \tag{2.29}$$

The asymptotics are given by

$$z \sim z_0 \pm I_d\,\Delta \mp \frac{1}{d-3}\,\Delta^{d-2}\,\theta^{3-d} \quad \text{as} \quad \frac{\theta}{\Delta} \to \infty\,. \tag{2.30}$$

2.B.3 Matching asymptotics

With solutions in the two approximation schemes in hand, it now remains only to fix constants of integration by matching the asymptotics. For definiteness, consider starting from a known (Δ_n, z_n) close to the North pole, computing the coefficients of the linear solution, and finally extracting the values of (Δ_{n+1}, z_{n+1}) close to the South pole.

The case $d = 3$ is a little different from $d > 3$, because of the logarithmic asymptotics, so we will stick to $d > 3$ here. A very similar analysis holds in this case, for which we will omit the details, and present only the results.

Recall the form of the linear solution $z(\theta) = A\,z_r(\theta) + B\,z_r(\pi-\theta)$, with the terms regular at $\theta = 0, \pi$ respectively. Consistency imposes that the coefficients have parametrically different sizes, with $A \gg B$. Towards the South pole, the B-mode is also suppressed relative to the A-mode by the growth of the latter, which renders B essentially irrelevant, and the turning points can be characterized by a single parameter. Indeed, by matching the asymptotics near the South pole, for consistency we need $z_{n+1} = I_d\,\Delta_{n+1}$. Finally, we match A with the relevant coefficients at each end,

$$z_n + I_d\,\Delta_n = A = \frac{\Delta_{n+1}^{d-2}}{\lambda\,(d-3)}, \tag{2.31}$$

which finally gives us our recurrence relation

$$\Delta_{n+1} = \left[2\,\lambda\,(d-3)\,\Delta_n\right]^{\frac{1}{d-2}}, \tag{2.32}$$

demonstrating a parametric growth in Δ, z at each step. This can be conveniently solved by changing to the coordinate x, defined by $r = r_+ + e^{2\,x}$, (or $\frac{1}{2}\,r_+\,\xi\,z^2 = e^{2\,x}$) from which we get

$$x_n = \frac{x_0 - \nu}{(d-2)^n} + \nu, \qquad \Delta_n = \sqrt{\frac{2}{\kappa}}\,\frac{e^{x_n}}{r_+\,I_d} \tag{2.33}$$

with the constant ν given by

$$\nu = \frac{1}{2}\,\log\left(\frac{r_+\,\xi}{2}\right) + \frac{\log[2\,(d-2)\,\lambda\,I_d^{d-2}]}{d-3}. \tag{2.34}$$

It should be noted that the initial value of x where the surface smoothly crosses the pole

$\theta = 0$, closest to the horizon, is not at x_0, but rather at $x = x_0 + \log 2$, since the solution here is not given by (2.28), but by that obtained in the linear regime: constant x to leading order.

For completeness, we record the computed values of λ here:

$$\lambda = \frac{\cosh(\pi \jmath)}{(d-3)\,\pi} \; \frac{2^{d-2}\left(\left(\frac{d-3}{2}\right)!\right)^2}{[(\frac{1}{2})^2+\jmath^2][(\frac{3}{2})^2+\jmath^2]\cdots[(\frac{d-4}{2})^2+\jmath^2]}\,, \qquad d \text{ odd}$$
$$\lambda = (d-3)\,\frac{\sinh(\pi\jmath)}{\jmath}\left(\frac{1}{1^2+\jmath^2}\right)\left(\frac{3}{2^2+\jmath^2}\right)\left(\frac{5}{3^2+\jmath^2}\right)\cdots\left(\frac{d-5}{\left(\frac{d-4}{2}\right)^2+\jmath^2}\right), \qquad d \text{ even} \tag{2.35}$$

where in both cases, $\jmath$ is defined by

$$\jmath^2 = (d-1)\,\xi - \frac{(d-2)^2}{4}. \tag{2.36}$$

In $d = 3$, we have slightly altered formulae:

$$\Delta_{n+1} = -2\,\lambda\,\Delta_n\,\log\left(\frac{\Delta_n}{2}\right) \tag{2.37}$$
$$x_n = \frac{1}{2}\log\left(\frac{r_+\,\xi}{2}\right) + \log\left[-\Delta_n \log\left(\frac{\Delta_n}{2}\right)\right] \tag{2.38}$$
$$\lambda = \frac{2\cosh(\pi\jmath)}{\pi}\,. \tag{2.39}$$

2.B.4 Validity

Having done this calculation, it is worth checking to see that we have done something sensible: there must be some overlap between the domains of validity of each approximation. It is straightforward to confirm this *a posteriori*, looking at overlaps near the North pole for definiteness. Firstly, the linear approximation is valid when $z, \dot{z} \ll 1$, which requires θ^{2-d} to be much smaller than the coefficient of the growing mode. This holds for $\theta \gg \Delta$. Secondly, the near-pole approximation is valid when $z \ll \dot{z}$, which requires $\theta^{d-2} \ll \Delta^{d-3}$ (as well as $\theta \ll 1$). This means that as long as Δ remain small, both approximations are simultaneously valid for

$$\Delta \ll \theta \ll \Delta^{\frac{d-3}{d-2}}. \tag{2.40}$$

For $d = 3$, we have a similar result, with overlap region

$$e^{-\frac{1}{\Delta}} \ll \theta \ll \frac{-1}{\log\Delta}\,. \tag{2.41}$$

2.B.5 Relation to θ_∞

We can straightforwardly use this analysis to show that for θ_∞ sufficiently close to $\pi/2$ there will be an infinite tower of minimal surfaces. This is because each branch of a surface is characterized by a single parameter. All surfaces with a particular turning point will

match closely outside that turning point. To obtain a surface on later branches of this tower requires the initial value of $-x$ to be exponentially large, so the distance from the horizon will be suppressed by an exponential of an exponential.

Incidentally, this gives an indirect demonstration of non-existence of connected minimal surfaces for sufficiently large θ_∞. The connected surfaces form a one-parameter family, parameterized by the maximum depth reached into the bulk, or equivalently the location x_0 of the surface at $\theta = 0$. Considering θ_∞ as a function of x_0, continuity implies that there are only two ways for θ_∞ to reach arbitrarily close to π. The first is for $\theta_\infty = \pi$ for some x_0, but this is precluded from happening by behaviour near the boundary. The other way is for a sequence x_m to exist such that $\theta_\infty(x_m)$ tends to π, but the self-similarity properties of surfaces indicated here will prevent this option also.

Chapter 3

Geometric probes of black hole formation

This chapter is a reproduction of the paper *Holographic probes of collapsing black holes* [2], written in collaboration with Veronika Hubeny.

It is striking that certain field theory observables, such as the entanglement entropy, are computed from very natural geometric objects. This observation suggests that they may be useful to help to understand how a bulk geometry emerges from a field theory, and even in principle to reconstruct that geometry from field theory data alone, without recourse to the bulk dynamics. But it may be the case that parts of the geometry are inaccessible to such geometric observables. For example, in static geometries a horizon provides a barrier to extremal surfaces. The paper aims to characterise how this is altered in a more physical situation, where the black hole forms dynamically. The question we therefore pose is "How far can geometric probes see inside a collapsing black hole?".

We addressed this by looking at a specific toy model. The spacetime we chose to look at was global Vaidya-AdS, describing the collapse of a spherical null shell. This is a simple model of a rapid quantum quench and thermalisation process in a field theory in finite volume, living on a sphere. We focus on the observables associated with geodesics and codimension two extremal surfaces, associated with correlation functions and entanglement entropy respectively.

The results show a fairly rich structure, dependent on the dimension of spacetime and of the probes themselves. The situation is particularly intricate for geodesics in four or more spacetime dimensions, and in particular, geodesics breaking the symmetry of swapping endpoints are important. An important aspect of the geometry, controlling the behaviour of some codimension two surfaces, is a 'radius of maximal area' within the event horizon. Here, two effects, of the surface becoming null at the horizon, and of it shrinking in the transverse direction, are balanced. This results in surfaces that are anchored on the boundary at arbitrarily late times, but hover close to this critical radius inside the horizon (though these are eventually of subdominant area).

Apart from the reach of the geometric objects themselves, the evolution of the field

theory quantities themselves is also interesting. A naïve interpretation of the geodesic lengths leads to particularly strange results, the thermalisation appearing to be neither monotonic nor continuous. This likely indicates that the geodesic approximation of correlation functions is inapplicable.

3.1 Introduction and summary

The gauge/gravity correspondence[1] has proved invaluable in providing useful insights into the behaviour of strongly coupled field theories, yet the converse quest of using the field theory to understand quantum gravity in the bulk is still far from reaching its fruition. Already from the outset, one of the key obstacles is our incomplete understanding of bulk locality. Questions of how the field theory encodes bulk geometry and causal structure, or how it describes a local bulk observer, remain opaque despite intense efforts of the last 15 years.[2] While the scale/radius duality provides us with valuable intuition in the asymptotic bulk region, the mapping becomes far more obscure deeper in the bulk, and inapplicable for bulk regions which are causally separated from the boundary. The question *'how does the gauge theory see inside a bulk black hole?'* has been foremost from the start.

Causal considerations aside, the context of black hole geometry holds a particularly sharp testing ground for quantum gravity, as the curvature singularity inside a black hole is a region near which classical general relativity breaks down. Nevertheless, the gauge theory contains the full physics – it understands what resolves or replaces this classical singularity in the bulk. But for learning the answer of the gauge theory, we first need to understand what question to ask in that language: in what field theoretic quantity can we isolate the near-singularity behaviour?

One of the approaches aimed at elucidating the encoding of bulk geometry in the dual field theory observables was recently undertaken in [61], which explored how much of the bulk spacetime is accessible to certain field theory quantities related to specific geometrical probes in the bulk. In particular, [61] focused on field theory probes characterised by bulk geodesics and more general extremal surfaces, anchored on the AdS boundary. Being geometrical by nature, such probes are well-suited for decoding the bulk geometry,[3] at least at the classical level. At the same time, they are related to well-defined CFT quantities: certain types of correlators for spacelike geodesics or bulk-cone singularities for null geodesics [88]; Wilson-Maldecena loops for 2-dimensional surfaces [91, 92]; and entanglement entropy for codimension-two surfaces [30, 31, 39]. Hence, postponing for

[1] For definiteness we'll mostly focus on the prototypical case of the AdS/CFT correspondence [24] which relates the four-dimensional $\mathcal{N} = 4$ Super Yang-Mills (SYM) gauge theory to a IIB string theory (or supergravity) on asymptotically $AdS_5 \times S^5$ spacetime.

[2] Early investigations of bulk locality from various perspectives include [74, 75, 76, 77, 78, 79] whereas more recent developments and reviews are given in e.g. [80, 81, 82, 83, 84, 85, 86, 87].

[3] For example, following [88], [89, 90] demonstrated that one can reconstruct the bulk metric for an arbitrary static and spherically symmetric bulk geometry, simply from knowing the proper length/area/volume of such surfaces along with where they end on the boundary.

the moment the discussion of the subtleties of the actual relation between these CFT 'observables' and the corresponding bulk geometrical constructs, we will follow the approach of [61] in asking how much these bulk geometrical quantities, i.e. extremal surfaces anchored on the AdS boundary, know about the geometry, now specifically focusing on the black hole interior.

Perhaps the most intriguing result of [61] was that extremal surfaces anchored on the boundary of AdS cannot penetrate through the horizon of a static black hole. Of the several arguments provided, the most general of these, which analysed the equation of motion near its turning point, applies to an extremal surface of any dimension, anchored on any shape of simply-connected boundary region, in any static asymptotically-AdS spacetime with planar symmetry and event horizon. Nevertheless, as emphasised in [61], extremal surfaces are able to penetrate the event horizon of a dynamically-evolving black hole. This is simply because the event horizon is a global construct whose location depends on the full future evolution, whereas the location of extremal surfaces is determined by the local geometry.

Indeed, this observation formed the basis of [93] which used it to argue that the event horizon by itself is not an obstruction to precursor-type CFT probes. In particular, [93] presented a simple gedanken experiment wherein a thin null shell implodes from the AdS boundary and forms a large black hole. The bulk geometry is pure AdS to the past of the shell and Schwarzschild-AdS to its future; however the horizon generators (outgoing radial null geodesics which define the late-time static horizon) originate at the center of AdS prior to the shell. In particular, a bulk constant-time[4] slice, anchored on AdS boundary shortly before the creation of the shell, passes through the AdS region enclosed by event horizon. Spacelike geodesics as well as higher-dimensional extremal surfaces in AdS which are anchored at this time will lie along the same time slice, and will therefore penetrate the black hole as long as their anchoring region is sufficiently large.

However, it not clear in this example what bulk regions remain inaccessible to such probes. In particular it is unclear whether it is possible to probe past the event horizon in the more genuine black hole geometry to the future of the shell, and to penetrate near the curvature singularity, and therefore to be useful in addressing the most interesting question of what happens there.[5] Nevertheless, this argument makes it clear that we should be able to use extremal surfaces to penetrate the horizon even after the shell has formed the black hole, as long as the black hole is still evolving. This is the question we set out to explore: *how deep into the collapsed black hole, and especially how close to the curvature singularity, can extremal surfaces penetrate?*

Although obtaining strongly time dependent black hole solutions in general relativity is typically a daunting process due to the non-linearity of Einstein's equations, there are certain solutions with sufficient symmetry which are known analytically. Perhaps the simplest and best-known of these is the Vaidya (in our gauge/gravity context, Vaidya-AdS)

[4] In a static part of the spacetime, there is a geometrically unique bulk foliation by 'constant time' surfaces that are anchored at a fixed boundary time.

[5] In fact, it would even be interesting to sample the late-time horizon itself, in the recently-explored context of firewalls [94, 95], where semiclassical physics breaks down already at the horizon.

class of solutions. These describe a spherically symmetric collapsing null dust, where we are free to specify the radial (or equivalently temporal) profile of the shell. Early-time geometry (inside or before the shell) is pure AdS, while at late times (outside or after the shell), it is Schwarzschild-*AdS*. The 'dust particles' making up the shell follow ingoing radial null trajectories, so the black hole forms maximally rapidly. This is particularly useful in the present context: since we seek a feature which is absent in static geometries, we are more likely to see a large effect for geometries which are as far away from being static as possible.

There is another motivation for probing duals of such collapse geometries coming from field theory. Since a large black hole in the bulk corresponds to a thermal state in the dual field theory, collapse to a black hole describes the process of thermalization. Moreover, if the collapse is rapid, the dynamics describes a far-from-equilibrium process. While we typically have a good handle on equilibrium situations, out of equilibrium processes are more interesting but far less understood. Sudden changes in the field theory Hamiltonian, known as "quantum quenches", and subsequent equilibration have been much studied in field theory, and have recently received mounting attention from holographic studies. The analysis of thermalization using (global, i.e. spherically symmetric) Vaidya-AdS as toy-model for quantum quench initiated in [39] was extended in the planar case by [62, 96] in 3-dimensional bulk, by [97] in 4 dimensions, by [98, 99] in 3,4 and 5 dimensions (the latter having used entanglement entropy as well as equal time correlators and Wilson loop expectation values), and more recently by [64, 100] with more general considerations.[6] While most of these works focused on the thermalization aspect, less attention was paid to the question of how much of the collapsing black hole can such probes access, along the lines motivated by [61]. Hence, apart from the interest in further exploration of holographic thermalization, the question of probing inside the black hole motivates us to continue the study of (global) Vaidya-AdS, using spacelike geodesics and codimension-two extremal surfaces as probes. The use of these probes was more fully justified in [61] and many of the references mentioned above; here we simply employ the same rationale in exploring them further.

Perhaps the greatest novelty in our findings stems from the fact that, motivated by creating a black hole with compact horizon, we are working with asymptotically globally AdS spacetimes, rather than the (geometrically simpler and more often studied) asymptotically Poincaré AdS spacetimes. While the global case includes the planar Poincaré case as a special limit, the converse is not true: the possibility of geodesics and surfaces which can 'go around' a spherical black hole allows for a vastly richer structure. This was evident already in the recent study [1] involving extremal surfaces in the static spherical Schwarzschild-AdS black hole, where it was demonstrated that in a wide region of parameter space, there are infinitely many of extremal surfaces anchored on the same boundary region, unlike

[6] See also [49, 101, 102, 103, 104, 105, 106] and references therein for other explorations of holographic entanglement entropy as a probe in different contexts. For a more extensive review of the earlier work, see e.g. [107] and references therein.

the planar case where there is just one. Correspondingly, working with field theory on the spatially compact Einstein Static Universe describing the boundary of global AdS allows us to explore interesting finite-volume effects, which would have been absent in the non-compact case.

Having motivated the spacetime of interest, specifically global Vaidya-AdS class of spacetimes (with shell thickness, final black hole size, and dimension of the spacetime left as free parameters that we can dial), let us now specify our probes. From previous studies such as [61], it is clear that extremal surfaces of different dimensionality can behave qualitatively differently from each other; nonetheless there is a certain 'monotonicity' of the behaviour in terms of dimension. The greatest qualitative difference occurs between geodesics and higher-dimensional surfaces, and the greatest difference from the geodesic behaviour occurs for surfaces of highest possible dimension. This motivates us to focus on spacelike geodesics and codimension-two extremal surfaces, whose lengths and areas characterize certain types of correlators and entanglement entropy respectively in the dual CFT. Note that in 3-dimensional bulk, spacelike geodesics coincide with codimension-two extremal surfaces. Although in this special case many features trivialize (and as has already been well-appreciated, the BTZ black hole singularity behaves fundamentally differently from higher dimensional black hole singularities [108]), it will nevertheless be instructive to include Vaidya-AdS$_3$ in our explorations, in order to draw contrast with the higher-dimensional case.

The plan of the paper is as follows. In section 3.2 we describe the class of bulk geometries which we will use, namely the Vaidya-AdS spacetimes, and explain the coordinates for presenting our results graphically. We then turn to examining the CFT probes of this geometry, starting with spacelike geodesics in section 3.3, first focusing on the higher dimensional case in section 3.3.1 and then contrasting this with the Vaidya-BTZ case in section 3.3.2, where we can supplement our numerical results by closed-form expressions for the key quantities. In section 3.4 we turn to bulk codimension-two extremal surfaces in higher-dimensional Vaidya-AdS, and we conclude with a discussion in section 3.5. The more involved technical details are relegated to the appendices so as to avoid breaking the flow of the presentation. In the remainder of this section we give a preview of the main results.

In section 3.3 we consider geodesics with both endpoints anchored on the boundary, which we dub 'boundary-anchored' geodesics. We observe that every point in the bulk spacetime (in the higher dimensional cases) lies along some boundary-anchored spacelike geodesic. However, the closer this point lies to the singularity, the more nearly-null will such a geodesic be, which in turn means that its endpoints will in general be temporally separated. This motivates us to restrict attention to spacelike geodesics, whose endpoints lie at equal time on the boundary (dubbed 'equal-time-endpoint boundary-anchored' or ETEBA for short), which are the ones relevant for encoding equal-time correlators in the field theory. One way to achieve this is for the geodesic to be symmetric under swapping the endpoints, though we find that this is not the only option.

When both endpoints are located prior to the shell, the entire geodesic remains in the AdS part of the spacetime. This means that such ETEBA geodesics are constant-time geodesics in the pure AdS part of the spacetime, which cannot penetrate near the singularity. For a short time soon after the quench, we find that there exists a class of geodesics which are not symmetric under reversing their affine parameter, and yet still have endpoints at equal times. This class is not only more novel than the symmetric geodesics (since it does not appear in static spacetimes), but also important, as in a certain regime of the parameter space such asymmetric geodesics can reach closer to the singularity, and furthermore are shorter than the symmetric ones. Indeed, if the endpoints are taken to be antipodal and occur arbitrarily soon after the shell, the corresponding shortest geodesic is nearly null, crossing the shell near its implosion at the origin, and has arbitrarily small length. While the asymmetric geodesics exist only up until some finite endpoint time, there are symmetric ETEBA geodesics reaching the boundary at arbitrarily late times, yet sampling the interior of the horizon. These necessarily cross the shell to circumvent the arguments of [61].

We then restrict the search further, to consider only the shortest geodesics for given endpoints. The intricate nature of the results is illustrated in fig. 3.7, which plots the regularised proper length ℓ along all families of geodesics which join antipodal points at time t, as a function of t. Curiously, the minimum ℓ for this set of curves jumps discontinuously, not once, but in fact four times (twice down and twice up), before the thermal value is achieved.

Having mapped the space of initial conditions for the geodesics to the space of corresponding boundary parameters (namely the length and position of the endpoints), we turn to identifying what part of the spacetime is actually probed by shortest ETEBA geodesics. We find that even this most restrictive class allows access to a spacetime region inside the horizon and simultaneously to the future of the shell, as indicated in fig. 3.8. However, this region is limited to relatively short time after the shell, and late-time near-singularity regions remain inaccessible.

These results are in contrast to the analogous results for the 3-dimensional (Vaidya-BTZ) spacetime. In this case, geodesics exhibit qualitatively different behaviour for small black hole ($r_+ < 1$) as opposed to large black hole ($r_+ > 1$) spacetimes. We first specialise to radial ETEBA (which in this case implies symmetric) geodesics. These only probe a part of the spacetime (except for the special case of $r_+ = 1$ when the entire spacetime is accessible), though the character of the unprobed region changes depending on whether the black hole is small or large. This behaviour is illustrated on spacetime plots in fig. 3.10 and on the corresponding Penrose diagrams in fig. 3.11. Adding angular momentum however has a dramatic effect: in the case of small black holes, the entire spacetime becomes accessible, even by ETEBA geodesics. On the other hand, for large black holes, the deepest boundary-anchored symmetric radial geodesic in fact bounds the region accessible to *any* boundary-anchored geodesic. In other words, for large black holes, a certain region of spacetime still remains inaccessible. However this region has very different – and almost

complementary – character from its higher-dimensional counterpart. Here it is confined to the vicinity of the shell inside the horizon, while the late-time near-singularity regions are fully accessible. (However, probing late-time near-singularity regions is not as useful as it would be in the higher-dimensional case, since the spacetime is locally AdS, so we cannot use such geodesics to directly probe the interesting strong-curvature effects; cf. [109].)

The behaviour of the length along shortest ETEBA geodesics, as function of time, is likewise very different for the BTZ case, as illustrated in fig. 3.13. For any-sized black hole, the length increases monotonically from the AdS value to the BTZ value, without exhibiting any remarkable features. This is consistent with the expectations for the behaviour of CFT correlator during thermalization, which we expect to be directly extractible from the shortest lengths geodesics for this 3-dimensional case (as argued in a similar context in [65]); we revisit this point in section 3.5.

Indeed, the consideration of boundary-anchored geodesics in 3-dimensional spacetime can be thought of as a special case of codimension-two extremal surfaces anchored on the boundary. Hence the shortest boundary-anchored geodesics have a bearing both on certain CFT correlators, as well as on entanglement entropy corresponding to a certain region (bounded by the geodesic endpoints). The results are compared quantitatively with the results of [64, 100] and [72], and agree with these in the regimes of early quadratic growth and intermediate linear growth of entanglement entropy.

In section 3.4 we turn to considering codimension-two extremal surfaces in the higher-dimensional case of Vaidya-AdS$_{d+1}$. While these share certain features in common with the 3-dimensional case, there are also important differences, as already exemplified by the fact that even in the static black hole geometry, higher dimensional surfaces have richer structure [1]. We demonstrate, both analytically and numerically, that for arbitrarily late boundary time, one can construct extremal surfaces anchored at that time which penetrate the black hole, an example of which is presented in fig. 3.16. These surfaces lie along a specific maximal area surface inside the horizon, as observed recently in a related context by [63, 64]. By studying the linearized perturbations of the surfaces away from this point, we obtain a good handle on what region of the bulk can be probed, indicated in fig. 3.19. We find that while we can probe to a finite depth inside the event horizon for arbitrarily late times, the near-singularity region of the geometry remains inaccessible. In this respect, the codimension-two extremal surfaces appear to be less suitable probes of the singularity than geodesics. This of course persists when we restrict to the surfaces of smallest area, which reach only a very limited region inside the horizon.

We also consider the evolution of the area of these surfaces for a fixed boundary region, which is directly related to the thermalization of the entanglement entropy, analogously to the recent examination by [64, 100] in the planar context. Since we expect the global geometry to offer richer structure than its planar limit, we focus on nearly-hemispherical boundary regions. This is presented in fig. 3.20 and (perhaps disappointingly) offers no new surprises: the entanglement entropy increases smoothly, and monotonically interpolates between the vacuum and thermal value. The growth is linear at intermediate times,

controlled by the surface hugging the maximal area constant-r surface, in agreement with [64, 100], and for a sufficiently thin shell also exhibits the early-time quadratic growth derived therein.

3.2 The Vaidya-AdS spacetime

To model a simple holographic thermalization process, we consider a bulk geometry given by a global Vaidya-AdS$_{d+1}$ spacetime, mostly for $d = 2, 4$. This is a solution to Einstein's equations with negative cosmological constant and a stress tensor for a spherically symmetric null gas, obtained by expressing the Schwarzschild-AdS metric in ingoing coordinates, and then allowing the mass to depend on the ingoing time v. The metric can be written as

$$ds^2 = -f(r, v)\, dv^2 + 2\, dv\, dr + r^2\, d\Omega_{d-1}^2 \ , \tag{3.1}$$

where $d\Omega_{d-1}^2$ is the round metric on the unit S^{d-1},

$$f(r, v) = r^2 + 1 - \vartheta(v) \left(\frac{r_+}{r}\right)^{d-2} (r_+^2 + 1) \ , \tag{3.2}$$

and $\vartheta(v)$ is monotonic function, increasing from 0 in the past to 1 in the future, characterising the profile of a spherical null shell collapsing from the boundary. We take the shell to be concentrated around $v = 0$, with a thickness of order δ, taking $\vartheta'(v)$ as a function with compact support $[0, \delta]$. We also consider the limit of a thin shell, for which $\delta \to 0$.

The metric interpolates between pure AdS inside (or to the past of) the shell and Schwarzschild-AdS outside (or to the future of) the shell. Concretely, away from $v = 0$, the metric inside and outside can be expressed separately in static coordinates,

$$ds_\alpha^2 = -f_\alpha(r)\, dt_\alpha^2 + \frac{dr^2}{f_\alpha(r)} + r^2 d\Omega_{d-1}^2 \ , \tag{3.3}$$

where the subscript α stands for i inside the shell and o outside. The event horizon is at $r = r_+$ at late times, and it originates from $r = 0$ at some $v = v_h < 0$. The origin of spherical coordinates $r = 0$ is smooth for $v < 0$, but forms a curvature singularity for $v > 0$.

Coordinates for spacetime plots: It is convenient to compactify the radial coordinate such that AdS boundary is drawn at finite distance, using $\rho \in (0, \pi/2)$ defined by

$$\rho = \tan^{-1} r. \tag{3.4}$$

A natural temporal coordinate is one which makes ingoing radial null curves always at 45 degrees. In terms of v and the compact radial coordinate ρ, the new temporal coordinate is

$$t = v - \rho + \frac{\pi}{2}, \tag{3.5}$$

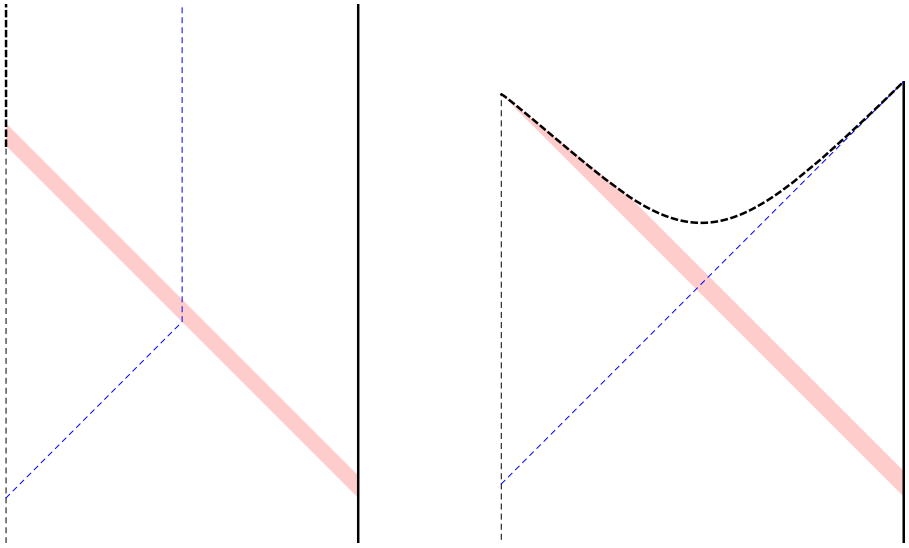


Figure 3.1: Eddington-Finkelstein (left) and Carter-Penrose (right) diagrams for the AdS-Vaidya spacetime, with $d = 4$ and $r_+ = 1$. The vertical black dashed line on the left side in each panel is the origin of spherical coordinates before the shell begins, and the thick dashed curve the singularity. The AdS boundary is the solid thick line on the right. The infalling shell of matter is indicated by the red shading (its width indicating the shell thickness δ used in our numerical calculations), and the blue dashed line denotes the event horizon.

the last term ensuring that t coincides with v on the AdS boundary. In pure AdS, $t = t_i$ is in fact the usual static coordinate, and both ingoing and outgoing radial null curves are at 45 degrees. However, in the black hole geometry, t is different from the static coordinate, $t \neq t_o$: indeed, t is a good coordinate on the whole spacetime, whereas the static t_o blows up at the horizon.

In plotting spacetime diagrams, with all relevant directions visible, we will use coordinates $(\rho \cos\psi\,,\ \rho \sin\psi\,,\ t)$, where we use ψ as a shorthand for an angular variable on the S^{d-1}. In particular, ψ will be either a longitude φ, or the colatitude θ, as appropriate to the symmetries of the problem in question. It will often be more convenient to consider 2-d projections by suppressing t or ψ. In particular, we will use ingoing Eddington-Finkelstein plots $(\rho\,,\ t)$ and 'Poincaré disk' plots $(\rho \sin\psi\,,\ \rho \cos\psi)$.

The Eddington diagrams are useful because the static nature of the pre- and post-collapse geometries is made manifest. It should be appreciated that they can be quite misleading in that they distort the geometry, particularly near to the shell, and do not represent the causal structure.

For these reasons, we complement them by showing plots on Carter-Penrose diagrams, in which radial null geodesics, both ingoing and outgoing, appear as 45 degree lines. To construct the necessary lightcone coordinates (U, V) for this purpose, we observe that every radial null geodesic has a past endpoint on the boundary. The lightcone coordinates are obtained by assigning a number to each outgoing and each ingoing radial null geodesic,

which can be done in a very general procedure.

Given a spacetime point p, we define U by constructing the radial outgoing null geodesic through p into the past, and noting the value of v when it meets the origin $r = 0$. The coordinate V can then be chosen to be constant along ingoing null geodesics, and such that the boundary lies at $V = U + \pi$. Operationally, this can be done by finding where an outgoing null radial geodesic ending on the boundary passes through the origin, which gives $V - \pi$ for that boundary point.

This construction has the advantage of ensuring that both the AdS boundary, and the smooth origin at $r = 0$ before the shell collapse begins, are straight lines on which $U - V$ is constant. The resulting coordinates in the pre-collapse part of the spacetime are identical to the standard (radially compact) coordinates on pure AdS, and hence noncompact in the past. The diagram is however compact in the future, terminating at the singularity.

For the lowest-dimensional case of Vaidya-BTZ, with $d = 2$, in the limit of a thin shell, this coordinate change can be computed explicitly, and the result is reproduced in Appendix 3.A. The metric in this case takes a particularly simple form:

$$ds^2 = \frac{-dT^2 + dR^2}{\cos^2 R} + r(T,R)^2 d\phi^2, \tag{3.6}$$

where $T = \frac{V+U}{2}$ and $R = \frac{V-U}{2}$, and the radius of the ϕ circle is given by

$$r(T,R) = \begin{cases} \frac{(1-r_+^2)\sin R - (1+r_+^2)\sin T}{2\cos R} & \text{if } R+T > 0 \\ \tan R & \text{if } R+T < 0 \end{cases} \tag{3.7}$$

with the coordinate ranges bounded by origin at $R = 0$, the boundary at $R = \frac{\pi}{2}$, and the singularity at $(1 - r_+^2)\sin R = (1 + r_+^2)\sin T$.

While the causal structure is made manifest in these diagrams, they hide the fact that the post-collapse geometry is static, and can be distorting because late times are compressed into a corner of the diagram.

3.3 Geodesics

We begin by studying geodesics, as the simplest example of extremal surfaces. We will restrict our considerations to spacelike geodesics, since these can end at the boundary but need not remain outside the event horizon. In contrast, null curves are causally prevented from entering the horizon and reemerging out to the boundary, while timelike geodesics cannot even reach the boundary.

By spherical symmetry, we can restrict to geodesics lying in the equatorial plane, reducing the spherical directions to a single relevant longitude φ. This simplifies the problem to a 3-dimensional one, described by Lagrangian

$$\mathcal{L} = -f\,\dot{v}^2 + 2\,\dot{r}\,\dot{v} + r^2\,\dot{\varphi}^2, \tag{3.8}$$

where the dots denote differentiation with respect to an affine parameter s, chosen so that $\mathcal{L}$ is constant at $+1$ on the curve. This parameter s is then an arclength.

The spherical symmetry also supplies us with a first integral for the angle, given by the conservation of angular momentum

$$L = r^2\,\dot{\varphi}. \tag{3.9}$$

Away from the shell, the spacetime is locally static, so there is in addition a conserved energy

$$E = f\,\dot{v} - \dot{r}, \tag{3.10}$$

though it is important that this is constant only locally in regions where f is independent of v, and changes whenever the shell is encountered.

The v equation of motion can be written to express this change, in the form

$$\dot{E} = \frac{1}{2}\,f_{,v}\,\dot{v}^2 \leq 0 \tag{3.11}$$

where the inequality uses the fact that the profile function ϑ is nondecreasing. As well as telling us the sign of the jump in the energy, it also confirms the natural expectation that it should be greater when the shell is more dense, at smaller r, when $d > 2$. In the limit of a thin shell, the discontinuity can be calculated exactly, as

$$E|_{v=0^+} - E|_{v=0^-} = \frac{1}{2}\,(f|_{v=0^+} - f|_{v=0^-})\,\dot{v}|_{v=0} \tag{3.12}$$

and is equivalent to the condition that $\dot{v}$ is continuous.

For numerics, we use second order equations of motion for v and r, given by

$$\begin{aligned}\ddot{v} &= -\frac{1}{2}\,f_{,r}\,\dot{v}^2 + \frac{L^2}{r^3}\\ \ddot{r} &= \frac{1}{2}(f_{,v} - f\,f_{,r})\,\dot{v}^2 + f_{,r}\,\dot{r}\,\dot{v} + f\,\frac{L^2}{r^3},\end{aligned} \tag{3.13}$$

integrating the definition of the angular momentum eq. (3.9) to solve for φ.

One useful fact that can be seen immediately from the equations of motion is that whenever $\dot{v}$ vanishes, $\ddot{v}$ must be positive. This, along with the fact that v must be increasing as the boundary is approached, implies that v has exactly one local (and hence also global) minimum along the geodesic. The uniqueness makes this a convenient point from which to start numerical integration.

Effective potential: Much of the qualitative behaviour of the geodesics in the static parts of the geometry can be understood from expressing the radial motion in the form of an effective potential, by eliminating $\dot{v}$ and $\dot{\varphi}$ in favour of the conserved quantities L and E:

$$\dot{r}^2 = E^2 - V_{\text{eff}}(r), \quad \text{where} \quad V_{\text{eff}}(r) = \left(\frac{L^2}{r^2} - 1\right) f_\alpha(r). \tag{3.14}$$

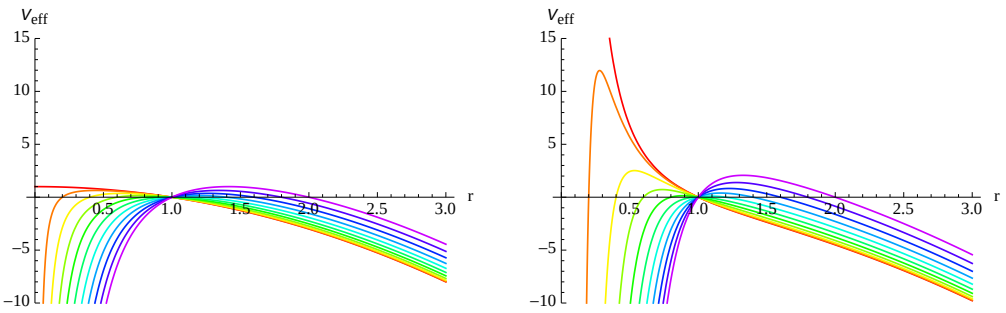


Figure 3.2: Effective potentials for spacelike geodesics in the BTZ (left) and Schwarzschild-AdS_5 (right) geometry with horizon radius $r_+ = 1$, for various values of the angular momenta: $L = 0$ (red) to $L = 2$ (purple), in increments of 0.2. The two cases are qualitatively different for low-L values.

Again the subscript α refers to i or o to distinguish the pre- and post-collapse static parts of the geometry. Most important is the form of this potential in the static Schwarzschild-AdS spacetime. Excepting the special case of radial geodesics ($L = 0$), it is unbounded from below as r tends to both zero and infinity, has exactly two zeroes at $r = r_+, L$, and has a single maximum between them.

As is well-known (see e.g. the discussion in [108]), the 2+1 dimensional case is qualitatively different from the higher dimensional cases. This is because in 3 dimensions, the BTZ black hole has locally the same geometry as pure AdS, so geodesics are not cognisant of the curvature singularity at $r = 0$. We can see this explicitly from the form of the effective potential, plotted in fig. 3.2. For $d > 2$, the height of the maximum grows without bound as L is taken to zero, with the potential for radial geodesics unbounded from above for small r. (For example, in $d = 4$, the maximum of V_{eff} scales as $\frac{r_+^2(r_+^2+1)}{2L^2}$ at small L.) The consequence is that the singularity repels even nearly-null geodesics, if they are sufficiently close to being radial. This simple observation turns out to be crucial to our considerations. The $d = 2$ (BTZ) case is qualitatively different from this, with the potential reaching a maximum of $(r_+ - L)^2$, which is bounded for small L. This means that geodesics must have small energies, or end up in the singularity, so the nearly null geodesics of relevance in higher dimension will not be relevant for the $d = 2$ case.

Lengths of geodesics: One natural observable associated with spacelike geodesics is their proper length. Since we are using arclength as a parameter, in principle we merely need to read off the difference Δs between the initial and final value on the curve. This is complicated by the fact that the length is infinite: r asymptotes to $e^{\pm s}$ as $s \to \pm\infty$. We will only need to compare lengths of geodesics with matching endpoints, so we need not worry about the details of choosing a renormalization scheme. We regulate in a simple way, cutting off at a large radius r_c, and subtract off the divergent piece $2\log(2r_c)$. Whenever length ℓ is referred to, it may be taken to mean this regularized version.

Terminology: As mentioned above, for purposes of relating the geodesic (length) to a natural CFT observable (i.e. a two-point function of high-dimension operators with insertion points at the geodesic endpoints on the boundary), we wish to restrict attention to spacelike geodesic with both endpoints on the (same) boundary. We will refer to these as *boundary-anchored* geodesics[7] and we will be primarily interested in the question of what part of the bulk spacetime is reached by the set of boundary-anchored geodesics. In particular, how deep into the black hole, and how close to the curvature singularity, can such boundary-anchored geodesics penetrate. Since one is often most interested in equal-time correlators on the CFT side, we will find it convenient to further refine our class of boundary-anchored geodesics to ones with both endpoints lying at the same boundary time; we will call these ETEBA (for 'equal-time-endpoint boundary-anchored') geodesics.

Initial condition space: A preliminary task is to find a convenient parameterization of the set of all geodesics. For this, we use the fact that on any given geodesic, v has exactly one stationary point (in contrast to r, for example, which may have more). With this in mind, we parameterize the set of geodesics by three parameters (v_0, r_0, E_0), respectively corresponding to the initial values of v and r when $\dot{v} = 0$, and the energy $E = f\dot{v} - \dot{r}$ at that point. We take E_0 to be nonnegative, since choice of sign corresponds only to choosing the direction of parameterization. The angular momentum follows from these; in fact $L = r_0$ (another choice of sign here corresponds to choosing the direction of increase of φ). These parameters are sufficient to give an initial unit tangent vector, from which the geodesic may be found, and no two different sets of parameters will give rise to the same geodesic. Of course, some of these will end up in the singularity, so can be disregarded. We have thus put the set of all geodesics (modulo symmetries) into one-to-one correspondence with the set of $(v_0, r_0, E_0) \in \mathbb{R} \times [0, \infty) \times [0, \infty)$, which we will henceforth refer to as 'initial condition space'.

3.3.1 Geodesics in higher dimensions

As already noted, the lowest dimensional case of BTZ is qualitatively different from higher dimensions, and the questions we are considering have correspondingly different answers. This section will focus on the case of Vaidya-Schwarzschild-AdS_{d+1} with $d \geq 3$, postponing the discussion of Vaidya-BTZ to section 3.3.2.

One natural question to ask is what spacetime region is accessible to spacelike geodesics with both endpoints anchored on the AdS boundary. Our first observation is that the answer to this question is in fact very simple: Every point in the spacetime has a boundary-anchored spacelike geodesic passing through it. For example, given any point (r_0, v_0) inside the horizon and after the shell, one may take a radial geodesic, picking the energy such that $E^2 = V_{\text{eff}}(r_0)$, so that r is at a local minimum. Constructing the geodesic in the maximally extended static black hole spacetime, it would join opposite asymptotic regions,

[7] In [61] these were referred to as 'probe geodesics'.

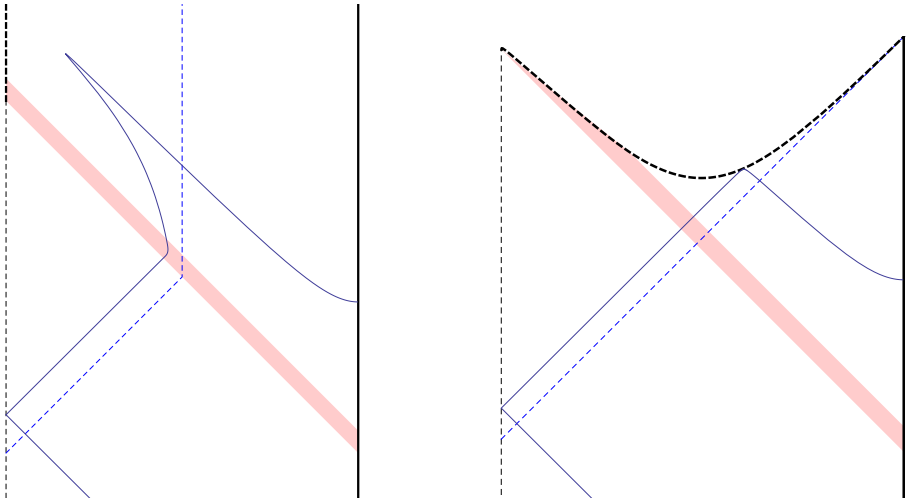


Figure 3.3: A radial geodesic (solid blue curve) with $v_0 = -1.4$ and $E_0 = 12$, in Schwarzschild-AdS_5 with $r_+ = 1$, plotted on Eddington (left) and Penrose (right) diagrams, as described in fig. 3.1. We have cut off the uninteresting bottom part of the geodesic; its continuation approaches the boundary in a similar manner to the top part. Note that on the Penrose diagram in the right panel, the geodesic looks like it reaches the singularity, but this is misleading effect of the coordinates, as evident in the Eddington diagram on the left panel.

geometrically encoding correlations between the two halves of the thermofield double state.[8] In the Vaidya-AdS spacetime, one end of this is altered since the second asymptotic region is replaced inside the shell by part of pure global AdS. Both ends must then lie on the (single) boundary, since there is no other boundary and no way it can reach the singularity since the effective potential is unbounded there. An example of such a geodesic is shown in fig. 3.3.

There are two noteworthy points. Firstly, geodesics reaching very close to the singularity must be nearly null, and will hence have arbitrarily short lengths. Indeed, this was the key observation used in the eternal Schwarzschild-AdS context in [108] to probe the black hole singularity (upon suitable analytic continuation). Secondly, geodesics reaching inside the horizon at late times will have one end close to $v = v_h$, the value of v at which the horizon is first formed. For strictly null geodesics, this observation was used in the context of bulk-cone singularities [88] to detect the horizon formation event: radial null geodesics whose earlier endpoint approaches v_h have the other endpoint $v \to \infty$. However, bulk-cone singularities arise from individual null geodesics which cannot penetrate the black hole by the usual causality constraints, so they are more limited probes of the bulk geometry [61].

With these observations made, we now restrict attention to the case of ETEBA geodesics, whose endpoints lie at matching times, corresponding to equal-time CFT correlators.

[8] In that context, such a geodesic would not qualify as boundary-anchored geodesic since it connects different boundaries; indeed, as argued in [61] for any static spacetime, boundary-anchored geodesics can only probe the spacetime region outside the black hole.

ETEBA geodesics: One obvious way to ensure a geodesic will have endpoints lying at equal times is to impose a $\mathbb{Z}_2$ symmetry under reflection, i.e. under swapping the endpoints. This is equivalent to setting $E_0 = 0$, so the initial conditions at the earliest part of the geodesic have this enhanced symmetry. In a globally static geometry, energy conservation implies that this is the only option, but it no longer needs to be the case in spacetimes with nontrivial time evolution. Indeed, the Vaidya geometry admits geodesics with equal-time endpoints which do *not* respect this symmetry.

A further refinement, relevant in cases with multiple geodesics joining the same endpoints, is to restrict to geodesics of shortest length for given time and angular separation of the endpoints, which are expected[9] to dominate the CFT correlators.

The classification of these classes of geodesics amounts to the following procedure:

1. Characterize the set of geodesics with both endpoints on the boundary.

2. Identify those geodesics with endpoints at equal times.

3. Compare lengths of such geodesics with matching endpoints.

Having identified the initial condition space (v_0, r_0, E_0), we must first find the region of this space for which both ends of the associated geodesic reach the boundary, and then find the two-dimensional surface in initial condition space for which the endpoints are at equal times. This is the level set $\Delta t = 0$, where Δt is the difference of times at final and initial endpoints (being the limits of v as $s \to \pm\infty$). This is a 2-parameter set of geodesics. One part of this surface will be the portion of the plane $E_0 = 0$ for which the geodesic reaches the boundary. Then, we find the time t_∞ and angular separation $\Delta\varphi$ of the endpoints for each such geodesic, along with the length ℓ. This amounts to finding the map from initial condition space to 'boundary parameter space' $(t_\infty, \Delta\varphi, \ell)$, which collects all the field theory data associated with a given geodesic. The image of the equal-time geodesics under this map is a two-dimensional surface in boundary parameter space, and comparing lengths for given endpoints will amount to understanding different branches of this surface.

Initial condition surface of ETEBA geodesics: We numerically undertook a systematic study of the geodesics in the Vaidya-Schwarzschild-AdS_5 spacetime, to find a representative sample of ETEBA geodesics. This was done by taking a fine grid of initial points (v_0, r_0), and for each of these points finding every initial energy which gives an appropriate geodesic, in the following process:

1. Identify the range of energies for which geodesics reach the boundary at both ends. This turns out to be an interval (possibly empty), which can be understood from the effective potential: the geodesic hits the singularity when the energy exceeds

[9] This expectation is subject to the assumption that this dominant saddle point lies on the path of steepest descent. For nearly-null geodesics bouncing off the singularity in the eternal Schwarzschild-AdS spacetime this does not happen as discussed in [108], so accessing the signature of this geodesic directly from the field theory is more subtle. We revisit this point in section 3.5.

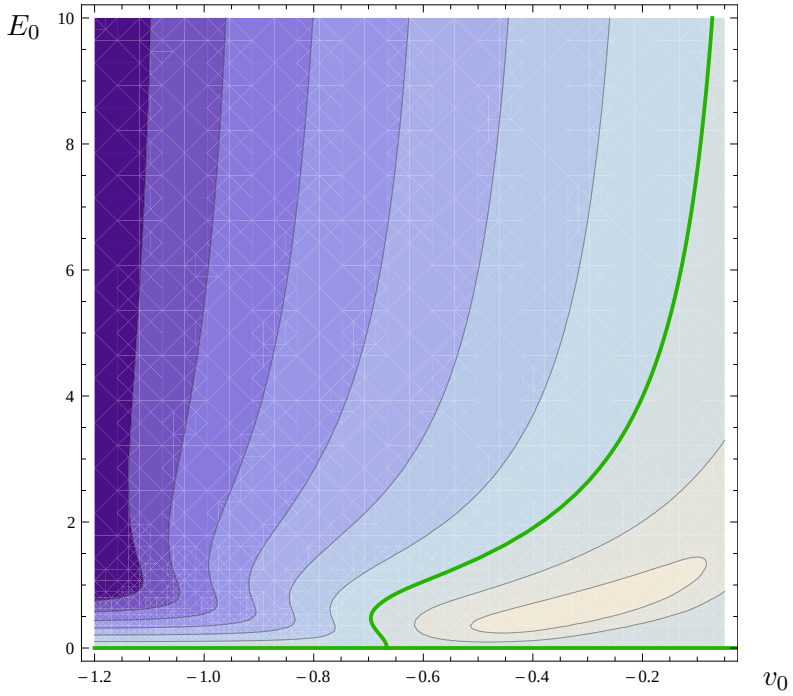


Figure 3.4: Contours of Δt for radial geodesics. They are parameterised by the value of $v = v_0$ and the energy E_0 when they pass through the origin. The green lines give the $\Delta t = 0$ contours, corresponding to ETEBA geodesics.

the maximum of $Veff$. This maximum energy is found by progressively bisecting between energies reaching the boundary or hitting the singularity.

2. Take a sample of geodesics reaching the boundary, and identify when the endpoints swap temporal order between adjacent energies. Each such occasion identifies an interval of energies containing a root of Δt.

3. Use a root-finding algorithm to find the appropriate initial energy within each such interval.

Each energy E_0 found in this manner gives a point (v_0, r_0, E_0) in the equal-time surface $\Delta t = 0$ of the initial condition space. Sufficiently many such points build up a complete picture of this surface.

The first piece of the picture can be obtained from looking at radial geodesics, for which $L = r_0 = 0$. Provided the initial point is regular (meaning $v_0 < 0$ here), these always end at the boundary, since the Schwarzschild-AdS effective potential is unbounded as $r \to 0$ in this case (cf. the red curve in right panel in fig. 3.2). The restriction to radial geodesics leaves us with two parameters to specify, namely (v_0, E_0), and the equal-time radial geodesics give a curve in this space. This turns out to have two branches, as shown in fig. 3.4, one the symmetric $E_0 = 0$ case, and another at nonzero initial energy.

The reason for the latter is a trade-off between two competing effects. At nonzero energy, as it goes away from the origin the geodesic moves into the future or past depending

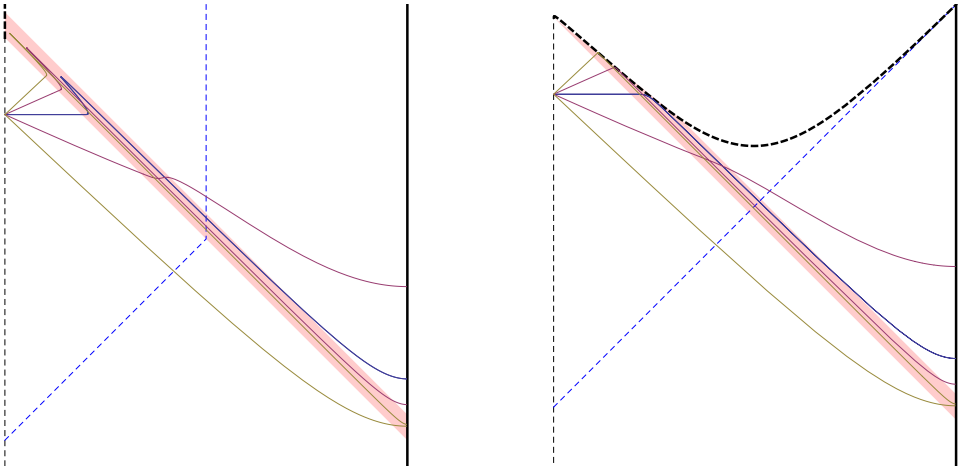


Figure 3.5: Radial geodesics passing through the origin at $v_0 = -0.3$, with increasing energy, plotted on Eddington diagram (left) and Penrose diagram (right), with $d = 4$ and $r_+ = 1$. The blue curve has zero initial energy, so is symmetric, the purple has initial energy $E_0 = 0.5$, and the yellow has $E_0 = 2.7$, close to the energy required to give equal-time endpoints.

on which direction is taken, and this separates the two branches in time. In a globally static geometry, the conservation of energy means that this separation persists to the boundary. This argument fails in the evolving geometry, but as long as the time-dependence is not too strong, this effect should still dominate.

The second effect is that the future branch encounters the shell of matter later and closer to the origin, when it has collapsed more, is more dense, and causes stronger curvature. This strongly influences the geodesic, and there is a large jump in energy as the shell is crossed, as implied from equation eq. (3.11). The future branch of the geodesic becomes nearly null, and hugs very closely to the shell. If this effect is different enough for past and future branches, it can cause the endpoints to exchange order in time.

It turns out that for sufficiently late initial conditions, it is the latter effect which dominates at low energies, the former taking over when the geodesic is nearly null, as illustrated in fig. 3.5.

This intuition for the existence of asymmetric equal-time geodesics also gives an indication of when they are unlikely to exist. Firstly, as we will argue in section 3.3.2, they do not exist in a 3-dimensional bulk. The effect of the shell on the energy is independent of the time at which the geodesic crosses it, because of the slow fall-off of gravity, so the competition is absent. Related to this, even radial geodesics of sufficient energy will not be prevented from ending in the singularity. Secondly, moving back to higher dimensions, the competition relies on the high energy, nearly null geodesics, which will fail for appreciable angular momentum. The maximum of the effective potential must be high enough to reflect the geodesics away from the singularity, but this maximum is reduced as L is increased. The result is that asymmetric geodesics only exist joining points of the boundary sphere that are close to antipodal.

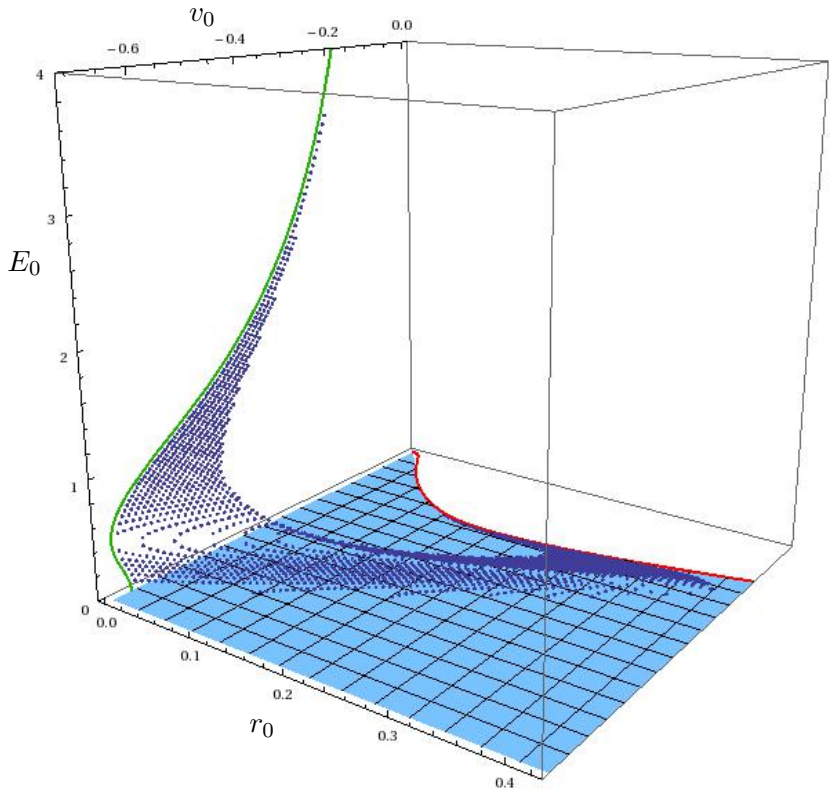


Figure 3.6: The surface in initial condition (v_0, r_0, E_0) space corresponding to ETEBA geodesics. The part of the plane $E_0 = 0$ for which geodesics are boundary-anchored, bounded by the red curve, gives symmetric geodesics. The blue points give asymmetric geodesics, and the curve for those which are radial is shown in green (c.f. fig. 3.4).

The full surface of initial conditions corresponding to ETEBA geodesics is shown in fig. 3.6.

Length of geodesics: The next stage is to map this surface (of initial conditions corresponding to ETEBA geodesics) into the boundary parameter space $(t_\infty, \Delta\varphi, \ell)$. This gives a complicated, multi-branched surface, but many of the salient features are revealed from taking a cross-section at $\Delta\varphi = \pi$, which corresponds to the set of geodesics joining antipodal points of the boundary sphere at equal times, shown in fig. 3.7. At early times, before the collapse begins, the only possibility is a simple straight line through the middle of AdS; this geodesic is both symmetric and radial. At late times, the only possibilities are again symmetric geodesics, lying at constant Schwarzschild-AdS time t_o, but these are not radial as they cannot penetrate the event horizon. This regime is then dominated by a geodesic simply deformed to one side of the horizon.[10] In the intermediate region, these families can be continued, and indeed meet, but there is also the additional possibility of the asymmetric geodesics presented in fig. 3.4 and the accompanying discussion. This additional family dominates for a short time immediately after the collapse; indeed for

[10] There are infinitely more possibilities, since the geodesic may wrap around the horizon arbitrarily many times, but such geodesics are of course longer.

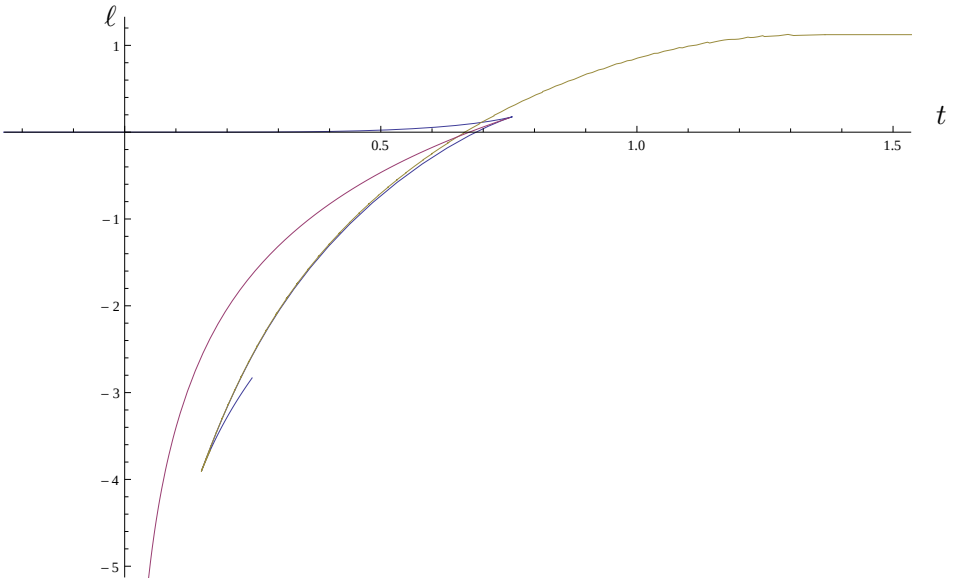


Figure 3.7: Regularised length of ETEBA geodesics joining antipodal points, plotted against the time at which the boundary is reached. The blue curve corresponds to radial, symmetric geodesics; the purple curve to symmetric but not radial, and the yellow curve to radial but not symmetric ones.

sufficiently early times the lengths may be arbitrarily short as the geodesics become very nearly null. On the other hand, there are no antipodal ETEBA geodesics which are neither symmetric nor radial.

The structure that fig. 3.7 reveals is surprisingly intricate. In the course of thermalization (i.e. between $t = 0$ when the shell starts imploding and $t \approx 1.3$ when the antipodal ETEBA geodesic remains entirely to the future of the shell), there are 4 'jumps' in the shortest length as different branches start or terminate. There are also several points where families of geodesics exchange dominance, but these kinks are hidden by the shorter ℓ families. The field theory interpretation of fig. 3.7 is, on the face of it, quite strange. It would seem to suggest that, during thermalization, the equal-time correlators of high-dimension operators of antipodal points correspondingly undergo no less than four discontinuous jumps. Furthermore, the first of these, at the start of thermalization, is an unbounded increase. However, as discussed in section 3.5, the shortest ℓ real-time geodesics may not actually be the ones to dominate the CFT correlator; such contingency arises in the simpler context of the eternal Schwarzschild-AdS geometry [108]. Nevertheless, even if the correlator is not dominated by these geodesics, their rich structure should still be subtly encoded in the correlation functions, possibly extractible by suitable analytic continuation.

We expect the geometry leading to this unexpected behaviour to be robust to changing many details of the collapse, depending rather only on the main features: spherical symmetry, and the formation of a spacelike singularity.[11] This is because any such

[11] The singularity however has to be 'black-hole-like' in the sense that it repels at least some class of spacelike geodesics; if the singularity were of the big crunch type (wherein all the transverse directions contract as the singularity is approached), then our spacelike geodesics would simply terminate in it.

geometry allows for nearly null radial geodesics, essentially following light rays except close to the singularity where they are repelled, with equal-time endpoints, by sending them into the corner of the Penrose diagram where the singularity is formed.

For geodesics joining points which are far from antipodal, the picture is simpler, with one family having the shortest length for all time, smoothly and monotonically interpolating between vacuum and thermal values. The asymmetric geodesics are absent entirely, this family disappearing very quickly on moving away from $\Delta\varphi = \pi$. The other parts of the curves visible in fig. 3.7 split into two families, one dominant, and the other corresponding to geodesics passing round the far side of the black hole. In the limit of large black hole and small angular separations, which recovers the planar black hole case, the picture becomes even simpler, since then even the possibility of passing on the other side of the black hole is not present. Hence the intricate structure observed in fig. 3.7 relies on both the black hole having compact horizon and on the geodesic connecting sufficiently far-separated points within the spherical boundary.

Region of geometry probed: The final task is to identify the region of spacetime covered by the ETEBA geodesics, both in totality, and also restricting to the shortest length for given endpoints. The latter region gives the part of the bulk on which the associated field theory observable is most sensitive. We find that the deepest probing geodesics are those connecting antipodal points $\Delta\varphi = \pi$, so restricting to these alone will not reduce the accessible region.

The region covered by the geodesics as a whole is illustrated in fig. 3.8, which shows the deepest points reached by asymmetric and symmetric geodesics. The symmetric geodesics are adequate to cover almost all of the accessible region. In particular they reach inside the horizon at arbitrarily late times, though only by a small distance, shrinking to zero as $v \to \infty$. They also cover the entirety of the spacetime inside the shell ($v < 0$), which includes points arbitrarily close to the singularity. From our numerics, it appeared that these geodesics did this in such a way as to remain at bounded curvature (considering, for example, the Kretschmann scalar $R_{abcd}R^{abcd}$, which goes like $r^{-8}\vartheta(v)^2$). This computation is rather sensitive to the fine details of the profile of the shell, so it is not clear how robust the conclusion is. Indeed, taking the limiting case of a shell of zero thickness, it is clear from considering symmetric radial geodesics passing through the origin immediately before collapse that unbounded curvature can be obtained.

This region close to the singularity is the only place where one may do better by including the asymmetric geodesics. These reach the region of small r to only slightly later times, but crucially appear to be able to get arbitrarily close to the singularity at some strictly positive v, where the curvature may become arbitrarily strong.

Including the restriction of considering only the shortest geodesics, we do not lose access to much of the region soon after formation of the black hole. In particular, the

This observation indicates that probing cosmological singularities (and correspondingly the resolution of a cosmological singularity in quantum gravity) would be expected to be drastically different from that of a black hole singularity.

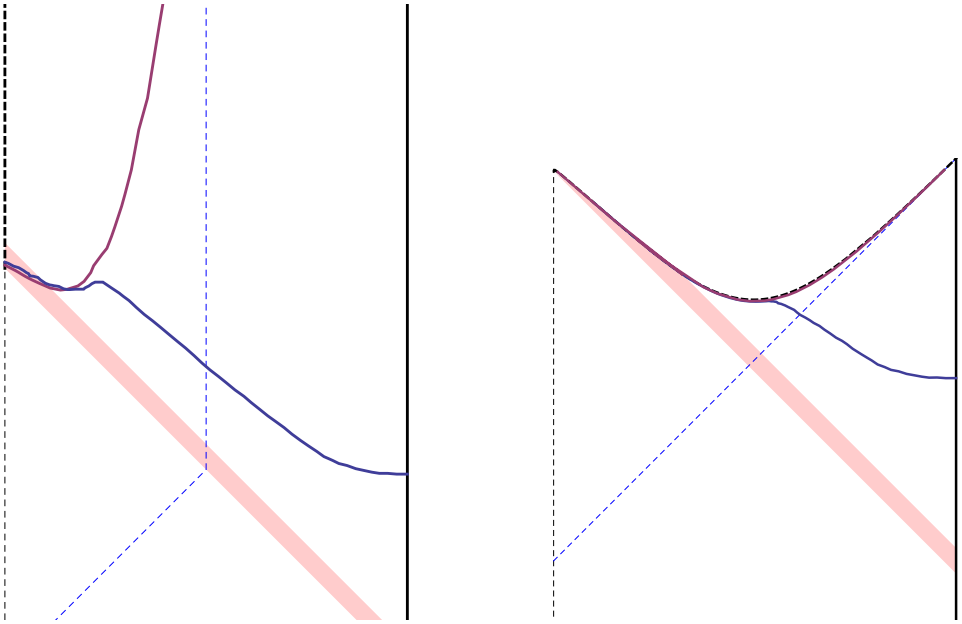


Figure 3.8: The region covered by all ETEBA geodesics, on Eddington and Penrose diagrams. The purple curves indicate the boundary of the region covered by only the symmetric geodesics, and the blue curves the region covered by the asymmetric geodesics. In particular, the asymmetric geodesics reach deeper, but only in a very small region.

same asymmetric geodesics that reach to regions of arbitrary curvature are also those of arbitrarily short length, and thus dominate.

Thereafter, we must consider what happens as dominance is exchanged between various families, as illustrated in fig. 3.7. The result is that we must exclude the geodesics reaching inside the horizon at late times, so the region after the shell and inside the horizon covered by shortest ETEBA geodesics is very limited, as shown in fig. 3.9. For example, in the case of $r_+ = 1$, $d = 4$, the latest time a shortest geodesic reaches the interior of the horizon is at $v \approx 0.4$. Thereafter, it should be emphasised that they reach not the whole exterior of the horizon, but only to the deepest radius of the shortest antipodal geodesic in Schwarzschild-AdS, which is at a finite though small distance above the horizon.

Apart from this region inside and close to the black hole, there is a distinct region which is not reached by shortest-length ETEBA geodesics. The shortest geodesics jump after $t = 0$ with the transition to the nearly-null geodesics, and because of this, a part of the pure AdS section of the geometry is also missed. This is the one place where including the geodesics which are not antipodal will allow access to a larger region. Despite this, there is still a small region remaining inaccessible, close to $r = 0$ and for some intermediate range of times, well after formation of the horizon but well before formation of the singularity.

3.3.2 Geodesics in Vaidya-BTZ

We have seen in section 3.3.1 that not every point in the Vaidya-AdS$_5$ spacetime is reached by the equal-time-endpoint boundary-anchored geodesics. In particular, events inside the

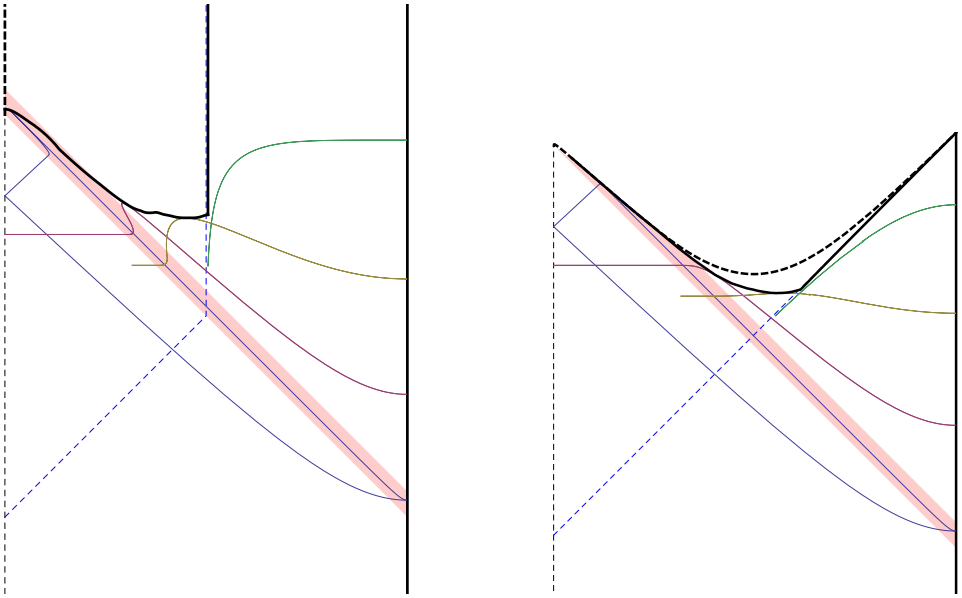


Figure 3.9: The region covered by shortest ETEBA geodesics, on Eddington and Penrose diagrams, bounded by the black curve, and examples of each family of such geodesics. Moving from early to late time, the blue curve is asymmetric and radial, the purple curve symmetric and radial, and the yellow and green symmetric but not radial. The green curve lies entirely in the Schwarzschild-AdS part, reaching not to the horizon but only to $r \approx 1.014$ in this case ($d = 4$, $r_+ = 1$)

black hole at late time (large v) do not lie on any ETEBA geodesic. However, at the same time, our geodesics probe arbitrarily close to the singularity just after its formation (though only traversing regions of bounded curvature). Here we wish to contrast this with the analogous set-up in $2 + 1$ bulk dimensions, i.e. the Vaidya-BTZ geometry. While as pointed out previously, this case is qualitatively different since the geometry is locally AdS_3 everywhere outside the shell and singularity, this case is most tractable by analytical means and most amenable to direct comparison to field theory. To take full advantage of the former, we also take the limit of a thin shell ($\delta \to 0$) in order to write simple closed-form expressions.

An additional curiosity in the case of BTZ is that for a time, the singularity is timelike. This can be seen from looking at outgoing radial geodesics: they may move away from $r = 0$ as long as $f(r = 0, v) > 0$, which happens for some window during the collapse. By making the collapse very slow, the singularity may even be made naked. Indeed, if the shell does not carry enough energy, having a BTZ black hole final state is not an option.[12] In the Vaidya case, while it starts out timelike, the singularity is of a particularly mild type, being only a spatial conical defect.

[12] See however [110] for a numerical study of scalar collapse in AdS_3 inducing turbulent instability which nevertheless remains regular.

Symmetric radial geodesics: Let us first consider the simplest case of symmetric radial geodesics in Vaidya-BTZ, starting at the origin $r_0 = 0$ before the implosion of the shell with $v_0 = -\tau$ where $\tau \in (0, \frac{\pi}{2})$, and with initial energy $E_0 = 0$. To simplify the computations, it turns out to be convenient to parameterize the final black hole size by a parameter μ defined by $r_+ = \sec\mu + \tan\mu$, where $\mu \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Note that $r_+ = 1$ corresponds to $\mu = 0$, which is a critical size separating qualitatively distinct types of behaviour.

The radial equation of motion outside the shell can be written as

$$\dot{r}^2 = r^2 + \frac{\sin^2\mu - \sin^2\tau}{(1-\sin\mu)^2} \ , \tag{3.15}$$

from which it is clear that when $\tau \geq |\mu|$, the geodesic can never reach the singularity at $r = 0$ since $\dot{r}^2$ would be negative for small r. When this fails ($\tau < |\mu|$), r has no turning points, so the fate of the geodesic depends on the sign of $\dot{r}$ just after crossing the shell: it will end in the singularity or on the boundary if it is negative or positive respectively. The calculations give

$$\dot{r}|_{v=0^+} = \frac{\sin^2\tau - \sin\mu}{\cos\tau\,(1-\sin\mu)} \ , \tag{3.16}$$

which for small black holes ($\mu < 0$) is automatically positive, so the geodesic continues to the boundary. On the other hand, large black holes $\mu > 0$ allow a regime for sufficiently small τ (i.e. later starting point, closer to the implosion of the shell) where $\dot{r} < 0$ outside the shell, so that the geodesic initially recedes to smaller r. If $\tau < \mu$, $\dot{r}$ remains negative for all r so the geodesic crashes into the singularity. On the other hand, if $\mu < \tau < \arcsin\sqrt{\sin\mu}$, it turns around at r_{tp}, where

$$r_{tp} \equiv \frac{\sqrt{\sin^2\tau - \sin^2\mu}}{1-\sin\mu} \ . \tag{3.17}$$

This can be made arbitrarily small by letting $\tau \to \mu^+$, so such boundary-anchored geodesic gets arbitrarily close to the singularity. Moreover, since $\dot{r}^2$ gets correspondingly small, the geodesic can remain in this vicinity for arbitrarily long span in v, and consequently make it out to the boundary arbitrarily late. In particular, the time at which it attains the boundary is given by

$$t = \frac{1-\sin\mu}{\cos\mu} \log\left[\frac{\cos\left(\frac{\tau+\mu}{2}\right)}{\sin\left(\frac{\tau-\mu}{2}\right)}\right] \ , \tag{3.18}$$

which is logarithmically divergent as $\tau \to \mu^+$.

From these considerations we can now determine what part of the spacetime is probed by these symmetric radial geodesics. The attainable region is bounded by the latest such geodesic, which originates inside the shell at $\tau \to 0^+$ for small black holes (i.e. when $\mu < 0$) and at $\tau \to \mu^+$ for large black holes (i.e. when $\mu > 0$). The limit $\mu \to 0$ agrees from both directions, and in this special case the entire spacetime is attainable. However, when $\mu \neq 0$, some spacetime regions are missing, the character of which depends on whether μ is positive or negative. This behaviour is illustrated in fig. 3.10 for small (left), intermediate (middle),

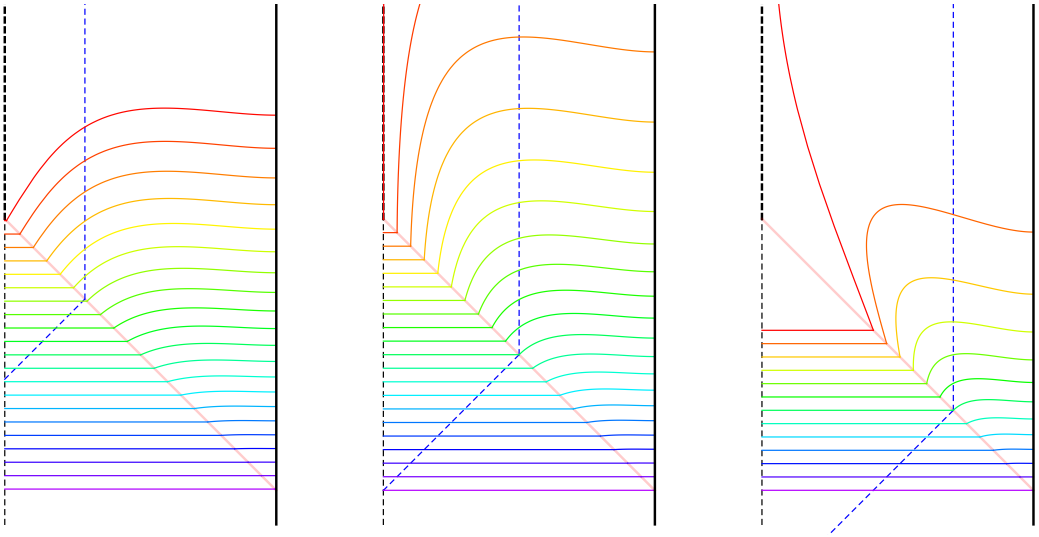


Figure 3.10: Radial symmetric ETEBA geodesics in Vaidya-BTZ, with horizon size $r_+ = 1/2$ (left), $r_+ = 1$ (middle), and $r_+ = 2$ (right) black holes. The red geodesic bounds the spacetime region which is attainable to this class of geodesics. We see that the unattainable region is above and to the left of this curve; for $r_+ = 1$ (i.e. $\mu = 0$) the entire spacetime is accessible.

and large (right) black holes on Eddington diagram, and in fig. 3.11 on the corresponding Penrose diagrams.

Small black holes ($\mu < 0$): The inaccessible region occurs to the future of the geodesic from $\tau \to 0^+$ (which is outgoing everywhere). This includes the entire interior of the black hole to the future of this geodesic. As $\mu \to -\pi/2$, this region is described by the line[13] $v > 2\tan^{-1} r - \frac{\pi}{2}$ on the Eddington plot. On the other hand, as $\mu \to 0^-$, the initial slope $\frac{dv}{dr}$ increases, and the time at which the boundary is attained eq. (3.18) diverges logarithmically. In this limit the unattainable region at large v gets pushed off to infinity.

Large black holes ($\mu > 0$): Now the inaccessible region occurs to the future of the geodesic from $\tau \to \mu^+$, which is initially ingoing, and turns around arbitrarily near the singularity, with arbitrarily small velocity. This means that the only unattainable region is the one between the shell and this geodesic. In the limit of very large black hole, $\mu \to \pi/2$, this region is described by the triangle bounded by $r = 0$, $v = 0$, and $v = \tan^{-1} r - \frac{\pi}{2}$, while as $\mu \to 0^+$ the region receded towards and gets elongated along the singularity $r = 0$.

These conclusions are made very clear by using the Penrose coordinates, which give

[13] This relation is simple in the tiny black hole limit since the spacetime region inside the horizon is so small that we can treat is as flat (recall that in BTZ the curvatures do not grow as the singularity is approached). On AdS scales the curvature is felt, though, and this bounding geodesic attains the boundary at $v = 2$. Note that, in contrast to the Eddington spacetime diagram, all geodesics are in fact straight horizontal lines in the Penrose diagram.

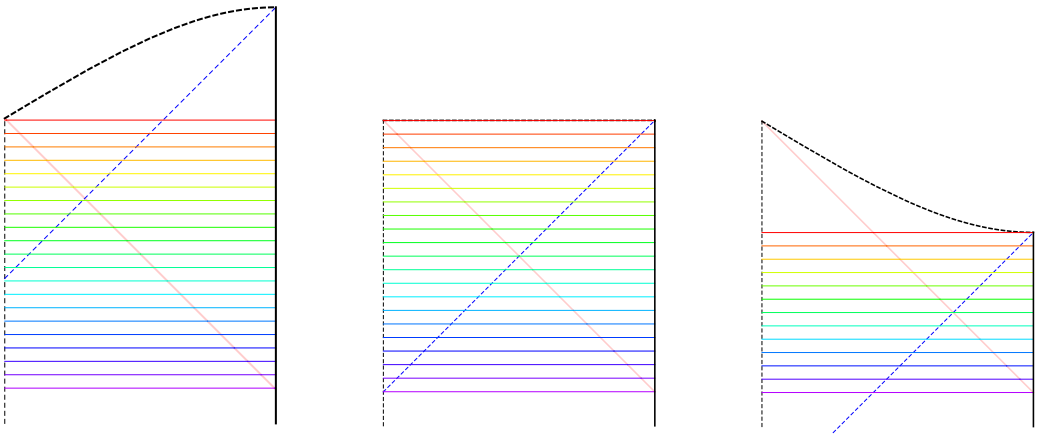


Figure 3.11: Radial symmetric ETEBA geodesics in Vaidya-BTZ as in fig. 3.10, now plotted on the Penrose diagram.

the metric of equation 3.7. In particular, it is manifest that the radial geodesics will follow identical curves to the case of pure AdS, and for the symmetric geodesics these are horizontal lines of constant T. The only remaining requirement is to know the shape of the singularity, given by $(1 - r_+^2) \sin R = (1 + r_+^2) \sin T$, which depends on the size of the black hole. For small black holes, this is at increasing T as R increases toward the boundary; for $r_+ = 1$ it is the horizontal line $T = 0$; and for large black holes it lies at decreasing T moving towards the boundary. Concretely, the singularity is between $R = T = 0$, and $R = \pi/2, T = 2 \tan^{-1} r_+ - \pi/2$. This alone is enough to reproduce the plots of figure fig. 3.11 and the associated conclusions.

The very restricted set of symmetric radial geodesics is a good starting point, but is too constraining. In particular, one might naturally expect that the region of spacetime covered will be increased by including more general classes of geodesic. As we demonstrate below, this expectation is only realised for small black holes.

For the small black holes, the result is analogous to the higher dimensions, in that boundary-anchored geodesics will cover the whole spacetime, though for a rather different reason – the mechanism can no longer rely on geodesics bouncing off the singularity. Indeed, we can use the same construction used in the previous section, of picking a radial geodesic passing through an arbitrary point at a local minimum of r, though it requires more work to argue that it will avoid the singularity. In fact, we can do better still in this case, since we can reach the same conclusion even with the restricted class of symmetric geodesics, once angular momentum is allowed. In particular, this means that ETEBA geodesics cover the whole spacetime.

This conclusion can be reached by considering a family of geodesics with initial conditions close to the singularity formation, with a small angular momentum. To fix notation, we will generalize the definition of the initial time τ to correspond to minus the initial AdS time, so that the shell is always reached at $r_s = \tan \tau$. This means that we must restrict $L < \tan \tau$ so that the geodesic actually starts inside the shell. We then consider the family

of geodesics with angular momentum $L = (-\sin\mu)\tau$. Taking τ to be small, there is a parametric separation between the radius where the geodesic crosses the shell r_s, the circular orbit radius r_0 at which the effective potential reaches its maximum, and the horizon r_+. Asymptotically as $\tau \to 0$:

$$r_s \sim \tau \ll r_0 \sim \sqrt{\frac{-\sin\mu\,\cos\mu}{1-\sin\mu}}\,\sqrt{\tau} \ll r_+ = \frac{\cos\mu}{1-\sin\mu}. \tag{3.19}$$

Moreover, the difference between E^2 and the maximum of the effective potential, which is the minimum of $\dot{r}^2$, is asymptotically

$$E^2 - V(r_0) \sim \frac{-2\,\sin\mu\,\cos\mu}{1-\sin\mu}\,\tau > 0. \tag{3.20}$$

This is positive but small, which means that the geodesic stays in the vicinity of $r_0 \sim \sqrt{\tau}$ for an arbitrarily long time Δv as $\tau \to 0$ before reaching the boundary. Since we can make r_0 arbitrarily small (and parametrically inside the horizon even for arbitrarily small black holes), and the radial velocity there likewise arbitrarily small, such geodesics penetrate arbitrarily close to the singularity at arbitrarily late time v. Details of the computation are included in section 3.A.

For the large black holes, the situation is entirely different, since the coverage of the radial symmetric geodesics is not improved by including even the most general boundary-anchored geodesics. The region covered by all geodesics is thus bounded by the innermost symmetric radial geodesic. This region includes points arbitrarily close to the singularity at late times, but is bounded away from its formation. This conclusion is easy to reach by using the Penrose coordinates once more. The equation of motion for geodesics associated with T in the BTZ part of the spacetime is

$$\ddot{T} + 2\dot{R}\dot{T}\tan R = \frac{1+r_+^2}{2}\frac{L^2}{r(T,R)^3}\cos R\cos T, \tag{3.21}$$

and the right hand side is positive for the corresponding range of coordinates. If $\dot{T} = 0$, $\ddot{T} \geq 0$, with equality only for the radial ($L = 0$) geodesics, so T can never have a local maximum on the geodesic. This means that if a geodesic lies above the critical curve $T = 2\tan^{-1} r_+ - \pi/2$ for any of its length, it must end in the singularity in at least one direction. The conclusion is that boundary-anchored geodesics see no more of the spacetime than the symmetric radial ones, namely the region $T \leq 2\tan^{-1} r_+ - \pi/2$.

Asymmetric ETEBA geodesics: In higher dimensions, we saw the novel feature of geodesics with endpoints at equal times, but nonetheless having no reflection symmetry. Our intuition for their existence relied on competition between two effects, one of which required nearly-null radial geodesics to be repelled from the singularity. In the case of BTZ, this effect is absent, since the effective potential is bounded, so it is a natural expectation that this class of geodesics does not exist.

If asymmetric ETEBA geodesics were to exist, it is expected that they would appear amongst radial geodesics, to give the largest potential barrier away from the singularity. With this simplification of assuming zero angular momentum, it is immediate from the metric in terms of Penrose coordinates that they may not exist. As already noted, in these coordinates the radial geodesics are identical to those in pure AdS (with the restriction that they must avoid the singularity), which move monotonically in T.

Allowing for angular momentum, this straightforward argument fails since T may have a minimum in the interior of the spacetime. The possibility that there may be asymmetric ETEBA geodesics is not in principle ruled out, but it seems highly unlikely that they would only appear for some intermediate L. This conclusion is supported by numerical calculations such as performed in higher dimensions, from which we find that for $d = 2$ there are indeed no asymmetric ETEBA geodesics.

Regions probed by ETEBA geodesics, and lengths: Our previous remarks have already answered the question of the region probed by ETEBA geodesics, being in the case of small black holes the entire spacetime, and in the case of large black holes the region outside the latest boundary-anchored radial symmetric geodesic. The final part of the picture is the refined question of the region covered by the shortest ETEBA geodesics.

The question of which geodesics dominate by virtue of having shortest length for given endpoints was investigated numerically, and turns out to have a simple answer, in contrast to the higher-dimensional cases. Because the only ETEBA geodesics are symmetric, we need only look at a two-parameter initial condition space, characterized by the location of the minimum of v.

We begin with the geodesics connecting antipodal points. There are two obvious candidates for such geodesics. Firstly, radial geodesics, with initial condition at $r = 0$, will automatically fall into this class. Secondly, in the static BTZ geometry there are antipodal geodesics passing outside the event horizon, with closest approach at $r = r_{\min}$, so in Vaidya-BTZ they must exist at late times, along with a continuation of the family to earlier times. This family in fact joins up continuously with the radial geodesics. Before this time, the only choice is the radial family, but after the nonradial family appears, there is a choice of two, of which the nonradial is always shorter. This means that the shortest antipodal geodesics follow a continuous curve in initial condition space as boundary time increases, starting at $r_0 = 0$, moving to nonzero r_0 when the new family appears, and following this to join the static BTZ geodesics at $r_0 = r_{\min}$. This outermost contour in initial condition space of $\Delta\phi = \pi$ turns out to be a boundary between initial conditions of shortest geodesics, lying outside it, and longer ones, lying inside it. In particular, the geodesics approaching close to the singularity are never shortest.

The region probed by these shortest geodesics is again covered by those with antipodal endpoints, with others reaching no deeper. In the case of small black holes $r_+ \leq 1$, it is simple to characterize, being bounded by two curves. The first is the latest radial geodesic of shortest length, with initial conditions at the critical point at which nonradial antipodal

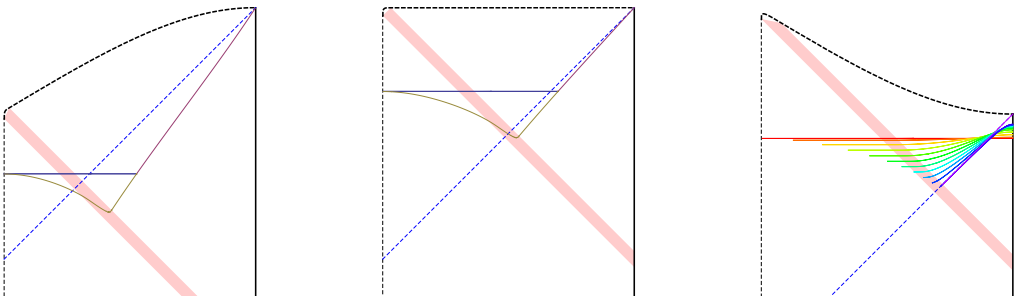


Figure 3.12: Region accessible by shortest ETEBA geodesics in Vaidya-BTZ as in fig. 3.10, plotted on the Penrose diagram. For large black hole, individual geodesics are plotted to illustrate the rounding of accessible region.

geodesics appear. The second curve is the deepest reach of the surfaces contained entirely in the static BTZ, at $r = r_{\min}$. In particular, from the time when nonradial geodesics become dominant, they never see deeper than the last radial geodesic, excepting for later points outside the minimal radius $r_{\min}$.

For large black holes, the situation is similar, with the difference that at intermediate times the geodesics 'cut the corner' inside these two curves, passing through a small additional portion of the spacetime.

These regions covered are shown in fig. 3.12 for small, critical and large black holes, along with the curve of initial conditions giving antipodal geodesics.

Finally, we take the opportunity to note how the lengths of the geodesics evolve with boundary time t_∞. In stark contrast to the higher-dimensional case, the length increases monotonically and smoothly with time, as shown in fig. 3.13. This is fortunate, as we have a more direct field theory interpretation for the observable associated with these lengths, postulated to be the entanglement entropy of the region between the endpoints. Furthermore, the early time growth, which in the case of antipodal points can be extracted from the expression in equation 3.47, agrees precisely with the results of [64, 100].

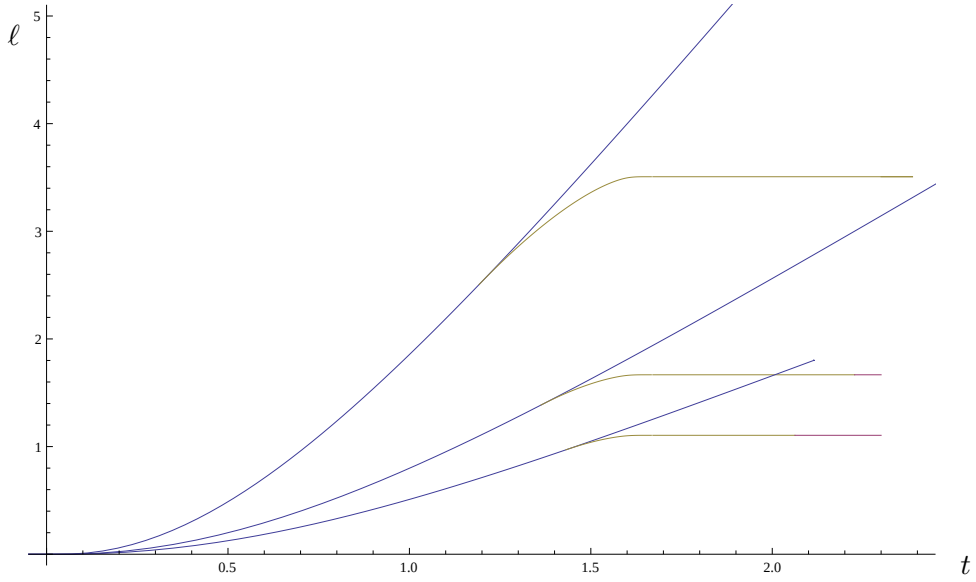


Figure 3.13: Regularised proper lengths along ETEBA geodesics in Vaidya-BTZ, plotted as a function of boundary time. Blue curves correspond to the radial geodesic branch.

3.4 Codimension-two extremal surfaces

Having considered the properties of ETEBA geodesics (which are simply one-dimensional extremal surfaces) in the previous section, we now turn to codimension-two extremal surfaces. As remarked previously, the 3-dimensional Vaidya-BTZ set up studied in section 3.3.2 is a special case of these. Here we generalise this case to higher dimensions, keeping the codimension fixed. We restrict exclusively to surfaces anchored to $(d-2)$-spheres at constant latitude, to retain an $O(d-1)$-subgroup of the $O(d)$ spherical symmetry. Further, we consider only surfaces that respect this symmetry in the bulk spacetime, which makes the great simplification of reducing the extremising equations from partial to ordinary differential equations. The experience from the geodesics in higher dimensions, where the boundary $O(1) = \mathbb{Z}_2$ symmetry of swapping insertion points need not be respected by the shortest geodesics in the bulk with given equal time endpoints, shows that this may be a genuinely restrictive assumption. However, regarding the surfaces more naturally as a generalisation of geodesics in 3 dimensions counters this concern. Moreover, in the case of extremal surfaces, the symmetry is continuous, which perhaps makes it less likely to be broken than for the discrete symmetry for geodesics.

The surface is parametrised by $d-2$ 'longitudes' ϕ on S^{d-2}, along with one other parameter, for now generically denoted by s, such that v, r, and the colatitude θ depend only on s. The area functional of the surface is given by

$$A = V(S^{d-2}) \int ds\, (r \sin\theta)^{d-2} \sqrt{-f(r,v)\,\dot{v}^2 + 2\,\dot{r}\,\dot{v} + r^2\,\dot{\theta}^2} \tag{3.22}$$

where dots denote differentiation with respect to s, and $V(S^{d-2})$ is the volume of the unit

S^{d-2}. The surfaces of interest are extrema of this functional, so the integrand acts as the Lagrangian from which the equations of motion my be obtained[14]. We have complete freedom in choice of the parameter s, setting it to be equal to one of the coordinates, or another convenient choice, for example a parameter analogous to the arc length of geodesics.

The surface must meet the poles of the S^{d-1} at $\theta = 0$ or π exactly once (excluding self-intersecting surfaces), so we can set boundary conditions at the North pole (WLOG), specifying that the surface must be smooth there. The equations of motion are singular at these points, so for the purposes of numerics, initial conditions are set by solving the first few terms in a series expansion near $\theta = 0$, and using this to begin the integration at a small positive value of θ. An exception to this rule is when the surface passes through the origin $r = 0$ (inside the shell, when $v < 0$), in which case the symmetry is enhanced, and the surface must lie entirely on the equatorial plane $\theta = \pi/2$. Details of parameters used, equations of motion, series solutions and initial conditions are given in Appendix 3.B.

One useful point from the equations of motion is that, like the geodesics, if $\dot{v} = 0$, then $\ddot{v} > 0$, so v can never have a local maximum. This tells us that the value of v on a particular surface is largest when $r \to \infty$, smallest at the initial point where it crosses the pole, and monotone between. In particular, a surface anchored to the boundary before the collapse begins can only sample the pure AdS part of the geometry, so entanglement entropy cannot evolve before the quench. It also means that v is a suitable parameter along the surface (unlike θ or r), useful because it reduces the order of the equations of motion to 4.

From the field theory point of view, there are three interesting pieces of data associated with each extremal surface ending on the boundary. The first two, θ_∞ and t_∞, respectively the colatitude and time at which the surface meets the AdS boundary (i.e. the asymptotic values of θ and v respectively as $r \to \infty$), characterise the region with which it is associated. The third, the area A, is a candidate for the entanglement entropy of the region, according to the conjecture in [39]. If there are several candidates for a given boundary region, the proposal specifies that we take the minimal area,[15] so comparison of areas in particular will be important for our purposes. As in the case of geodesics, we can think of each surface as giving a point in the 'boundary parameter space' $(\theta_\infty, t_\infty, A)$.

The area itself is divergent due to the portion of the surface heading to $r = \infty$. We regulate by cutting off the surface at a large but finite $r = r_c$, corresponding to a UV cutoff in the field theory. The leading order divergence can be computed as $V(S^{d-2})(r_c \sin\theta_\infty)^{d-2}/(d-2)$, though for $d > 3$ there are additional divergences: logarithmic for $d = 4$, and stronger as d increases. Hence, as a simple universal prescription to renormalise the area, we use background subtraction. We compute the area as a function of the cutoff radius r_c, and

[14] One may worry that an extremum of this functional may not be a true extremal surface, since there are variations away from spherical symmetry. However, we have verified that stationarity with respect to variations preserving the symmetry implies stationarity with respect to all variations.

[15] A summary of alternate definitions and subtleties in specification of the covariant entropy proposal, especially the role of the homology constraint, have been recently discussed in [1].

subtract the same quantity computed for an extremal surface in the static AdS spacetime with the same θ_∞, which can be computed analytically. As the cutoff is taken to infinity, this tends to a constant, which is defined as the renormalised area. The practicalities of this are outlined in the appendix.

The equations of motion along with smoothness condition at $\theta = 0$ have a 2-parameter family of solutions, labelled by the point (r_0, v_0) where the surface crosses the pole of the sphere. Similarly to the case of geodesics, we can think of this as a 2-dimensional 'initial condition space', parametrizing the set of spherically symmetric extremal surfaces.

Integration of the equations proceeds from this initial point, and either ends in the singularity, or continues to the AdS boundary $r \to \infty$. We are interested only in the latter, so one requirement is to find the region of the initial condition space (r_0, v_0) corresponding to these probe surfaces.

The problem of computing the entanglement entropy of regions bounded by such a sphere of constant latitude, according to the proposal in [39], is then one of identifying the surfaces with appropriate $(\theta_\infty, v_\infty)$, and finding their areas. For any given region, we expect some discrete set of surfaces; those of most interest will be the ones of minimal area amongst that set.

The main problem thus amounts to finding first the domain of the function $(r_0, v_0) \mapsto (\theta_\infty, t_\infty, A)$, for which the minimal surface ends at the boundary rather than in the singularity, and then understanding the image, a surface in $(\theta_\infty, t_\infty, A)$-space. We must bear in mind that for the purposes of computing entanglement entropy, there is an equivalence $\theta_\infty \sim \pi - \theta_\infty$, corresponding to a choice of which side surfaces may pass round the origin and related by a rotation. In particular, this is not disrupted by a homology constraint as in the globally static case [1, 73], and only connected surfaces need be considered, which is a reflection of the CFT being in a pure state.[16]

The result must interpolate between the two static geometries, namely pure AdS for $t_\infty < 0$ and Schwarzschild-AdS for large t_∞. The surface in the $(\theta_\infty, t_\infty, A)$ boundary parameter space must smoothly match up these early and late parts.

For pure AdS, this is all known analytically: the surfaces lie on constant time slices, and in fact are surfaces of revolution of geodesics. Explicitly, the map is given by $\theta_\infty = \cot^{-1} r_0$, $t_\infty = v_0 + \cot^{-1} r_0$, and $A = 0$ as a consequence of the renormalisation prescription. The domain of relevant initial conditions is the whole spacetime, and the resulting surface in $(\theta_\infty, t_\infty, A)$ space is the plane $A = 0$, for $0 < \theta < \frac{\pi}{2}$.

In the case of Schwarzschild-AdS, again the staticity greatly simplifies, since the surfaces are orthogonal to the timelike Killing field. This causes the dependence to decouple, so that θ_∞ depends only on r_0, and t_∞ reduces to the value of the static coordinate t_o on

[16] Geometrically, the homology constraint being satisfied (even in the strong form of there existing a bulk codimension-one smooth achronal surface whose only boundaries are the anchoring region and the extremal surface), follows from the fact that from arbitrarily late section of the boundary, there exists a smooth spacelike surface stretching to pre-singularity-formation era. This is most readily apparent by considering the Penrose diagram, and follows directly from the spacelike nature of the curvature singularity and the fact that there is only a single asymptotic region.

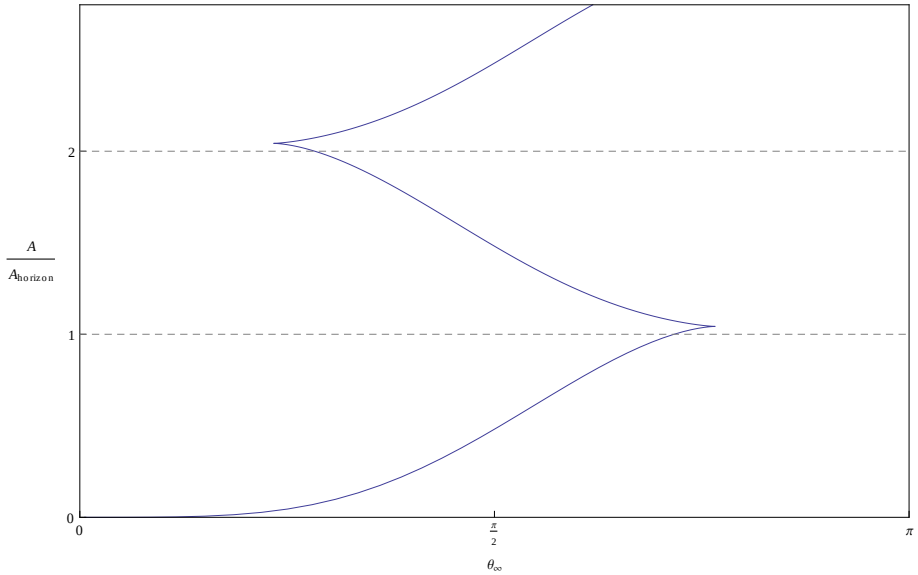


Figure 3.14: Renormalised areas of connected minimal surfaces relative to the area of the event horizon plotted against their colatitude θ_∞ at the boundary. The tower of surfaces continues indefinitely, repeating in an essentially periodic pattern.

which the surface lies. The naïve expectation here is that any boundary region should be matched by two minimal surfaces, one passing round each side of the event horizon, but the actual story is more complicated, and is described in detail in [1]. There is an infinite tower of surfaces for sufficiently large boundary regions, shown in fig. 3.14. Only the lowest branch of this tower, up to $\theta_\infty = \pi/2$, is directly relevant for computing entanglement entropy in the static case, since the others all have larger area. The domain of relevant initial conditions is the exterior of the event horizon, and the surface in $(\theta_\infty, t_\infty, A)$ space is a translation in the t-direction of the multi-branched shape of fig. 3.14.

It is useful to bear these static cases in mind, since during the collapse, the surface in $(\theta_\infty, t_\infty, A)$ must interpolate between these regions, and join the simple plane pre-collapse to the complicated folded surface post-collapse. To solve the puzzle of how this can occur, we turn to numerical studies. The computations were made almost exclusively in AdS_5 (i.e. $d = 4$), with horizon radius $r_+ = 1$, though comparisons with other parameters indicate that the results are generic. Details of the methods are given in Appendix 3.B.

Domain of relevant initial conditions, and an interesting class of surfaces Numerical studies to identify the region of relevant initial conditions show that it is characterised by a very simple critical curve, which can be defined by a function $r_c(v_0)$. This is shown in fig. 3.15. Initial conditions (r_0, v_0) give a surface ending on the boundary when $r_0 > r_c(v_0)$. (Note that the actual surfaces themselves can penetrate past $r_c(v_0)$, as we will see momentarily, but their 'initial condition point' is restricted by $r_c(v_0)$.) After the collapse of the shell, the critical curve coincides with the event horizon, so r_c equals r_+. Before the collapse, the curve meets the origin, so r_c vanishes at a particular value of

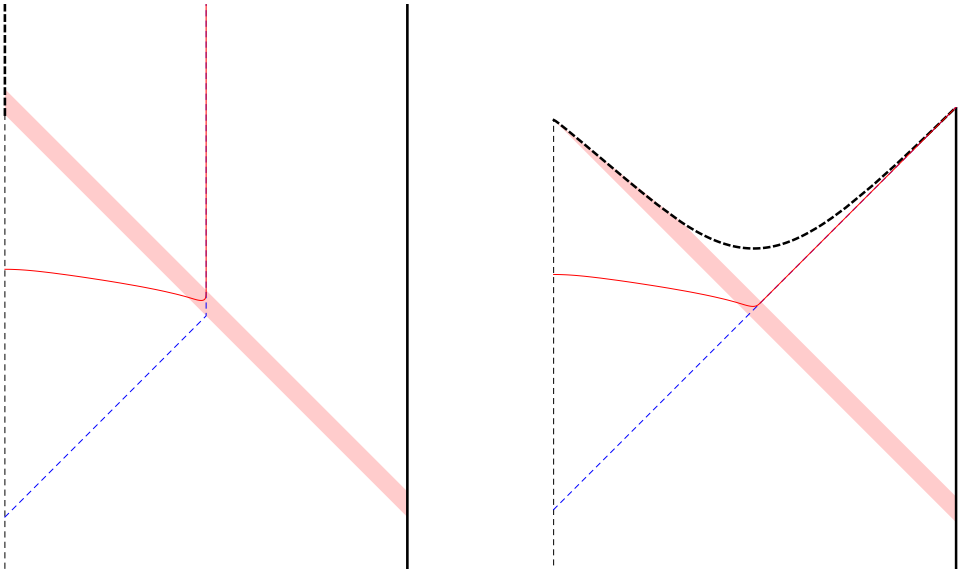


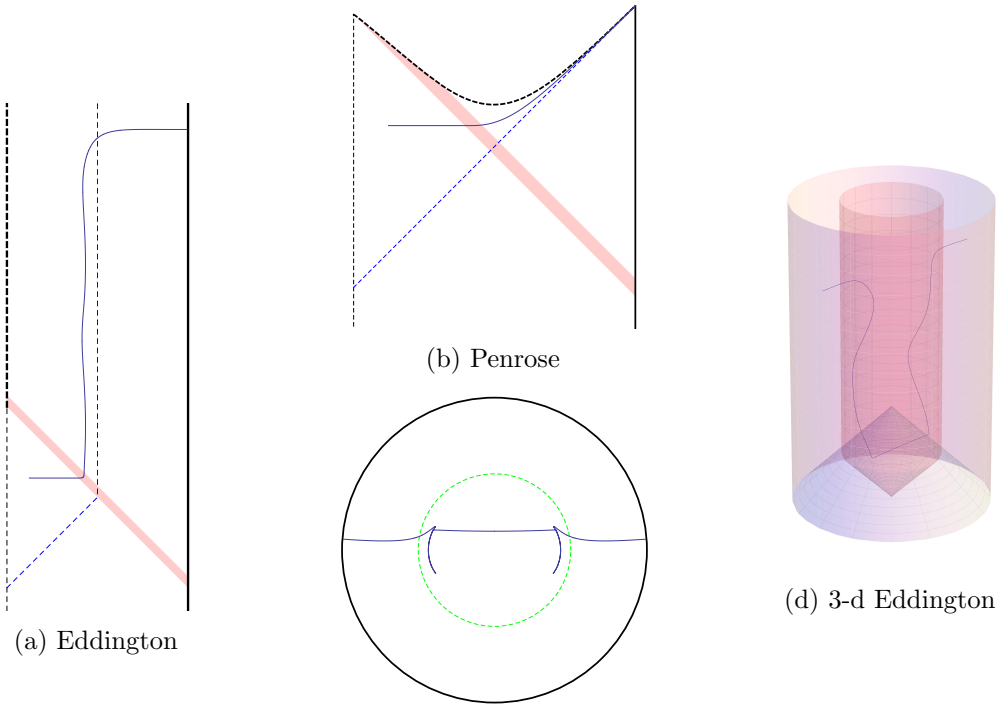
Figure 3.15: The region of relevant initial conditions. Surfaces whose earliest point (with respect to the ingong time v) lie outside or inside the curve reach the boundary or end in the singularity respectively.

v_0, before which time all surfaces will reach the boundary, for any r_0. This part of the critical curve lies entirely inside the event horizon. In particular, this shows already that codimension-two extremal surfaces anchored on the boundary do reach within the horizon.

In the process of finding this critical surface, we found a particularly interesting class of solutions, a typical example of which is shown in fig. 3.16. If the initial conditions are chosen close to the critical curve, the resulting solution lies inside the event horizon along much of its length, at nearly constant r, moving outside the horizon and to the boundary at a late time. By careful tuning of the initial conditions, extremal surfaces meeting the boundary at arbitrarily late times can be constructed. These surfaces link up to the multiply wrapping surfaces in the static black hole geometry; indeed it was these surfaces in the dynamic Vaidya geometry which led to the discovery of the tower of surfaces in the static case. They can be understood from an analytic study, to which we now turn.

Consider first surfaces in the Schwarzschild-AdS part of the spacetime which lie at constant radius r, on the equator $\theta = \pi/2$, extended in the v (or t) direction. Inside the event horizon, these are spacelike. Taking r to be small, approaching the singularity, their area (per unit extent in the v direction) reduces to zero by virtue of the shrinking in the spherical directions (for $d > 2$). Taking r larger, approaching the event horizon, the area also reduces to zero, this time by virtue of the surface becoming null. Between these two extrema, there must be a maximal area surface[17].

[17] The former effect of shrinking the sphere is absent for geodesics, which is essentially why they are different. We would expect surfaces between the two cases discussed here, with dimension larger than one, but smaller than $d-1$, to behave more like the codimension-two surfaces than the geodesics for essentially this reason.



(a) Eddington

(b) Penrose

(c) Poincaré

(d) 3-d Eddington

Figure 3.16: A surface whose initial conditions are very close to the critical value in the fig. 3.15. By tuning this closeness to be exponentially small, the boundary time t_∞ may be made as late as desired.

Indeed, the equations of motion for extremal surfaces in the static black hole geometry admit an exact solution at constant $r = r_* < r_+$, on the equatorial plane $\theta = \pi/2$. In the globally static spacetime, a perturbation of this surface must either meet the singularity, or form a tube connecting the two asymptotically AdS regions, as in [63]. In constrast, the Vaidya geometry allows for extremal surfaces lying close to this critical radius for much of their extent, but smooth everywhere, and terminating on the boundary, as seen in the numerics.

We look next at the equations of motion, in the Schwarzschild-AdS geometry, linearised around this constant radius solution. We parameterise by v, primes denoting differentiation with respect to v. The radius of maximal area r_* satisfies

$$f'(r_*) + 2\,\frac{d-2}{r_*}\,f(r_*) = 0, \tag{3.23}$$

which has a unique solution (when $d > 2$). Linearising around this, with $r = r_* + \rho$ and $\theta = \frac{\pi}{2} + \eta$, the equations of motion decouple, and reduce to

$$\rho'' - \lambda^2\,\rho = 0 \tag{3.24}$$

$$\eta'' + \omega^2\,\eta = 0 \tag{3.25}$$

with parameters given by

$$\lambda^2 = \frac{(2d-3)(d-2)f(r_*)^2}{r_*^2} - \frac{1}{2}f(r_*)f''(r_*) \tag{3.26}$$

$$\omega^2 = -\frac{(d-2)f(r_*)}{r_*^2}, \tag{3.27}$$

which are both positive when f has the form of Schwarzschild-AdS.

A case of particular interest is when the surface meets the boundary at late times. For these, the growing mode for ρ will be tuned close to zero, so ρ will be much smaller than η, and η^2 terms become relevant at leading order for ρ. We include this forcing for ρ by solving for η as $\theta = \pi/2 + a\cos(\omega v)$, keeping terms up to order a^2, to get

$$\rho'' - \lambda^2\rho = -a^2(d-1)(d-2)\frac{f(r_*)^2}{r_*}\sin^2(\omega v), \tag{3.28}$$

which has a particular solution

$$\rho = a^2(d-1)(d-2)\frac{f(r_*)^2}{2r_*}\left(\frac{1}{\lambda^2} - \frac{\cos(2\omega v)}{\lambda^2 + 4\omega^2}\right). \tag{3.29}$$

This explains the oscillations in the radial direction visible in fig. 3.16, showing that r is largest when $\theta = \pi/2$, and smaller where the surface is further from the equator. Of particular note is that this particular solution is strictly positive, which constrains the surface to lie outside $r = r_*$, a point to which we return later.

Not every possible solution in Schwarzschild-AdS will give an extremal surface when

continued into the full spacetime, since generically this continuation will not be smooth at the pole of the sphere: indeed, in Schwarzschild-AdS, there are four parameters describing the solutions, but there is only a two parameter family smooth solutions in Vaidya. Consider starting at an initial point inside the shell, with the two initial condition parameters (r_0, v_0), and integrating until reaching the outside of the shell. After this, our surface is well described by our analytic solution. In this way, the initial conditions map to values for the four constants of integration: the amplitude a and phase of the angular oscillations, and the growing and shrinking modes of ρ. Further, this map should be smooth. For surfaces for which the linearisation is a good approximation, this observation alone will tell us much.

The fate of an extremal surface is characterised, in this approximation, by the sign of the coefficient of the growing mode $e^{\lambda v}$ of ρ, which we denote by g. It will escape to the boundary if g is positive, or end in the singularity if g is negative. The limiting case, when $g = 0$, will give the critical curve in the (v_0, r_0) initial conditions plane separating surfaces ending on the boundary from those ending in the singularity.

Boundary region from initial conditions Having characterised the domain of relevant initial conditions, we would now like to understand how they correspond to a region on the boundary. Initially, the analytic study will take us a long way towards this goal.

When $g > 0$, it will roughly tell us the time at which the surface meets the boundary, since we expect $g \approx e^{-\lambda v_\infty}$. Since g changes smoothly with the initial conditions, barring any coincidences, it should go to zero linearly as the critical curve is approached, along a line of constant v_0, say. This implies that t_∞ should diverge logarithmically near the edge of initial condition parameter space:

$$t_\infty \sim -\frac{1}{\lambda} \log(r_0 - r_c(v_0)) \tag{3.30}$$

In particular, the surface can reach the boundary at arbitrarily late time, by tuning initial conditions exponentially accurately.

This also corresponds nicely to what happens for inital conditions near the horizon in the static black hole part of the geometry. Surfaces fully in this region lie on a constant-time slices with respect to the static coordinate t_o, and, approaching the future horizon, t_o blows up logarithmically.

Close to the critical curve, changes in the other parameters will be unimportant relative to the effect of g going to zero. Fixing v_0, and considering r_0 very close to r_c, the shape of the surface should be almost unchanged, except for the time at which the growing mode of ρ kicks in and drives the solution out to the boundary. Hence, we expect the value of θ_∞ to be determined largely by the phase of η when the growth of ρ begins.

Since η undergoes oscillations in v, this means that as the critical curve is approached, the blow-up of v_∞ is accompanied by increasingly rapid oscillations in the value of θ_∞.

The angular frequency in t_∞ will be ω, so we expect

$$\theta_\infty \sim a \cos\left(\frac{\omega}{\lambda}\log(r_0 - r_c(v_0)) + \text{phase}\right) \tag{3.31}$$

with the phase depending on the choice of v_0 along which the critical curve is approached. The range of θ_∞ covered is determined by the amplitude a of the η oscillations, which depends most strongly on r_0. It must vanish, by the enhanced symmetry, when $r_0 = 0$, and should increase with increasing r_0, as the surfaces depart further from the equatorial plane $\theta = \pi/2$.

Again, this matches the behaviour seen in the static Schwarzschild-AdS spacetime, where there are similar oscillations as the initial conditions approach the event horizon.

The numerical results show all these features. In particular, the details close to the edge of the initial condition space match the expectation from the analytic calculations, including the rate of blow-up of t_∞, and the period of oscillations of θ_∞.

Contour plots, representing how the boundary region associated to a given surface corresponds to its initial point where it crosses the pole of the S^3, are shown in fig. 3.17.

The contours of constant θ_∞ are particularly interesting, as they correspond to a family of curves associated with entanglement entropy for a specific region of space in the field theory. We see the first few curves of what we expect to be an infinite collection for a given region (as long as that region is not too small). These continue into the black hole part of the geometry, where they correspond to the tower of surfaces of [1].

In terms of the surface in boundary parameter space $(\theta_\infty, t_\infty, A)$ we have enough to build a qualitative picture of what goes on. The surface will have an edge corresponding to the equatorial surfaces lying entirely on $\theta = \pi/2$, which is the image of the initial conditions $r_0 = 0, v_0 < 0$. The other boundary of initial condition space (the curve of fig. 3.15) maps to $t_\infty \to \infty$, so the only other edge of the surface is at $\theta_\infty = 0, A = 0$, when the initial conditions approach the AdS boundary. The surface thus looks like a strip, which for $v < 0$ is just the flat plane $0 < \theta < \pi/2, A = 0$, and thereafter progressively folds over itself to link up with the tower of fig. 3.14 one branch at a time. The beginning of the first such folding is shown in fig. 3.18.

Spacetime region covered Of particular interest is the region in spacetime covered by the probe extremal surfaces. As long as there are no unexpected departures from the approximation scheme, the analytic study gives us a complete characterisation of this. After the collapse of the shell, extremal surfaces reach precisely the region of the bulk outside the radius $r = r_*$ (perhaps excepting 'cutting the corner' very near the shell, allowing for a negative decaying mode of ρ). Given this, it is not unreasonable to expect that the covered region of spacetime is bounded by the surface on the equator $\theta = \pi/2$ (so $r_0 = 0$), with v_0 chosen such that $r \to r_*$ as $v \to \infty$. This is the critical value between surfaces ending on the boundary or in the singularity. This expectation is borne out by the numerics, and the region covered is shown in fig. 3.19.

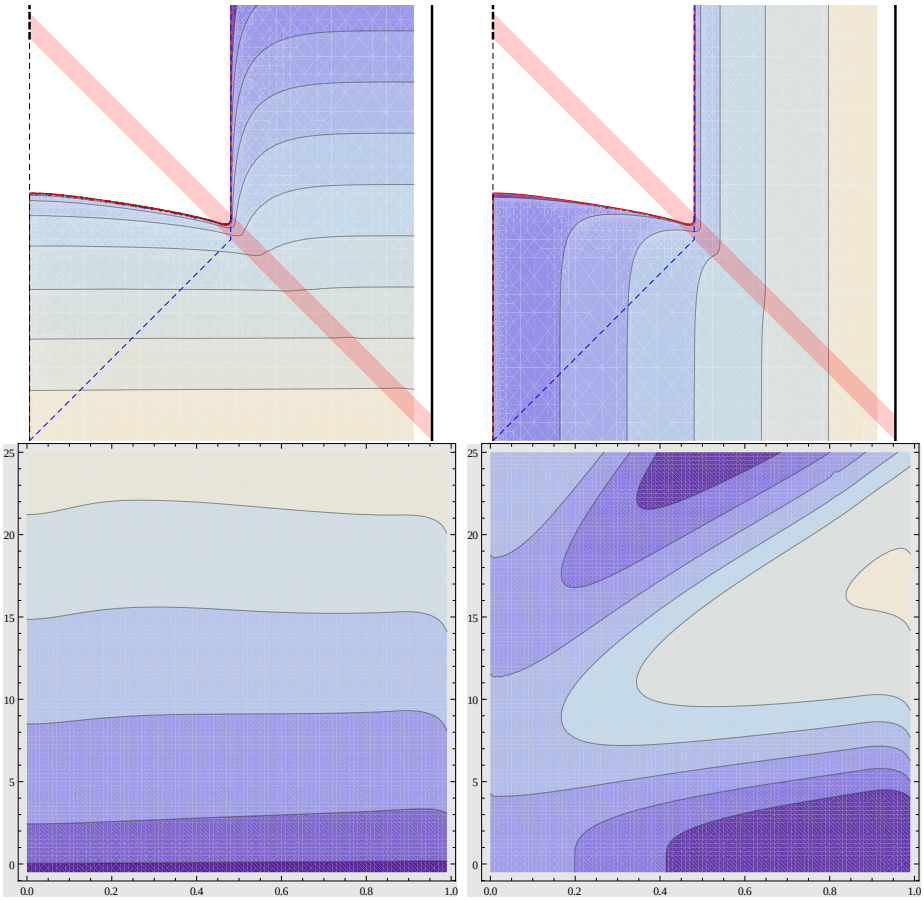


Figure 3.17: Contour plots for t_∞ (left) and θ_∞ (right) as a function of initial conditions (v_0, r_0). The red curve is the edge of the relevant initial conditions (as in fig. 3.15). The values of the contours in the top right figure are at multiples of $\pi/12$; part of the outermost contour for $\theta_\infty = \pi/2$ is just visible. The bottom figures show the detail close to the edge of the region of relevant initial conditions. The horizontal coordinate is r_0, and the vertical coordinate $-\log(v_c(r_0) - v_0)$, where $v_c(r_0)$ gives the latest relevant initial v_0 for given r_0.

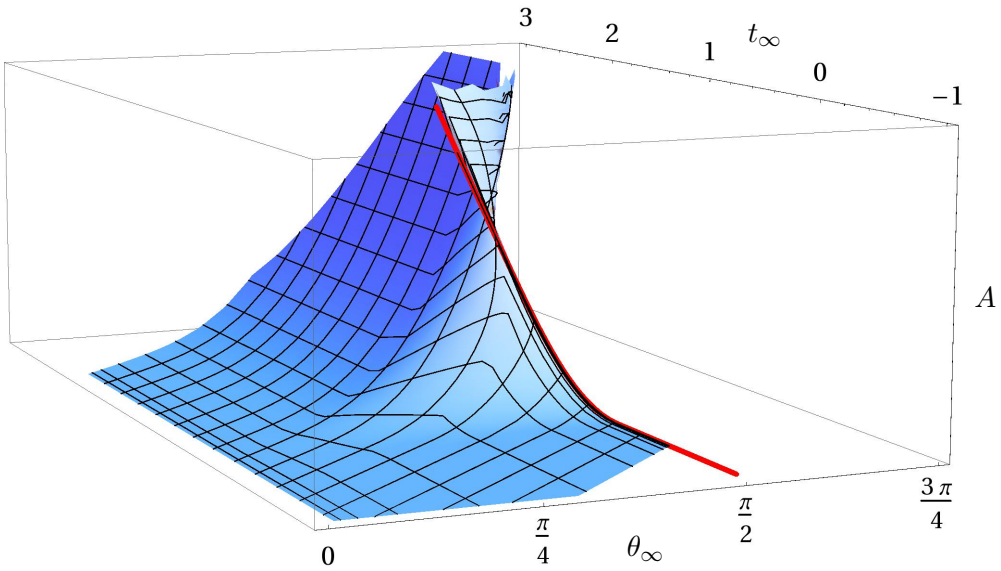


Figure 3.18: Part of the surface in boundary parameter space $(\theta_\infty, t_\infty, A)$ induced by the extremal surfaces. The red curve shows the edge of the surface, which corresponds to surfaces lying on the equator $\theta = \pi/2$, and passing through the origin with $r_0 = 0, v_0 < 0$.

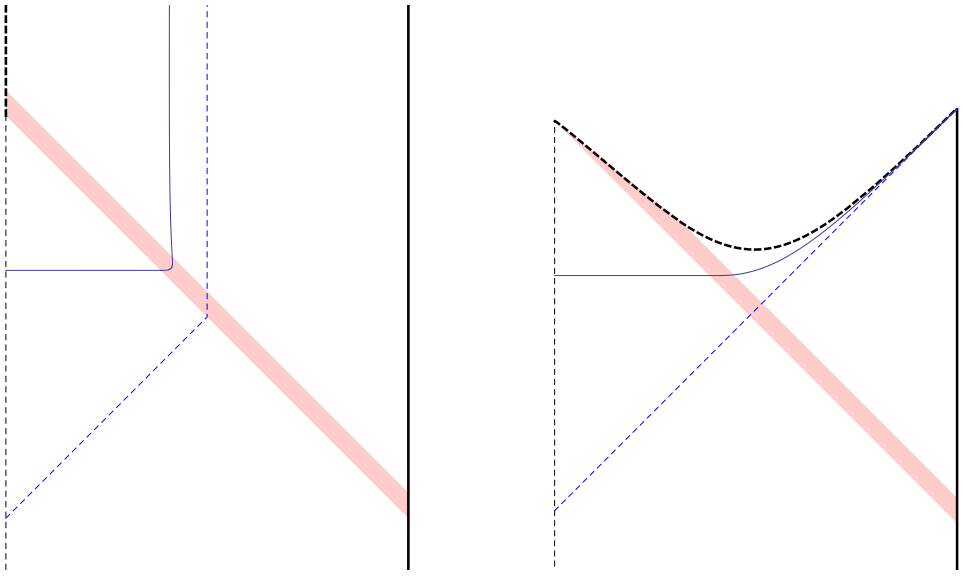


Figure 3.19: The region covered by all extremal surfaces reaching the boundary. The edge is the same as the limiting surface between surfaces ending on the boundary or in the singularity, lying on the equatorial plane, and at constant $r = r_*$ after the collapse.

Surfaces of minimal area, and entanglement entropy We now turn to the measurement of the field theory quantity of interest, namely the area of the surfaces. This is interesting for at least two reasons. Firstly, we can learn about the thermalization of entanglement entropy for a field theory on a sphere. Secondly, we would like to compare areas of surfaces anchored to the same boundary region, because the extremal surface of least area is most directly associated with the field theory observable. This will allow us to refine our description of the spacetime region covered by surfaces, to include only those of minimal area for a given boundary region.

The simplest possibility is that the relevant surfaces of minimal area are those arising from a continuous deformation between the static parts of the geometry. In terms of the initial conditions, this will correspond to the outside of the outermost $\theta_\infty = \pi/2$ contour (see fig. 3.17). In principle, this could be spoilt by the higher branches of surface: for example, one could imagine a case where the folding portion of the surface in boundary parameter space (fig. 3.18), where $\theta_\infty > \pi/2$, dips below the corresponding piece under the equivalence $\theta_\infty \sim \pi - \theta_\infty$. If this were to occur, it would allow for some novel behaviour of entanglement entropy, such as kinks, discontinuities, and non-monotonicity as a function of time, where two branches of surfaces exchange dominance or new branches appear. This turns out not to be realised in the case at hand, though we know of nothing which would prevent it; it would not be entirely dissimilar to what we have found in the case of geodesics. It would be interesting to see if these possibilities can be excluded, or instead found to be present in an altered geometry.

With the absence of these complications, the thermalization of entanglement entropy offers nothing new, as shown in fig. 3.20, smoothly and monotonically increasing from the vacuum to the thermal value. This matches well with the findings of [64, 100], who undertake a similar study in the planar case. Here we might a-priori have expected the physics to be much richer, but that expectation has not been realised. In particular, there is an intermediate regime where the area grows linearly, controlled by the surface extending along the critical radius $r = r_*$. Additionally, if the is shell sufficiently thin, at early times there is a quadratic growth with known coefficient, proportional to the area of the bounding region (in our case an S^{d-2}) and the energy density, since the calculation of [64] goes through unchanged. Unsurprisingly, this is modified to a slower growth if the collapse is more gradual. We found no clear robust law found governing the final approach to equilibrium, excepting that it appears to be smooth, though a more thorough investigation of this would be worthwhile.

The part of spacetime covered by the extremal surfaces of minimal area is accurately characterised as the outside of either of two regions. The first is given by the deepest point of the minimal surface giving the entanglement between hemispheres in Schwarzschild-AdS, which is a value of r strictly larger than r_+, important after the collapse. The second is the latest surface giving entanglement between hemispheres which passes through the origin. The initial conditions of this are given by the meeting of the outermost $\theta = \pi/2$ contour of fig. 3.17 with the origin. This surface samples the inside of the horizon, including for a

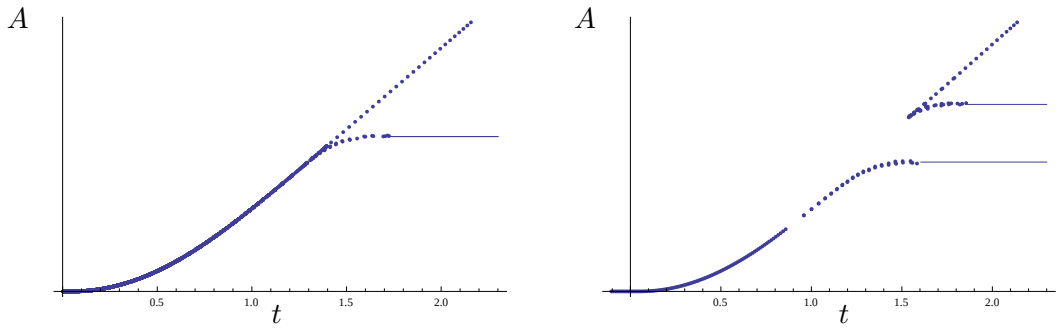


Figure 3.20: Thermalization of entanglement entropy for a hemisphere (left), and a slightly smaller region (right). Data points for extra branches of surfaces are shown, and can be seen to have larger areas.

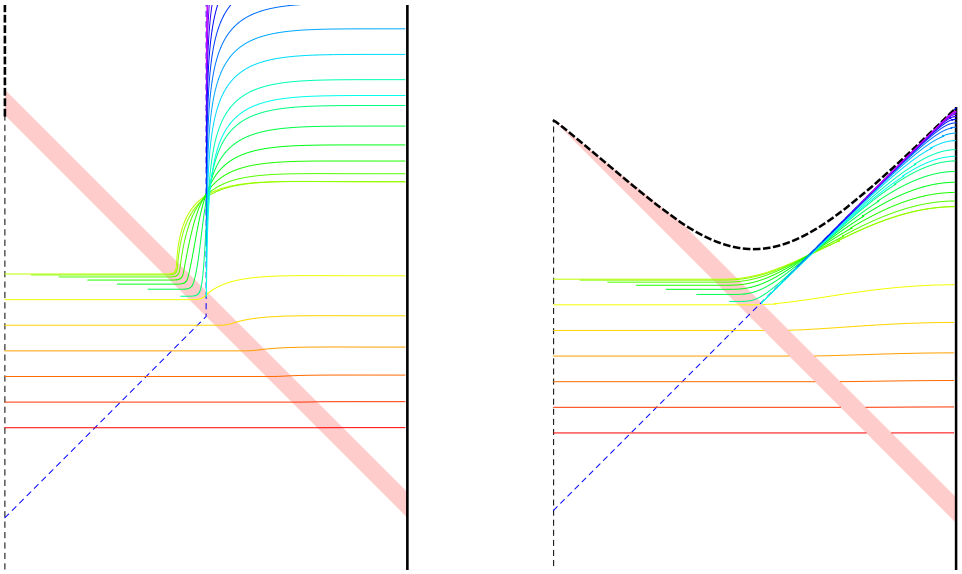


Figure 3.21: A selection of the surfaces corresponding to entanglement between hemispheres. The innermost of these characterises the region of the bulk probed by these observables.

significant time outside the shell, but is bounded well away from the singularity and hence is protected from regions of strong curvature. This is illustrated in fig. 3.21.

3.5 Discussion

We have explored the behaviour of spacelike boundary-anchored geodesics and codimension-two extremal surfaces in a spherically symmetric Vaidya-AdS bulk spacetime. The main goal was to assess how deep past the horizon can such surfaces reach in this simple model of a collapsing black hole. This was motivated in part by the observation of [61], that unlike for a static spacetime wherein such boundary-anchored extremal surfaces cannot penetrate a black hole,[18] in dynamically evolving spacetime, the event horizon, being

[18] In fact, there is an interesting subtlety in this argument, partly analogous to the geodesic behaviour we exploited in the present context: [61] used conservation of 'energy' along a geodesic to argue that in

globally defined, does not pose a fundamental obstruction. We realized this expectation both with geodesics as well as with codimension-two extremal surfaces, and in all cases there are even surfaces passing inside the horizon at arbitrarily late times. Having said this, the actual bulk regions probed are qualitatively rather different.

Geodesics with endpoints on the boundary at equal times are able to reach arbitrarily close to the singularity when it forms, accessing regions of arbitrarily strong curvature. The same is not true of codimension-two extremal surfaces, which remain bounded away from the singularity. However, these access a much larger region later on, being constrained only to lie outside the surface of maximal area at constant radius, while the geodesics probe a region shrinking to the horizon at later times.

The exception to this is in the special case of 3-dimensional bulk, where the two classes coincide, and the behaviour depends crucially on the final size of the black hole. For black holes no larger than the critical radius $r_+ = 1$ in AdS units, even the geodesics with endpoints at equal times cover the whole spacetime. For black holes larger than this, there is a region inaccessible to all geodesics with both endpoints on the boundary, even without the restriction to equal-time endpoints, around the singularity when it first forms. At late times, the geodesics still can approach arbitrarily close to the singularity.

In all cases, restricting considerations to the surfaces expected to dominate the CFT variable, of smallest area or shortest length, puts constraints on the probed region. In particular, none of these surfaces reach inside the horizon for an extended time after the shell passes.

Unsurprisingly, the surfaces reaching the deepest are consistently those corresponding to the largest possible length scales, namely geodesics connecting antipodal points, and surfaces corresponding to entanglement between hemispheres. Access to the largest part of the bulk requires knowledge of the correlations on the biggest scales.

Let us pause briefly to consider what attributes of the geometry enabled the novel features described above such as ETEBA geodesics penetrating to regions of arbitrarily high curvature. The most obvious feature is the rapid time-dependence. Indeed, the shell has a dramatic effect on geodesics which traverse it, especially near the point of implosion. But compared to previous studies of extremal surface probes in Vaidya-AdS, further novel features arise due to the compactness of the horizon, i.e. by considering collapsing black hole with spherical, rather than planar, symmetry. From the field theory point of view, this enables us to access finite-volume effects. Indeed, we have seen that some of the surprising features arise only when the boundary endpoints are sufficiently nearly antipodal. From the bulk standpoint, there are two different effects of the spherical geometry. The one

static spacetimes, once a geodesic enters the future event horizon, it cannot exit back through the future horizon. This argument didn't prevent the geodesic from reaching the same boundary via the past horizon, but [61] argued that in order to do that, it would have to turn around in the 'right' exterior region, which, *assuming it has the same radial profile of the metric as the left region*, would contradict the assumption that the geodesic reached from the left boundary to the horizon. One may, however, consider a more contrived spacetime with a static shell on the right side of Schwarzschild-*AdS* beyond which the right boundary is replaced by e.g. de Sitter region with smooth origin as in [53]. In such a situation, as discussed in that work, there certainly do exist boundary-anchored spacelike geodesics passing through the black hole.

which is most crucial for the asymmetric radial geodesics is the fact that prior to the shell, the geometry has a smooth origin through which the geodesic may pass, heading back out to the boundary rather than through the Poincaré horizon which replaces it in the planar geometries. The other effect, which is most crucial for the extremal surfaces, is that surfaces can pass around the black hole.

In [1] we considered the question of whether in a fixed bulk geometry, the area of smallest-area extremal surface can jump discontinuously as a function of the parameters specifying the surface (namely the size and time of the boundary region on which this bulk surface is anchored). We argued that in the case of static spacetimes, where the extremal surface is in fact a *minimal* surface on a constant-time slice as required by the Ryu-Takayanagi prescription [30, 31], the area must vary continuously, which implies that correspondingly the entanglement entropy must vary continuously.[19] The argument however relied on minimality, so that the corresponding issue was not clear for the broader class of extremal surfaces which is relevant for time-dependent bulk geometries.

Indeed, the most general statement to the effect that areas of smallest-area extremal surfaces vary continuously with the boundary conditions is false. The results of section 3.3 provide a manifest counter-example in the case of geodesics, being one-dimensional extremal surfaces, since we saw in fig. 3.7 that the minimal length $\ell(t)$ is discontinuous. This is in sharp contrast to the naive thermalization picture where we would expect this quantity to grow monotonically and thermalize.

On the other hand we found that found that, in all situations we considered, codimension-two extremal surfaces do vary continuously, thus rescuing entanglement entropy from the bizarre contingency of discontinuous jumps. The question of whether this is true in general remains open, though its failure for geodesics provides some guidance for any attempt at a proof, by ruling out arguments that would encompass all extremal surfaces. On the other hand, if there are situations in which the area may vary discontinuously, looking at asymptotically locally AdS spacetimes with non-planar boundary topology seems a good place to search for counterexamples. As seen here, this allows for a richer structure, with multiple branches of surfaces, which is likely to be a minimum requirement for a discontinuous exchange of dominance.

We now turn to the question of interpreting the results obtained for the lengths of the geodesics. Using spacelike geodesics to probe the black hole has a long history; see e.g. [112], [113], [109], [108], [114], [115], and more recently revisited in [65], [116], [117]. In the present work, perhaps the most fascinating result is the striking contrast between the conventional thermalization picture and the non-monotonic, discontinuous behaviour of the length $\ell(t)$ along shortest ETEBA geodesics with endpoints at time t, antipodally-separated, as illustrated in fig. 3.7. Translated directly to the corresponding expectation for the equal-time CFT correlator of high-dimension operators, $\langle\phi\phi\rangle(t) \sim e^{-m\,\ell(t)}$, this would be very bizarre. However, we do not expect this to hold due to the subtlety that the geodesic may not lie on the path of steepest descent.

[19] See also the discussion in [111] which appeared concurrently with our work.

As argued in [108] in the eternal Schwarzschild-AdS context (see also [115] for a complementary approach and [65] for more recent discussion in a broader context, more germane to the present case of interest), if the CFT correlator were dominated by the shortest spacelike geodesic (which bounces off the black hole singularity), the correlator would become singular when the insertion points are such that the joining geodesic approaches being null, which is ruled out by direct considerations in the CFT. The resolution of this apparent puzzle comes from the fact that there are multiple (complexified) geodesics connecting the boundary points. At the time-reflection-symmetric point they coincide, signalled by a branch point in the correlator. By considering the resolution of this branch point, [108] was able to show that the correlator is given by a sum of two complex branches. But since the correlator is an analytic function in the position and time of the insertion points, one can recover the expected 'light cone singularity' by analytic continuation.

Here the situation is more complicated, since these methods are explicitly inapplicable if the spacetime itself is not analytic – as in the present case of the shell having compact support. This by itself might be circumvented by considering an analytic spacetime (which can be arbitrarily close to the present geometry and therefore the behaviour of the real geodesics will likewise be arbitrarily close to the present case), by making the shell profile analytic. However, in this dynamical situation we will have lost the standard crutch of being able to use the Euclidean continuation. In [65] the authors consider such a situation, involving a shock wave in BTZ which is nonanalytic, and analytic approximations do not have real Euclidean continuations. For small non-analytic perturbations of the metric, the authors argue that indeed the saddle point represented by the perturbed geodesic continues to give the dominant contribution to the two-point function when the unperturbed one does. On the other hand, in the higher dimensional case where the shortest nearly-null geodesic of [108] did not dominate the correlator of the eternal Schwarzschild-AdS geometry, a 'corresponding' geodesic continued from Schwarzschild-AdS to our Vaidya-AdS geometry will probably likewise not dominate. It would be useful to develop a robust and universal criterion for directly determining when a given shortest geodesic dominates the corresponding CFT correlator, without recourse to solving the wave equation.

In the second part we focused on spacelike surfaces which are anchored on a round $d-2$ sphere at constant time in d-dimensional boundary. The motivation for this restriction was two-fold: from the pragmatic standpoint, this is the case which is simplest to solve when the bulk spacetime is spherically symmetric. Although the specification of the entangling surface necessarily breaks the full symmetry, for spherical regions we retain $SO(d-1)$, which is inherited by the full extremal surface. This is an enormous simplification, since the extremal surface is determined by coupled ODEs rather than PDEs. By itself, this is a looking under the lamppost type motivation; however, choosing spherical regions has a separate reason, based on the expectation that for a fixed extent of the entangling surface on the boundary, the corresponding extremal surface reaches the deepest into the

bulk.[20] Hence for the question of how deep into the black hole can extremal surfaces reach, spherical entangling regions seem like the 'best bet'. However it would be useful to verify this expectation explicitly, by considering other entangling surfaces. It would also be interesting to consider disconnected boundary regions, with multiple entangling surfaces. Is there a constellation allowing the corresponding extremal surfaces to probe still deeper?

We have seen that the two types of probes we focused on, spacelike geodesics and codimension-two extremal surfaces, both probe inside the genuine black hole, but that each class accesses a different region inside the horizon. One might then ask if there is a natural geometrical characterization of the region probed, without making direct reference to the probes. In other words, is there any special meaning to this region, especially from the CFT standpoint? Such a characterization cannot be global like the event horizon, nor can it be quasi-local (spacetime foliation-dependent) like the apparent horizon.[21] It should also be something which only requires knowledge of the local part of the geometry, but at the same time it should allow for the richness of changing with dimension.

A weaker version of this question is whether there may be simple considerations giving bounds on the accessible region. In the present work, one likely candidate for such a bound was the critical surface corresponding to the maximal-area constant-r surface inside the black hole, which we saw characterized the region covered very simply. Some steps in this direction have been made in [119], which provided criteria for surfaces bounding the region accessible by extremal surfaces anchored on a boundary. While no bound exists for geodesics, which access the whole spacetime, in may be possible to use their results to say something about codimension-two surfaces. It would be interesting to see how much can be said on such general grounds, and in particular how closely any bounds thus constructed come to characterizing the actual region probed.

Apart from the understanding they might provide in their own right, such bounds have a possible practical purpose in a study of thermalization requiring numerical evolution of the spacetime, such as formation of a black hole after sourcing some specific CFT operator for a time. The numerics do not allow for evolving the spacetime all the way to the singularity, but to study entanglement entropy via extremal surfaces, stopping at the horizon will be inadequate. The radius of maximal area, or any other more general bound that could be found, give a natural intermediate place to stop integration.[22]

Since geodesics (1-dimensional extremal surfaces) and codimension-two (in $d+1$ dimensional bulk) extremal surfaces have such different behaviour in terms of their reach when $d > 2$, one might naturally ask what happens to n-dimensional extremal surfaces with $1 < n < d-1$ when $d > 3$. For example in $d = 4$, string worldsheets corresponding to a Wilson loop would constitute such an intermediate case. While we expect the qualitative behaviour to be close to the case of the codimension-two surfaces, a more full comparison

[20] This was argued in [61] in case of planar AdS: deforming the entangling surface on the boundary, while keeping its extent (or volume enclosed) fixed, makes the bulk extremal surface recede towards the boundary.

[21] A similar issue for planar charged collapsing shell was recently considered in [118] which discussed connection between surfaces reaching past apparent horizon and strong subadditivity violation.

[22] We thank M. Headrick for this suggestion.

would be needed to check whether this is borne out. Another natural generalization to consider would be more general spacetimes, for example adding charge, as in [118], and even causally trivial spacetimes may give interesting results (for example, see work on scalar solitons as in [120, 121]).

3.A Geodesics in Vaidya-BTZ

In this appendix we collect the details of the calculations of geodesics in the $d = 2$ case of BTZ, in the limit of a thin shell, used in section 3.3.2.

Firstly, we give the change of coordinates to the Penrose diagram as described in section 3.2. In the BTZ part of the spacetime, after the collapse, for $v > 0$, the coordinate transformation is given by

$$v = \frac{2}{r_+} \coth^{-1}\left(\frac{1}{r_+}\cot\frac{V}{2}\right) \tag{3.32}$$

$$r = -\frac{r_+^2 \tan\frac{V}{2} + \tan\frac{U}{2}}{1 + \tan\frac{U}{2}\tan\frac{V}{2}} \tag{3.33}$$

and in the pure AdS by

$$v = V \tag{3.34}$$

$$r = \tan\left(\frac{V-U}{2}\right) \tag{3.35}$$

Writing $T = \frac{V+U}{2}$ and $R = \frac{V-U}{2}$, the metric is

$$ds^2 = \frac{-dT^2 + dR^2}{\cos^2 R} + r(T,R)^2 d\phi^2. \tag{3.36}$$

One point that this choice of coordinates makes clear is that the metric is in fact continuous, which is not evident from the original coordinates. This implies that the tangent vectors of geodesics will change continuously across the shell, with no kink.

A striking feature of these coordinates is that the $T - R$ part of the metric is identical to pure AdS. It should be emphasized that this does not happen in higher dimensions, but is special to the BTZ case.

The most useful equations of motion will be:

$$\dot{r}^2 = E^2 + \left(1 - \frac{L^2}{r^2}\right) f(r) \tag{3.37}$$

$$\dot{v} = \frac{\dot{r} + E}{f(r)} \tag{3.38}$$

To match cross the shell, we use the fact that $\dot{v}$ is continuous. To get energy after shell

crossing, we eliminate $\dot{r}$ from the above and use

$$E = \frac{f(r)\dot{v}}{2} - \frac{1}{2\dot{v}}\left(1 - \frac{L^2}{r^2}\right). \tag{3.39}$$

We first record the solutions for symmetric geodesics in the pure AdS part of the geometry, corresponding to zero energy there. The initial condition will be parameterized by the angular momentum L, which is also the minimum of r, and the static time slice it starts on, labelled by $\tau = \tan^{-1} r_0 - v_0$, lying in the range $(\tan^{-1} L, \pi/2)$. For radial geodesics, starting at the origin, this gives the time before the formation of the shell; it is larger for starting points further in the past. The solution is

$$\begin{aligned} r &= \sqrt{L^2\,\cosh^2 s + \sinh^2 s} \\ \dot{v} &= \frac{\tanh s}{\sqrt{L^2\,\cosh^2 s + \sinh^2 s}} \\ v &= \tan^{-1}\sqrt{L^2\,\cosh^2 s + \sinh^2 s} - \tau \\ \varphi &= -\tan^{-1}(L\,\coth s) \end{aligned}$$

The shell is hit at

$$s = \cosh^{-1}\frac{\sec\tau}{\sqrt{1+L^2}} = \frac{1}{2}\,\cosh^{-1}\frac{1 - L^2 + 2\,\tan^2\tau}{1+L^2} \tag{3.40}$$

at which point $v = 0$, $r = r_s = \tan\tau$, and

$$\dot{v} = \cos\tau\,\sqrt{1 - L^2\,\cot^2\tau}$$

The next step is to extend into the BTZ part of the spacetime. Note that the s in what follows is not the same parameter, but differs by some shift. This only matters for measuring length, where the two parts need to be added separately.

It turns out that it's very convenient to parametrize the radius of the BTZ horizon as $r_+ = \sec\mu + \tan\mu$, with $-\frac{\pi}{2} < \mu < \frac{\pi}{2}$. With this, we get the energy outside the shell as

$$E = -\frac{\cos\tau\,\sqrt{1 - L^2\,\cot^2\tau}}{1 - \sin\mu}. \tag{3.41}$$

A useful piece of information is also the value of $\dot{r}$ after shell crossing, which is

$$\dot{r}(r_s) = \frac{\sqrt{1 - L^2\,\cot^2\tau}}{\cos\tau}\,\frac{(\sin^2\tau - \sin\mu)}{(1 - \sin\mu)} \tag{3.42}$$

3.A.1 Symmetric radial geodesics

We now specialize further to consider just the radial geodesics, with $L = 0$. Outside the shell, the radial equation of motion can be obtained from the energy, and is

$$\dot{r}^2 = r^2 + \frac{\sin^2\mu - \sin^2\tau}{(1-\sin\mu)^2}$$

Combining this with the value of $\dot{r}$ after the shell, we find that the boundary is reached if and only if $\tau > \mu$, as argued in the text. In particular, if $\mu < 0$ (corresponding to $r_+ < 1$), the geodesics will reach the boundary for all positive τ.

We must now split into two cases, depending on whether τ is greater than or less than $|\mu|$.

$\tau > |\mu|$: We begin with the case of earlier geodesics, relevant for any size of black hole. The simpler parts of the solution to obtain are

$$r = \frac{\sqrt{\sin^2\tau - \sin^2\mu}}{1-\sin\mu}\cosh s$$
$$\dot{v} = \frac{1-\sin\mu}{\sqrt{\sin^2\tau - \sin^2\mu}\sinh s + \cos\tau}$$

and at the shell, we have

$$s = \log\left[\frac{1+\sin\tau}{\cos\tau}\sqrt{\frac{\sin\tau - \sin\mu}{\sin\tau + \sin\mu}}\right]. \tag{3.43}$$

Note that this can be positive or negative, depending on whether the geodesic is going inwards or outwards after crossing the shell.

By integrating $\dot{v}$ from this value to ∞, using the substitution

$$x = \frac{\sqrt{\sin^2\tau - \sin^2\mu}\, e^s + \cos\tau}{\cos\mu}, \tag{3.44}$$

we eventually get the time at which the boundary is reached:

$$t_\infty = \frac{1-\sin\mu}{\cos\mu}\log\left[\frac{\cos\left(\frac{\tau+\mu}{2}\right)}{\sin\left(\frac{\tau-\mu}{2}\right)}\right], \tag{3.45}$$

or back in terms of r_+,

$$t_\infty = \frac{1}{r_+}\log\left[\frac{\sec\tau + \tan\tau + r_+}{\sec\tau + \tan\tau - r_+}\right]. \tag{3.46}$$

Finally, we extract the length from this, by the value of s at the (large) cutoff $r = R$,

minus the value of s at the shell, plus the length from the pure AdS region:

$$\begin{aligned}\ell &= 2\log\left[\frac{2(1-\sin\mu)R}{\sqrt{\sin^2\tau-\sin^2\mu}}\right] - 2\log\left[\frac{1+\sin\tau}{\cos\tau}\sqrt{\frac{\sin\tau-\sin\mu}{\sin\tau+\sin\mu}}\right] \\ &\quad + 2\log(\sec\tau+\tan\tau) \\ &= 2\log\left(\frac{1-\sin\mu}{\sin\tau-\sin\mu}\right) + 2\log(2R),\end{aligned}$$

where the second term in the last line is the result in vacuum.

We can eliminate τ between these, to get

$$\ell = 2\log\left(\cosh^2\frac{r_+t}{2} + \frac{1}{r_+^2}\sinh^2\frac{r_+t}{2}\right) + 2\log(2R) \tag{3.47}$$

for $t > 0$.

$\tau < |\mu|$: The second case is only relevant for small black holes ($\mu < 0$). The solution is

$$\begin{aligned}r &= \frac{\sqrt{\sin^2\mu-\sin^2\tau}}{1-\sin\mu}\sinh s \\ \dot{v} &= \frac{1-\sin\mu}{\sqrt{\sin^2\mu-\sin^2\tau}\cosh s+\cos\tau}\end{aligned}$$

and at the shell, the parameter is

$$s = \log\left[\frac{1+\sin\tau}{\cos\tau}\sqrt{\frac{\sin\mu-\sin\tau}{\sin\mu+\sin\tau}}\right]. \tag{3.48}$$

The time at which the boundary is reached, as well as the length, are computed in a similar way to the first case, and the resulting expressions are identical.

An alternative way to reach the same results is to calculate directly in the Penrose coordinates, which reduces to the simpler computation in pure AdS. The only extra work required is to check that the geodesics remain away from the singularity, and to work out how the coordinates transform at the boundary, to obtain t_∞ and to make the correct regularization of the lengths.

3.A.2 Region covered by geodesics

We here flesh out the arguments in the text describing the regions of spacetime accessible to geodesics of various classes.

The radial motion in the BTZ part of the geometry is described by

$$\dot{r}^2 = r^2 - L^2 - \frac{\cos^2\mu\left(1-\frac{L^2}{r^2}\right) - \cos^2\tau\,(1-L^2\,\cot^2\tau)}{(1-\sin\mu)^2}, \tag{3.49}$$

and $\dot{r}^2$ has a minimum at $r = r_0 = \sqrt{L\, r_+}$. This minimum value $\dot{r}^2_{\rm min} \equiv \dot{r}^2(r_0)$ is given by

$$\dot{r}^2_{\rm min} = \frac{\cos^2\tau\,(1 - L^2\,\cot^2\tau) - (\cos\mu - L\,(1-\sin\mu))^2}{(1-\sin\mu)^2} \tag{3.50}$$

which can take either sign, depending on the specific values of the parameters.

The fate of a given geodesic depends on the interplay of r_+, L and r_s. In particular, it can make it out towards the boundary only if either $\dot{r}^2_{\rm min} > 0$ and $\dot{r}(r_s) > 0$, or if $\dot{r}^2_{\rm min} \leq 0$ but $r_s > r_0$. The geodesic can remain for an arbitrarily long time Δv in the vicinity of r_0, which happens in the former case as $\dot{r}^2_{\rm min} \to 0^+$, or in the latter case when $\dot{r}(r_s) < 0$ and $\dot{r}^2_{\rm min} \to 0^-$.

In either case, the desired fine-tuning is one which makes the magnitude of $\dot{r}^2_{\rm min}$ very small. For a fixed black hole size μ, the condition $\dot{r}^2_{\rm min} = 0$ specifies a curve in the $\tau - L$ plane. Solving for $L = L_0(\tau)$, we obtain

$$L_0(\tau;\mu) = \frac{\pm\cot\tau(\sin^2\tau - \sin\mu) - \cos\mu(1-\sin\mu)}{(1-\sin\mu)^2 + \cos^2\tau\,\cot^2\tau}. \tag{3.51}$$

In particular, asymptotically as $\tau \to 0$, we find that $L_0 \sim \pm(\sin\mu)\,\tau$.

This motivates us to look at the family of geodesics with $L = -(\sin\mu)\tau$ for small black holes. We expand the relevant quantities for small τ, and conclude that the whole spacetime is accessed by this family of geodesics as described in the text.

3.B Details of extremal surface computations

Extremal surfaces with the appropriate symmetries, parametrized by generic parameter s, are found from the Lagrangian

$$\mathcal{L} = (r\sin\theta)^{d-2}\sqrt{-f(r,v)\dot{v}^2 + 2\dot{r}\dot{v} + r^2}. \tag{3.52}$$

The equations of motion, at a point where $\dot{v} = 0$, give $\ddot{v} = (d-1)r\dot{\theta}^2 > 0$, which implies that any critical point for v must be a minimum. This, along with the fact that v must be increasing as the boundary is approached, tells us that v has exactly one local minimum. Furthermore, the symmetries imply that this must occur when the surface crosses the pole of the sphere at $\theta = 0$. This makes v an appropriate candidate for a parameter along the surface.

Denoting differentiation with respect to v by primes, the Lagrangian with the parameter v is

$$\mathcal{L} = (r\sin\theta)^{d-2}\sqrt{-f(r,v) + 2r' + r^2\theta'^2}, \tag{3.53}$$

which gives the equations of motion

$$r'' = \frac{3r' - f}{2}\frac{df}{dr} + (f - r')\left((d-1)r\theta'^2 + \frac{d-2}{r}(-f + 2r')\right) \tag{3.54}$$

$$\theta'' = \left(\frac{1}{2}\frac{df}{dr} - \frac{f}{r}\right)\theta' - \left(\frac{d-1}{r}\theta' - \frac{d-2}{r^2}\cot\theta\right)(-f + 2r' + r^2\theta'^2). \tag{3.55}$$

Initial conditions are chosen as the values of v and r as the pole $\theta = 0$ is crossed. Since this is a singular point of the equations, for numerics we must start integration slightly away from this point, picking initial conditions by a series solution:

$$\theta(v_0 + h) = \sqrt{\frac{2h}{r_0}}\,(1 + O(h))\,, \quad r(v_0 + h) = r_0 + f(r_0, v_0)h + O(h^2). \tag{3.56}$$

This is altered in the special case where the surface is equatorial; this means that it passes through $r = 0$ and has $\theta = \frac{\pi}{2}$. The solution will always be the same close to the origin, namely lying on a static slice of pure AdS until it meets the shell, which, and this is how initial conditions are specified.

The areas are found by numerically integrating the difference between the Lagrangian and

$$\frac{(r\sin\theta)^{d-2}}{\sqrt{1 + r^2\sin^2\theta}}(r'\sin\theta + r\theta'\cos\theta), \tag{3.57}$$

which is the derivative of the function which gives the area of a minimal surface in AdS passing through the point r, θ. This automatically regularizes the area by subtracting off the background vacuum value.

Chapter 4

Gravity in three dimensions

This chapter is a reproduction of the paper *Entanglement entropy in three dimensional gravity* [3].

This paper looks at holographic entanglement entropy in pure gravity in three dimensions, which is a major simplification, because there are no bulk dynamics. In fact, the theory is topological, being equivalent to an $\mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})$ Chern-Simons theory. This simplicity is related to the structure of conformal field theories in two dimensions, where the conformal symmetry is enhanced to the infinite-dimensional Virasoro algebra. As a result, this is an excellent testing ground for new ideas, and in many cases analytic techniques are available to make calculations tractable.

Since there are no local degrees of freedom, the Einstein equations imply that the geometry is locally isometric to AdS_3. Despite, this, there are still nontrivial geometries, which can be formed by taking a quotient of part of AdS_3. The paper shows how the regularised lengths of geodesics, relevant to the RT and HRT calculations, can be computed in a very simple, purely algebraic way in such quotients. This was applied to a few example geometries.

In time-reflection symmetric geometries, the length computation can be analytically continued in a natural way from Euclidean signature, where proposed derivations of the holographic entanglement entropy result apply, to Lorentzian signature, where time evolution of the geometry is implemented. I describe how such a continuation can be used, in this limited context, to justify the HRT formula, for which there is no more general derivation.

Finally, I also discuss an issue which arises if one takes the Euclidean quantum gravity calculation seriously. From this point of view, a minimal surface corresponding to a saddle point computing entanglement entropy need not lie on the time-reflection invariant slice of the Euclidean geometry. This implies that it is not given by the area of any surface in the Lorentzian spacetime. I give one example in three dimensions, using the Euclidean version of the length calculation, where the relevant geodesics are subdominant, though this requires knowledge of a Hawking-Page phase transition between bulk geometries.

4.1 Introduction

A central feature of quantum systems, distinguishing them from classical ones, is entanglement. In holographic theories, there is an intimate relationship between this quantum mechanical property and geometry. One concrete realisation of this relationship is embodied in the Ryu-Takayanagi (RT) prescription [30, 31] and its covariant Hubeny-Rangamani-Takayanagi (HRT) generalisation [39], which equates the entanglement entropy of a spatial region in the field theory to the area of an extremal surface in the bulk geometry. The entanglement entropy is a natural and useful quantity in field theory, but is notoriously difficult to compute in most circumstances. The gravitational calculation is in many cases much more tractable, so offers a practical way to study the properties and dynamics of entanglement.

Three dimensional gravity has no local degrees of freedom, which makes it a particularly useful setting in which to obtain analytic results. One way of understanding this is that the Ricci tensor specifies the Riemann tensor completely, a consequence of which is that solutions of pure gravity with negative cosmological constant are necessarily isometric to pure AdS_3 locally. Despite this, there is still a rich set of solutions that differ from AdS globally, obtained by taking a quotient of a subset of AdS_3 by some group of isometries. The simplest such solution is the BTZ black hole [122, 123], later generalised to 'wormholes' with any number of asymptotic boundaries, joined through regions with possibly nontrivial topology [124, 125, 126], and possibly rotating [127]. Further generalisations include nonorientable manifolds, see [128] for example. In many cases, there is a known candidate to construct the CFT state dual to these geometries [129, 130], generalising the interpretation of an eternal black hole as a thermofield double state [131].

A priori, the calculation of entanglement entropies, even with holography and with the simplifications of three dimensions, appears challenging in these geometries. It requires finding geodesics, imposing an appropriate infrared cutoff and then integrating their lengths, possibly in a space with complicated topology and no symmetry, and requiring several coordinate patches. The main result of this paper is to show that this apparent difficulty can be circumvented, and in fact the geodesic lengths can be found purely from algebra. With a description of a spacetime as a quotient in hand, the lengths are computed simply from the trace of a product of 2×2 matrices, representing points on the boundary and elements of the quotient group.

From a computational point of view, the advantages of this approach are obvious: many lengths can be calculated very quickly and algorithmically, allowing for very detailed investigations of entanglement entropy in these geometries. We demonstrate this utility in section 4.3, by computations in several examples. We begin for illustration by reproducing known results in rotating BTZ, a single line computation once the black hole has been written as a quotient. We then move on to the $\mathbb{RP}^2$ geon to demonstrate the generalisation to nonorientable spacetimes, and describe the structure of entanglement in the associated pure CFT state. The final two examples are more complicated wormholes, the first with

three asymptotic regions and the second a black hole with a single exterior but a torus behind the horizon. Entanglement in multi-boundary wormhole geometries has been considered already in [132], but only between entire boundaries; the techniques here allow a relatively straightforward extension of this to subintervals.

From a more conceptual standpoint, the description of entanglement entropies obtained may be useful for understanding the status and origin of the RT and HRT prescriptions. In section 4.4, we discuss one such use, namely to show how HRT may follow from an analytic continuation of a Euclidean quantum gravity computation [36] in some circumstances. This also throws up a related issue, that if geodesics computing entanglement entropy are to be understood as coming from such a Euclidean path integral, there are circumstances where there may be no geometric realisation of a relevant Euclidean geodesic in the Lorentzian spacetime. This leads to a very precise and specific notion of complexified (from the Lorentzian standpoint) entangling surface. We discuss how this may occur, and use the Euclidean version of the technology developed earlier to show in a simple example that the complex geodesics we describe are subdominant.

We conclude with a discussion of the results, and indicate some possible directions of future study. These include studies of entanglement in a variety of interesting geometries, as well as some avenues to better understand the underlying origin and properties of RT and HRT.

4.2 Lengths of geodesics in quotients of AdS_3

This section constructs the procedure for calculating entanglement entropy in constant negative curvature geometries. The approach we use relies on a description of AdS_3 as the $SL(2,\mathbb{R})$ group manifold, in which the isometries and boundary have a particularly nice algebraic structure, lending itself to construction of quotients and description of geodesics. We begin by reviewing this presentation of AdS_3, before describing the natural representation of the boundary in this picture. We then look at geodesics, and define a useful notion of regularised length of boundary-to-boundary geodesics. We then put this to work in the context of quotient spacetimes, which leads us to the simple algebraic computation of entanglement entropies.

There is also an analogous construction in the Euclidean setting of quotients of H^3, which we briefly describe in section 4.2.5. We then review considerations relating Lorentzian and Euclidean computations.

4.2.1 The structure of AdS_3

The AdS_3 spacetime can be described conveniently as the group manifold of $SL(2,\mathbb{R})$. Concretely, write $\mathbb{R}^{2,2}$ as

$$\mathbb{R}^{2,2} = \left\{ p = \begin{pmatrix} U+X & Y-V \\ Y+V & U-X \end{pmatrix} \right\}, \quad ds^2 = -L_{AdS}^2 \det(dp), \tag{4.1}$$

and then AdS_3 is the embedded submanifold given by the one sheeted hyperboloid $\det(p) = 1$. Herein, we will choose units to set $L_{AdS} = 1$. Parametrising the hyperboloid by coordinates (t, r, ϕ), we recover the more familiar metric

$$\left.\begin{array}{ll} X = r\cos\phi, & U = \sqrt{1+r^2}\cos t \\ Y = r\sin\phi, & V = \sqrt{1+r^2}\sin t \end{array}\right\} \Rightarrow ds^2 = -(1+r^2)dt^2 + \frac{dr^2}{1+r^2} + r^2 d\phi^2 \tag{4.2}$$

with both ϕ and t periodic[1] in 2π.

The connected part of the isometry group is $SO(2,2) \cong SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\mathbb{Z}_2$, which acts as $p \mapsto g_L p g_R^t$, for g_L and g_R in $SL(2,\mathbb{R})$, with the equivalence $(g_L, g_R) \sim (-g_L, -g_R)$. In addition to this there are two distinct discrete $\mathbb{Z}_2$ symmetries, a spatial reflection and a time reversal, the latter acting by transposition $p \mapsto p^t$.

Of particular importance is the diagonal subgroup $g_L = g_R$. This action fixes symmetric matrices, of which there are two disconnected components, being the static slices $t = 0, \pi$ (distinguished by the sign of the trace). This diagonal $PSL(2,\mathbb{R})$ constitutes the set of isometries commuting with time reversal. We may identify the $t = 0$ slice with the upper half plane model of the hyperbolic plane, by

$$p = \frac{1}{\Im(z)} \begin{pmatrix} |z|^2 & \Re(z) \\ \Re(z) & 1 \end{pmatrix}, \tag{4.3}$$

where z is in the upper half plane and $\Re(z), \Im(z) > 0$ denote its real and imaginary parts. Then these isometries act on the half plane by fractional linear transformations $z \mapsto \frac{az+b}{cz+d}$ with real coefficients.

The Penrose diagram in section 4.2.1 is useful to describe both the spacetime itself, as well as the elements of $SL(2,\mathbb{R})$ making up the isometries. These are divided into hyperbolic, parabolic and elliptic, depending on the trace, as described in the caption of the figure.

There is also a very natural description of the asymptotic boundary in this picture, represented by the entries of the matrix becoming large. By pulling out a factor which blows up, chosen to leave a piece with a finite limit, the points approached on the boundary can be represented by the space of singular real 2×2 matrices modulo overall rescaling with an arbitrary positive constant. Any such matrix can alternatively be written as $p = \vec{v}\vec{u}^t$ for some nonzero two-dimensional vectors $\vec{u}, \vec{v}$. One choice of gauge fixes the overall scaling by choosing the vectors $\vec{v}$ and $\vec{u}$ to be unit, which leaves only a $\mathbb{Z}_2$ ambiguity of changing the signs of both vectors. This makes the boundary space a torus $S^1 \times S^1/\mathbb{Z}_2$, and the $t = 0$ slice is given by the circle $\vec{u} = \vec{v}$.

In terms of the coordinates, this corresponds to taking $r \to \infty$ at fixed t, ϕ. With null

[1] Usually, it is necessary to take the universal cover by unwrapping the time circle. We won't need to do this, since the quotient construction (save for the case of rotating BTZ) will remove the closed timelike curves.

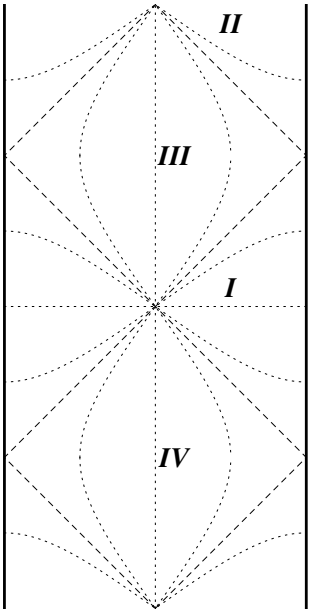


Figure 4.1: The Penrose diagram of AdS_3, or equivalently $SL(2,\mathbb{R})$, formed from the rotation around the central vertical line, and identifying top ($t = \pi$) and bottom ($t = -\pi$). The boundary is at the bold lines to the left and right. The identity element is in the centre. The hyperbolic transformations are in regions ***I*** ($\mathrm{Tr} > 2$) and ***II*** ($\mathrm{Tr} < -2$), the elliptic transformations in regions ***III*** and ***IV*** ($-2 < \mathrm{Tr} < 2$), and the parabolic transformations ($\mathrm{Tr} = \pm 2$) on the lightcones between them. For the diagonal subgroup, hyperbolic, parabolic and elliptic classes are distinguished by having two, one or no fixed points on the boundary of the $t = 0$ Poincaré disc respectively. Elliptic isometries also have a single fixed point in the interior of the disc.
Dotted lines show geodesics.

coordinates $t = v + u$, $\phi = v - u$, the resulting matrix is

$$p \propto \begin{pmatrix} \cos v \\ \sin v \end{pmatrix} \begin{pmatrix} \cos u & -\sin u \end{pmatrix} \tag{4.4}$$

and the 2π periodicities of u, v, along with the $\mathbb{Z}_2$ of flipping signs $(u, v) \sim (u + \pi, v + \pi)$, reproduces the 2π periodicities of t, ϕ. From this form it is clear how the $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\mathbb{Z}_2$ acts on the boundary, with each $SL(2,\mathbb{R})$ factor acting separately as projective transformations on left- or right-moving null coordinates. Fixed points on the boundary of an isometry (g_L, g_R) are points where $\vec{u}$ and $\vec{v}$ are eigenvectors of g_L and g_R respectively, with eigenvalues of the same sign.

The choice of positive scaling coincides precisely with a conformal factor in the boundary metric, with the unit vector choice above corresponding to the usual flat metric on the cylinder. We will review this in more detail after discussing the geodesic lengths, to which we now turn.

4.2.2 Lengths of geodesics

One advantage of this group manifold description is that the length of a geodesic between two points has a simple form. There are three cases, understood easily from the Penrose diagram section 4.2.1. We are mostly interested in spacelike separated points $p, q \in SL(2,\mathbb{R}) \equiv AdS_3$, which means $\mathrm{Tr}(p^{-1}q) > 2$, so if p is at the identity, q is in region ***I***. There is a unique spacelike geodesic between two such points, with length

$$\ell(p,q) = \frac{1}{2}\cosh^{-1}\left(\frac{\mathrm{Tr}(p^{-1}q)^2}{2} - 1\right) = \cosh^{-1}\left(\frac{\mathrm{Tr}(p^{-1}q)}{2}\right). \tag{4.5}$$

This can be proven by using isometries to first bring p to the identity, and then using the residual $SL(2,\mathbb{R})/\mathbb{Z}_2$ (with $g_R^t = g_L^{-1}$) to bring q to diagonal form so the length calculation is simple. The trace requirement ensures that this is possible, since it implies that q has two distinct real positive eigenvalues e^ℓ and $e^{-\ell}$. Since $\mathrm{Tr}(p^{-1}q)$ is invariant under the action of the isometries, we recover the stated formula in general.

For completeness, we note that the proper time along geodesics connecting timelike separated points p and q with $|\,\mathrm{Tr}(p^{-1}q)| < 2$, (second point in region ***III*** or ***IV***) is

$$\tau(p,q) = \cos^{-1}\left(\frac{\mathrm{Tr}(p^{-1}q)}{2}\right) \tag{4.6}$$

and the choice of the branch of $\cos^{-1}$ allows for cycling round the periodic time direction any number of times. Finally, null separated points have $\mathrm{Tr}(p^{-1}q) = 2$, and if $\mathrm{Tr}(p^{-1}q) \leq -2$ there is no geodesic between p and q (with the second point in region ***II***). The exception is the case $p = -q$ when all timelike geodesics through p also pass through q, after proper time π.

This can be usefully applied to define a regularised length of geodesics between points on the boundary. For any boundary point, we may choose a curve $p(\rho)$ in $SL(2,\mathbb{R})$ such that at least one entry tends to infinity as $\rho \to \infty$, and $p(\rho)/\rho$ has a finite but nonzero limit p_∂. This curve approaching the boundary can of course be chosen in many ways, and the resulting p_∂ will differ by a positive constant of proportionality; the choice will amount to picking a regularisation scheme.

Given two such curves approaching spacelike separated points on the boundary, the distance between the points $p(\rho)$ and $q(\rho)$ is

$$\begin{aligned}
\ell(p(\rho),q(\rho)) &= \cosh^{-1}\left(\frac{\mathrm{Tr}(p(\rho)^{-1}q(\rho))}{2}\right) && (4.7)\\
&= \cosh^{-1}\left(\frac{1}{2}\,\mathrm{Tr}\left(R_\perp p(\rho)^t R_\perp^t q(\rho)\right)\right) && (4.8)\\
&= \cosh^{-1}\left(\frac{\rho^2}{2}\,\mathrm{Tr}\left(R_\perp p_\partial^t R_\perp^t q_\partial\right) + o(\rho^2)\right) && (4.9)\\
&= \log\rho^2 + \log\left(\mathrm{Tr}\left(R_\perp p_\partial^t R_\perp^t q_\partial\right)\right) + o(1) \quad \text{as } \rho\to\infty && (4.10)
\end{aligned}$$

where $R_\perp = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ is the matrix rotating by angle $\pi/2$, used to implement the inverse.

Motivated by this, we define a regularised length between the boundary points by subtracting off the divergent piece:

$$\ell_{\mathrm{reg}}(p_\partial, q_\partial) = \log\left(\mathrm{Tr}\left(R_\perp p_\partial^t R_\perp^t q_\partial\right)\right). \tag{4.11}$$

Notice that this depends on the choice of singular matrix to define each point, but the boundary is defined as the set of such matrices modulo positive rescalings; this is the dependence on the choice of regularisation.

If we choose to express the boundary points as $p_\partial = \vec{u_1}\vec{v_1}^{\,t}$, and $q_\partial = \vec{u_2}\vec{v_2}^{\,t}$, then this

can also be expressed very simply as

$$\ell_{\text{reg}}(p_\partial, q_\partial) = \log\left((\vec{u}_1^{\perp} \cdot \vec{u}_2)(\vec{v}_1^{\perp} \cdot \vec{v}_2)\right), \tag{4.12}$$

where $\vec{u}^{\perp} = R_{\perp}\vec{u}$ represents the rotation of $\vec{u}$ by $\pi/2$. This could alternatively be expressed with two-dimensional cross products of $\vec{u}$s and $\vec{v}$s, or determinants.

The choice of regularisation we describe here maps very simply to the usual IR cutoff prescription, where the lengths are measured up to some fixed radius in a Fefferman-Graham expansion. This expansion depends on the choice of boundary metric, which is fixed simply by the conformal factor relating it to the flat metric on the cylinder. Indeed, one may choose ρ to be the radial Fefferman-Graham coordinate with the preferred choice of boundary metric, in which case the prescription is manifestly equivalent to the usual procedure of regulating at constant ρ and dropping the cutoff dependent piece.

The flat metric comes from choosing $\vec{u}$ and $\vec{v}$ to be unit vectors as described above. To illustrate, on the static $t = 0$ slice choosing $\vec{u} = \vec{v} = (\cos(\phi/2), \sin(\phi/2))^t$ gives $\ell_{\text{reg}} = \log\sin^2(\Delta\phi/2)$, which, up to a factor of $\frac{1}{4G_N} = \frac{c}{6}$ and the constant non-universal piece, reproduces the universal 2D CFT entanglement entropy on a circle in vacuum as expected.

When we move to discussing quotients, it will be necessary to make another choice of boundary metric, in particular one that is invariant under the action of the quotient group. If this is related to the flat metric as

$$ds^2 = \Omega^2 ds^2_{\text{flat}} \tag{4.13}$$

then the representative matrix on the boundary is just a factor of Ω larger than one built from unit vectors.

As a simple example of an alternative regularization on the $t = 0$ slice, natural in Poincaré patch or planar coordinates, choose $\vec{u} = \vec{v} = (x, 1)^t$. This differs from the vacuum normalization by a conformal factor of $\Omega = (1 + x^2)$, which is just right to give the metric of Minkowski space when extended to the diamond $-\pi < t \pm \phi < \pi$. The regularised length gives the entanglement entropy result $\frac{c}{3}\log\Delta x$, being the universal CFT result on a line in vacuum.

4.2.3 Quotients

Any constant negative curvature geometry can be obtained from a quotient of some subset of AdS_3 by a discrete group of isometries[2]. The technology we develop for computing regularised lengths of geodesics relies of having such a description for the spacetime of interest. We only briefly outline the aspects of the construction we need; for the full details see [127] for the general case, and [124, 125, 126] for more on the special non-rotating case. An extension to allow orientation reversing isometries is relatively straightforward; we will

[2]Some, such as planar BTZ, must be obtained as a limit of such quotient geometries.

not describe the details here, though an example will be given in section section 4.3.2.

Suppose we have a discrete group $\Gamma \subseteq SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\mathbb{Z}_2$ of hyperbolic isometries. The hyperbolic condition here means that the trace of both $SL(2,\mathbb{R})$ elements is greater than two, and ensures that the spacetime is free of conical singularities, closed timelike curves, or other singular features[3]. We can form a spacetime as the quotient $\widehat{AdS}/\Gamma$ of a particular covering space $\widehat{AdS}$ by the isometries. The covering space is formed from AdS_3 by removing all points to the future and past of any fixed point of an isometry lying on the boundary. In particular, the covering space on the boundary will be a (generically infinite) collection of diamonds with corners at the fixed points of some $\gamma \in \Gamma$. The resulting bulk covering space is simply connected, so the fundamental group of the resulting geometry can be identified with Γ.

It is easiest to understand the construction in the special 'non-rotating' case of Γ being a subgroup of the diagonal $PSL(2,\mathbb{R})$, acting on the $t = 0$ slice, modelled as the upper half plane, by fractional linear transformations. The quotient of the upper half plane obtained, which is a time-reflection invariant slice of the quotient geometry, will be a surface with some number of boundaries and handles. This can be thought of as any number of black hole exteriors joined behind the horizon with an arbitrary topology. See sections 4.3.3 and 4.3.4 for examples.

A geodesic in the quotient is represented by a set of geodesics in the covering space, related by elements of Γ. Between any two spacelike separated boundary points p and q in $\widehat{AdS}/\Gamma$, there will be many geodesics, in one-to-one correspondence with elements of Γ. To see this, pick some representatives $\hat{p}$ of p and $\hat{q}$ of q in covering space. Each geodesic in the quotient will have a unique lift ending on $\hat{p}$. The other end of this lifted geodesic may be at any of the points $\gamma\hat{q}$, with $\gamma \in \Gamma$, and these are all distinct since fixed points have been removed from the covering space. Topologically, this correspondence between geodesics and elements of Γ results from the fact that each homotopy class of curves has a unique geodesic representative.

There is a subtlety here, that the lifted geodesic could leave the covering space $\widehat{AdS}$, and hence not be an admissible geodesic in the quotient, but we show that this is impossible. This relies on a property of any spacelike geodesic in AdS. Begin by constructing the pair of diamonds of boundary points spacelike separated from the endpoints of the geodesic. Now take the 'causal wedges' (see [54]) in the bulk formed by the union of causal curves beginning and ending in one of these diamonds. The causal wedges are bounded by null surfaces to the past and future, and these intersect precisely at the geodesic. The consequence of this is that the past of the geodesic on the boundary is exactly the past of its endpoints, being the past of the pair of diamonds. Another way of stating this is that the geodesics satisfy causality in the sense of [41], but only marginally. Now if there is a fixed point of Γ on the boundary in the past of the geodesic, then it is also in the past of at least one of the endpoints. But in that case, the endpoint is not in the covering space, and

[3]Some other special classes of isometries including parabolic elements are important, for example for the extremal BTZ black hole (see [123, 127]), but only minor details change in what follows.

does not represent a point on the boundary of the quotient geometry. Identical remarks apply with 'past' replaced by 'future', so no point of the geodesic leaves $\widehat{AdS}$.

Now, we wish to compare the lengths of the various geodesics between two boundary points in the quotient. The lift of these geodesics to covering space will connect different endpoints, related by elements of Γ. To make sure we are regulating consistently, we must ensure that the points at which we impose the infrared cutoff in different regions are images of one another under the appropriate element of Γ. The result is that the singular matrices representing points on the boundary may not be chosen independently, but must be invariant under the action of Γ: we may only choose a regularisation for a single lift of each point, and this will determine all its images' regularisations. This choice of singular matrices can be considered as a choice of boundary metric, which must be invariant under the action of Γ on the boundary.

So for each pair of points in $\partial\widehat{AdS}/\Gamma$, having chosen singular matrices p, q to represent a lift of each, the consistent regularised lengths of geodesics between the points are given by

$$\ell_{\text{reg}}(p,q,\gamma) = \log\left(\text{Tr}\left(R_\perp p^t R_\perp^t g_L q g_R^t\right)\right) = \log\left((\vec{u}_1^\perp \cdot g_L \vec{u}_2)(\vec{v}_1^\perp \cdot g_R \vec{v}_2)\right) \tag{4.14}$$

where $\gamma = (g_L, g_R) \in \Gamma \subseteq SL(2,\mathbb{R}) \times SL(2,\mathbb{R})/\mathbb{Z}_2$ and $p = \vec{u_1}\vec{v_1}^t$, $q = \vec{u_2}\vec{v_2}^t$.

In any quotient, the asymptotic regions approaching each boundary component are isometric to the external region of the BTZ black hole up to the horizon. Hence, a natural way to choose the regularisation is to reproduce the BTZ result for lengths of geodesics remaining outside the horizons, picking flat cylinder metrics on each boundary component. Locally, near each boundary, there will be two Killing vectors generating translations in space and time (though not extendible consistently into the whole interior). With this regularisation choice, matrices representing different points in the same asymptotic boundary will be related by the flow of these local symmetries. This leaves only one overall constant to be determined, fixed by matching to some standard normalisation for lengths between points at small separation, say $S \sim \frac{c}{3}\log l$ for small intervals of length l.

For computing entanglement entropy holographically, apart from geodesics ending on the boundary it is also required to be able to compute lengths of closed spacelike geodesics. A closed geodesic in the quotient again lifts to a set of geodesics in the covering space $\widehat{AdS}$, each lift connecting a pair of fixed points on the boundary of some $\gamma \in \Gamma$. This γ is the element which takes any point on the curve in covering space to a corresponding point obtained by following the geodesic for a single cycle. A different lift of the closed geodesic will connect fixed points of some γ' conjugate to γ in Γ. Hence, the closed geodesics are in one to one correspondence with conjugacy classes of Γ. From a topological point of view, this says that each free homotopy class of loops[4] has exactly one geodesic representative.

This also gives us a simple way to compute the length of the closed geodesic. We can

[4]The standard definition of homotopies in the definition of the fundamental group requires beginning and ending loops on some fixed basepoint, necessary to give π_1 its group structure. Free homotopy here means homotopies of loops without the requirement to fix the basepoint. The equivalence classes under this larger set of homotopies are conjugacy classes of the fundamental group, conjugation corresponding to taking different paths linking a loop back to the basepoint.

use the isometries of AdS to send the fixed points of γ to $t = \phi = 0$ and $t = 0, \phi = \pi$, and hence the geodesic to the straight line across the $t = 0$ slice (being the diagonal matrices). It is then straightforward to compute

$$\ell = \cosh^{-1}\left(\frac{\operatorname{Tr} g_L}{2}\right) + \cosh^{-1}\left(\frac{\operatorname{Tr} g_R}{2}\right), \tag{4.15}$$

where $\gamma = (g_L, g_R)$ for g_L and g_R in $SL(2, \mathbb{R})$, with traces greater than 2 as required for the transformation to be hyperbolic.

4.2.4 Homology

In this description it is very easy to compute the homology classes of curves in order to algorithmically decide which combinations of open and closed geodesics are topologically admissible. We will take the criterion to be that the geodesics in the bulk, together with the intervals on the boundary whose entanglement we compute, when given appropriate orientations, should have trivial homology[5]. In other terms, there must be a codimension-one oriented submanifold whose boundary is the interval along with the geodesics. This is essentially the prescription used to prove that known quantum mechanical properties of entanglement entropy, such as strong subadditivity, follow from the holographic prescription [32, 40]. In the case of time-reflection symmetric geometries at the moment of time symmetry, where a Euclidean quantum gravity approach is available, this form of the homology constraint can be proven [4] under certain assumptions about the holographic replica trick [36].

As already commented upon, the fundamental group of the quotient space is isomorphic to the quotient group Γ itself. The first homology group H_1 is the abelianisation of the fundamental group, so homology is obtained from homotopy simply by counting the number of each generator contained in any element $\gamma \in \Gamma$.

For closed geodesics, this is entirely straightforward: closed curves are associated with conjugacy classes of Γ, any representative of a conjugacy class will have the same nett number of each generator, and so any representative will do to compute the homology.

For open geodesics, it is equally straightforward if the representatives of boundary points are chosen appropriately. Any interval on the boundary of the quotient lifts to a collection of intervals on the boundary of covering space, and if the representatives of the endpoints are chosen from ends of the same component of the lifted interval, then the element γ associated with a geodesic will correspond to the homology class of the geodesic along with the interval. This means that the interval will be homotopic to the geodesic associated with the identity in Γ. This will happen automatically if the representatives are obtained by translating with local Killing vectors in each component.

If there are geodesics between different asymptotic boundaries, and the endpoints of the intervals are chosen as above, this can be straightforwardly extended by combining the

[5]This is in fact not quite the right thing if the spacetime is nonorientable. The required modification will be discussed in section 4.3.2.

abelianisations of all elements of γ from the geodesics, taking care with signs[6].

4.2.5 The story in H^3

There is an analogous picture in the Euclidean hyperbolic 3-space, by taking Hermitian (instead of real) 2×2 matrices of unit determinant. In fact, this gives two disconnected copies of H^3, which can be distinguished as positive and negative trace components.

$$H^3 = \{M \in SL(2,\mathbb{C}) | M = M^\dagger, \operatorname{Tr} M > 0\} \tag{4.16}$$

Explicitly, this can be seen as the top sheet of the two-sheeted hyperboloid embedded in $\mathbb{R}^{3,1}$ written as

$$\mathbb{R}^{3,1} = \left\{ p = \begin{pmatrix} T+Z & X-iY \\ X+iY & T-Z \end{pmatrix} \right\}, \quad ds^2 = -\det(dp), \tag{4.17}$$

given by $\det(p) = 1$. The normal ball model coordinates are

$$\left.\begin{array}{c} (X,Y,Z) = \frac{2}{1-r^2}(x,y,z) \\ T = \frac{1+r^2}{1-r^2} \end{array}\right\} \Rightarrow ds^2 = \frac{4(dx^2+dy^2+dz^2)}{(1-r^2)^2} \tag{4.18}$$

where $r^2 = x^2 + y^2 + z^2 < 1$.

Now the connected part of the isometry group is given by $SO(2,1) \cong PSL(2,\mathbb{C})$, acting by $p \mapsto gpg^\dagger$. There is one orientation reversing isometry, acting by transposition. This fixes an equatorial copy of the hyperbolic plane at $Y = 0$, where the matrices are real symmetric, matching the static $t = 0$ slice of AdS_3 precisely, and it is fixed by the obvious $PSL(2,\mathbb{R}) \subseteq PSL(2,\mathbb{C})$. This matching between the diagonal symmetry group in the Lorentzian case and real part of the symmetry group in the Euclidean case, along with identification of the static slice, will play an important rôle in certain contexts, as reviewed in section 4.2.6.

Once again, the boundary is approached when any entry of the matrix becomes large, so is represented by singular 2×2 Hermitian matrices of positive trace, modulo real rescaling. These can be written as $\vec{u}\vec{u}^\dagger$ for some $\vec{u} \in \mathbb{C}^2 - \{0\}$, with an equivalence under scaling $\vec{u}$ by a nonzero complex number. In other words, the boundary is just the Riemann sphere $\mathbb{CP}^1$. One particular canonical choice of $\vec{u} = (z, 1)$ can be made, for $z \in \mathbb{C} \cup \{\infty\}$, in terms of which the isometry group then simply acts on the boundary by fractional linear transformations in the usual way.

Lengths of geodesics are again given by the formula

$$\ell(p,q) = \cosh^{-1}\left(\frac{\operatorname{Tr}(p^{-1}q)}{2}\right) \tag{4.19}$$

[6] An individual geodesic going between different asymptotic regions does not have a meaningful homology class on its own, since without other bulk geodesics it can never form a closed loop. Only combinations that can be summed to make cycles have any invariant meaning.

and the computation for regularised length between boundary points remains essentially unchanged, giving

$$\ell_{\text{reg}}(p,q) = \log\left(\text{Tr}\left(R_\perp p^t R_\perp^t q\right)\right) \tag{4.20}$$

$$= \log|\vec{u}_\perp \cdot \vec{v}|^2 \tag{4.21}$$

where $p = \vec{u}\vec{u}^\dagger$, $q = \vec{v}\vec{v}^\dagger$ and $\vec{u}_\perp = R_\perp \vec{u}$ as before.

Hyperbolic manifolds can be obtained by taking a quotient H^3/Γ with a discrete subgroup Γ of $PSL(2,\mathbb{C})$, known as a Kleinian group, consisting of loxodromic or parabolic elements (excepting the identity) so that the quotient group acts without fixed points in the bulk. This is simpler than the Lorentzian case, since no points of the bulk need to be removed from the covering space. The boundary of the space will be a quotient by Γ of a set of points on which Γ acts 'nicely', known as the domain of discontinuity (which may be empty, giving a 'closed universe') and will result in a Riemann surface or orbifold.

Two special cases are worth mentioning here. Handlebody geometries, which loosely speaking are obtained from 'filling in' a closed Riemann surface of genus g, are obtained from Schottky groups with g generators, as discussed extensively in [130]. There are more ways to obtain a bulk with a given Riemann surface as its boundary, though there is a conjecture that the handlebodies are the dominant geometries holographically [128]. Secondly, the groups which fix the equatorial slice are discrete subgroups of $PSL(2,\mathbb{R})$, known as Fuchsian groups, and these offer the most obvious mapping between Lorentzian and Euclidean spacetimes. These two classes overlap, so a group may be both Fuchsian and Schottky.

The lengths of geodesics in quotients are found in much the same manner as the Lorentzian case. The relevant formulae for geodesics joining boundary points represented by singular matrices $p = \vec{u}\vec{u}^\dagger, q = \vec{v}\vec{v}^\dagger$ is

$$\ell_{\text{reg}}(p,q,\gamma) = \log\left(\text{Tr}\left(R_\perp p^t R_\perp^t g q g^\dagger\right)\right) = \log|\vec{u}_\perp \cdot g\vec{v}|^2 \tag{4.22}$$

for $g \in \Gamma$. The lengths of closed geodesics are given by

$$l = \cosh^{-1}\left[\left|\frac{Trg}{2}\right|^2 + \left|\left(\frac{Trg}{2}\right)^2 - 1\right|\right] \tag{4.23}$$

where g is a representative of the conjugacy class of $PSL(2,\mathbb{C})$ associated to the curve. Note that in the real $PSL(2,\mathbb{R})$ case, these formulae all match the Lorentzian formulae in the diagonal $PSL(2,\mathbb{R})$ case.

4.2.6 Matching between Euclidean and Lorentzian descriptions

The Euclidean geometries offer a natural way to describe the CFT state dual to a given quotient geometry, at least if the quotient group Γ is in the diagonal $PSL(2,\mathbb{R})$ [129, 130], generalising the duality between the eternal black hole and the thermofield double state

[131].

In this case there is a moment of time-reflection symmetry, which automatically has zero extrinsic curvature, so can also be described as a surface in a Euclidean geometry, which can be obtained by the quotient of H^3 by the same group Γ, now in the real $PSL(2,\mathbb{R})$ subgroup. The state on the $t=0$ slice can now be defined by a Hartle-Hawking procedure as being prepared by a path integral over the bottom half of the Euclidean spacetime.

The interpretation of the CFT state is the boundary analogue of this construction: a path integral over a Riemann surface with boundaries defines a state on those boundaries, as the wave functional evaluated on the field configuration given by the boundary conditions. The partition function is obtained from gluing a pair of these surfaces together at their boundaries, to give the Schottky double.

However, from any given Riemann surface, there will in general be many different bulk geometries with that boundary, and the partition function may be computed from any one of them depending on which is the dominant saddle point with smallest Euclidean action. As the moduli vary, the dominant saddle point will change, and the Euclidean bulk description will change. These transitions are a generalisation of the Hawking-Page phase transition. Assuming that the bulk does not spontaneously break the time reflection symmetry, the appropriate Lorentzian geometry can then be obtained by evolving from the time-reflection invariant slice of the dominant Euclidean saddle point. This may be disconnected, which results in a Lorentzian spacetime consisting of several disjoint components, though they are all connected in the Euclidean section. See also [132] for a more detailed discussion.

4.3 Examples

4.3.1 BTZ

The simplest example of a quotient of AdS is the BTZ black hole [122, 123], generated by a single hyperbolic element γ of $SL(2,\mathbb{R})\times SL(2,\mathbb{R})/\mathbb{Z}_2$. By conjugation, this can be chosen to be the result of exponentiating the $\mathfrak{sl}(2,\mathbb{R})\oplus\mathfrak{sl}(2,\mathbb{R})$ Lie algebra elements (or equivalently Killing vectors)

$$\xi_L = \frac{r_+ - r_-}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \quad \xi_R = \frac{r_+ + r_-}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{4.24}$$

with $r_+ \geq r_- \geq 0$, so $\gamma = (g_L, g_R) = (\exp(2\pi\xi_L), \exp(2\pi\xi_R))$. The cases $r_+ = r_-$ and $r_- = 0$ correspond to extremal and static cases respectively.

This isometry fixes the two points $t=0,\phi=0$ and $t=0,\phi=\pi$ on the boundary, and so the boundary covering space is the pair of diamonds spacelike separated from both these points. As a result, the spacetime has two asymptotic boundaries, with the field theory interpretation of the two noninteracting but entangled halves of a thermofield double state.

This spacetime has two independent Killing vectors, generated by the $\mathfrak{sl}(2,\mathbb{R})$ matrices

commuting with ξ_L and ξ_R, acting on the left and right respectively. These can be chosen as $\xi = (\xi_L, \xi_R)$ itself, generating spatial translations on the boundaries, and the orthogonal combination $\eta = (\xi_L, -\xi_R)$ generating time translations, acting in opposite directions in the two asymptotic regions.

The $SL(2,\mathbb{R})$ description of the spacetime can be related to the more usual coordinates (t, r, ϕ) in one exterior region $r > r_+$ by

$$p = \frac{1}{\sqrt{r_+^2 - r_-^2}} \begin{pmatrix} e^{-r_- t + r_+ \phi}\sqrt{r^2 - r_-^2} & e^{-r_+ t + r_- \phi}\sqrt{r^2 - r_+^2} \\ e^{r_+ t - r_- \phi}\sqrt{r^2 - r_+^2} & e^{r_- t - r_+ \phi}\sqrt{r^2 - r_-^2} \end{pmatrix}. \tag{4.25}$$

Acting with γ simply maps ϕ to $\phi + 2\pi$, so implements the periodicity of the angular direction. The metric is then

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2\left(d\phi - \frac{r_+ r_-}{r^2}dt\right)^2, \quad \text{where } f(r) = \frac{(r^2 - r_+^2)(r^2 - r_-^2)}{r^2}. \tag{4.26}$$

In terms of the inner and outer horizon radii, the physical parameters of the black hole are its mass $M = r_+^2 + r_-^2$ and angular momentum $J = 2r_+ r_-$.

There is one closed geodesic for each element γ^n of Γ, according to the correspondence with conjugacy classes of Γ. These are geodesics wrapping the bifurcation circle n times, and have length $2\pi|n|r_+$, as can be checked from equation (4.15).

The length calculations require choosing a singular matrix, or a pair of vectors, to represent one point on each boundary, normalised in some convenient way. A simple set of choices is at $t = \phi = 0$, with

$$\vec{u}_1 = \vec{v}_1 = (r_+^2 - r_-^2)^{-1/4}\begin{pmatrix}1\\1\end{pmatrix}, \quad \vec{u}_2 = \vec{v}_2 = (r_+^2 - r_-^2)^{-1/4}\begin{pmatrix}-1\\1\end{pmatrix} \tag{4.27}$$

with the subscripts labelling the region. The factor is chosen for a convenient standard normalisation. The regularised lengths are now extremely simple to compute, from equation (4.14), and by translating the points to the required locations with the Killing vectors ξ and η. Since these are Killing vectors, this translation need only be done to one endpoint, the other being placed at $t = \phi = 0$. Furthermore, the inclusion of group elements of Γ labelling different geodesics corresponds to adding multiples of 2π to ϕ to wrap round the spatial circle, so the lengths are all obtained from computing

$$\ell = \log\left((\vec{u}_i^{\perp} \cdot \exp(\phi\xi_L + t\eta_L)\vec{u}_j)(\vec{v}_i^{\perp} \cdot \exp(\phi\xi_R + t\eta_R)\vec{v}_j)\right). \tag{4.28}$$

The result between points on the same exterior region is

$$\ell = \log\left[\frac{4}{r_+^2 - r_-^2}\sinh\left(\frac{r_+ + r_-}{2}(\phi - t)\right)\sinh\left(\frac{r_+ - r_-}{2}(\phi + t)\right]\right) \tag{4.29}$$

and for geodesics between opposite regions it is

$$\ell = \log\left[\frac{4}{r_+^2 - r_-^2}\cosh\left(\frac{r_+ - r_-}{2}(\phi + t)\right)\cosh\left(\frac{r_+ + r_-}{2}(\phi - t)\right)\right] \tag{4.30}$$

where it should be borne in mind that the time evolution here is by the '$H_L - H_R$' Killing field which acts in opposite directions on the two boundaries. The first of these results reproduces the answer obtained with the original HRT proposal [39], slightly generalised to allow for endpoints of the interval to lie at different times. The second can in fact be found from the first by analytically continuing in t, and the result is given for example in [133].

With these expressions we could go on to study entanglement entropies of intervals, mutual information, including entanglement between the two halves of the thermofield double, and so forth. This has already been considered in detail in special cases, for example [63, 70], so we will not dwell on this but instead move on to our next example.

4.3.2 $\mathbb{RP}^2$ geon

We turn now to a simple example of a quotient including orientation reversing isometries, the $\mathbb{RP}^2$ geon, first discussed in relation to holography in [134].

Spatial reflection in AdS_3 is implemented by conjugation with a traceless matrix of determinant -1, hence with eigenvalues ± 1 and squaring to the identity, for example $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Combining this with a hyperbolic element gives an isometry

$$\gamma : p \mapsto gpg^t, \quad g = \begin{pmatrix} \exp\left(\frac{\pi r_+}{2}\right) & 0 \\ 0 & -\exp\left(\frac{-\pi r_+}{2}\right) \end{pmatrix}. \tag{4.31}$$

The geon is obtained by taking the quotient under the group generated by γ, isomorphic to $\mathbb{Z}$.

The element γ^2 is an orientation preserving isometry, and the quotient under the subgroup of the even elements generated by it gives the nonrotating case of the BTZ black hole. Adding in the odd elements corresponds to taking the quotient of the black hole under the involution which swaps the two asymptotic regions, and also rotates half way round the spatial circle. This involution acts without fixed points, so the result is a smooth spacetime (except at the usual BTZ singularities). This leaves a geometry with a single exterior region, isometric to the black hole until the centre of the Einstein-Rosen bridge where the spatial circle has antipodal points identified, as an $\mathbb{RP}^1$. Any time slice thus has the topology of a cylinder with the boundary at one end and a cross-cap inserted at the other, which gives a Möbius strip, or alternatively the real projective plane $\mathbb{RP}^2$ with a boundary, whence the name.

In $d+1$ dimensions, this construction can be generalised to an $\mathbb{RP}^d$ geon by taking a quotient of a Schwarzschild or other more general spherically symmetric black hole with two exterior regions, by the involution swapping the asymptotic regions combined with the antipodal map on the sphere [135, 136].

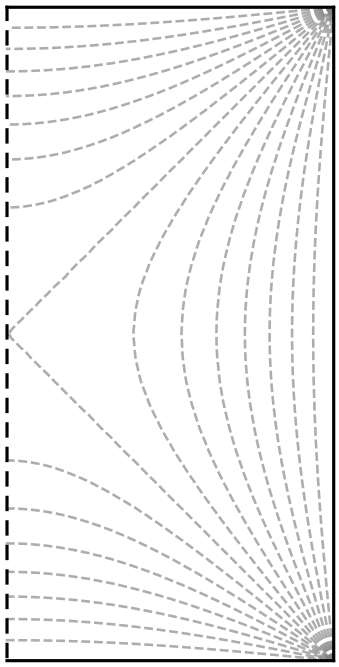


Figure 4.2: The Penrose diagram for the $\mathbb{RP}^2$ geon. The light dashed lines are of constant radius r, including the event horizon at $r = r_+$. Each point on the diagram is a circle, with antipodal points identified at the dashed line on the left. The boundary is on the right, and past and future singularities at the bottom and top.

The $\mathbb{RP}^2$ geon geometry has a single Killing vector field,

$$\xi_L = \xi_R = \frac{r_+}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.32}$$

generating spatial translations. The time translation symmetry of the BTZ black hole is broken by the reflective involution, so is not smoothly extendible into the whole spacetime, so the geon is not stationary. It still generates time translations on the boundary, and exists outside the event horizon as the local Killing vector field $\eta = (\xi_L, -\xi_R)$.

From a field theory point of view, the state at $t = 0$ can be prepared by a path integral over a Möbius strip, exactly as in the orientable case described in section 4.2.6. It is therefore a pure state on a single copy of the CFT. From the Euclidean point of view, swapping the asymptotic regions of BTZ can be viewed as a refection in the Euclidean time, $t_E \mapsto \beta/2 - t_E$, which fixes $t_E = \pm\beta/4$. Taking the quotient of the torus by this reflection along with $\phi \mapsto \phi + \pi$ gives the resulting boundary geometry of a cylinder of length $\beta/2$, with cross-caps at both ends: a Klein bottle. The partition function is given by a path integral over this Klein bottle, a surface with a single real modulus identified with r_+.

Unlike the torus, holographically this state has only one phase. Any phase must be obtainable from a Euclidean bulk with toroidal boundary (the oriented double cover) by taking a quotient by the appropriate involution, which restricts to the correct thing on the boundary torus. This boundary involution extends into the bulk in the black hole phase (giving the geon), and also in the 'thermal AdS' phase where the spatial circle is contractible in the bulk. But in the latter case the involution acts with fixed points at $t_E = \pm\beta/4$, and the resulting Euclidean bulk is not smooth. The consequence is that there is a single phase

at all moduli of the boundary, and there is no analogy to the Hawking-Page transition.

One qualitative way to think of the state, made more precise in [134], is that it is very similar to a thermofield double, being made up of two halves that look thermal individually, but purify one another when taken together. But now the halves are not two copies of the CFT, but the left- and right-movers in a single copy.

Now computation of the lengths of geodesics can be done using the methods described above essentially without modification. In particular, the equation (4.14) still applies, excepting that the matrices describing the quotient may have determinant -1 for orientation reversing elements.

A pair of representative vectors for the boundary is $\vec{u} = \vec{v} = (1,1)^t/\sqrt{r_+}$, and representatives at other times can be found by inserting exponentials of Killing vectors as for BTZ. There are two classes of geodesics, one staying away from the horizon, and one passing through it, which corresponds to inserting a factor of g. The results for lengths of geodesics connecting points at equal times, separated by an angle ϕ, are

$$\ell_1 = \log\left(\frac{4}{r_+^2}\sinh^2\left(\frac{r_+\phi}{2}\right)\right) \tag{4.33}$$

$$\ell_2 = \log\left(\frac{2}{r_+^2}\left(\cosh(r_+(\pi-\phi)) + \cosh(2r_+t)\right)\right), \tag{4.34}$$

which can also be obtained directly from the appropriate results for BTZ. Additionally, there is one simple closed geodesic, going half way round the bifurcation circle, closed due to the identifications, with length πr_+.

This state throws up an interesting question of how precisely to state the topological constraint on allowed geodesics, since if the statement is that the boundary region together with the geodesic should have trivial homology class in $H_1(\mathcal{M},\mathbb{Z})$, we arrive at a contradiction. The state is pure, so computing the entropy of the whole boundary should allow for no geodesic at all: the boundary itself should have trivial homology. But this is not the case in the given definition, since the interpolating manifold filling a constant time slice bulk is not a chain over $\mathbb{Z}$, being nonorientable. The solution[7], described in detail in [4] motivated by a holographic replica trick derivation [36], is to insist on the existence of a spacelike 'two-sided' codimension-1 manifold, where two-sided means it has a globally defined continuous unit normal vector. In an orientable spacetime, this is equivalent to the homology constraint as stated above, but does not follow from such a homology theory in nonorientable spacetimes. It is this version that we will employ in what follows.

The result for the entanglement entropy of a single interval is identical to the answer in the thermal case when the interval is less than half of the boundary. The answers differ when the interval is larger than half since the state is pure, so entanglement entropy of a region equals the entanglement entropy of the complement. The length of the geodesic staying away from the horizon is always shorter than the geodesic passing through it, when

[7]One proposed solution might be to take homology with coefficients in $\mathbb{Z}_2$. This turns out to be a necessary condition, but it is not a strong enough constraint in general.

combined with the closed geodesic required by homology. This can be seen directly from the expressions obtained: the difference $\ell_2 + \pi r_+ - \ell_1$ is smallest when $t = 0$ and $\phi = \pi$, equalling $r_+\pi - \log \sinh^2\left(\frac{r_+\pi}{2}\right)$. This is bounded below by $\log 4$, so always positive.

In fact this can be seen directly from the geometry without any calculations. On the $t = 0$ slice, the geodesic passing through the centre can be joined to the closed geodesic at the bifurcation circle, but the combination is a single curve with sharp corners, and so can be reduced in length by smoothing these out. Hence there is a shorter curve, and hence a shorter geodesic, on the $t = 0$ slice. For nonzero times, it is not even clear that the geodesic passing through the centre is allowable, since there is no closed geodesic to add that will allow for a spacelike interpolating surface, required in a version of the topological constraint including causality. Even allowing it, the geometric argument can be straightforwardly adapted in any case, since the central $\mathbb{RP}^1$ is largest at the bifurcation point.

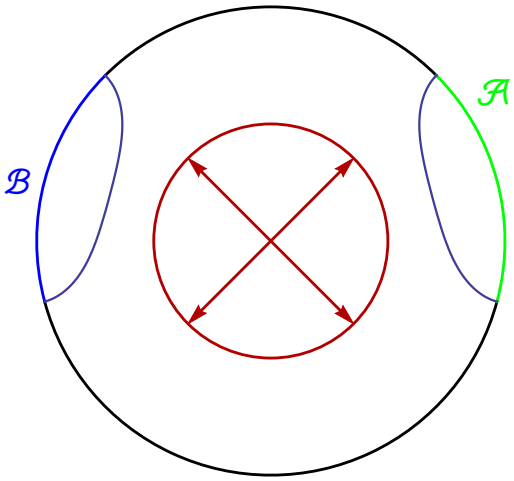
More interesting is the question of the mutual information $I(\mathcal{A} : \mathcal{B}) = S(\mathcal{A}) + S(\mathcal{B}) - S(\mathcal{A} \cup \mathcal{B})$ of two intervals $\mathcal{A}$ and $\mathcal{B}$. Consider the special case where the intervals have equal lengths $l < \pi$, with the centres of the intervals separated by $l < \Delta < \pi$. For the thermal state, this would have three different possible phases. The purity of the state changes one of these, much as in the case of a single interval, by removing the necessity of including the event horizon. Perhaps more interestingly, there is now an additional fourth phase, where the two geodesics pass through the centre. No closed geodesic is required to satisfy the homology constraint, so the argument that these are not shortest does not apply, and in fact for large enough r_+ it turns out that this phase can dominate. The geodesics relevant to the four phases[8] are shown in fig. 4.3.

At $t = 0$, the results for the mutual information in the different phases are

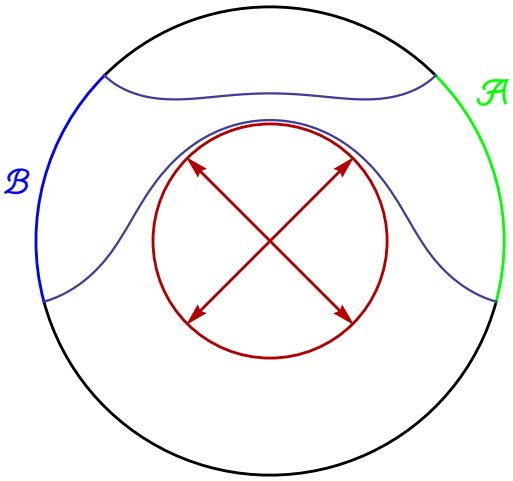
$$
\begin{aligned}
I_1 &= 0 \\
I_2 &= 2\log\left(\frac{\sinh^2\left(\frac{lr_+}{2}\right)}{\sinh\left(\frac{r_+}{2}(\Delta - l)\right)\sinh\left(\frac{r_+}{2}(\Delta + l)\right)}\right) \\
I_3 &= 2\log\left(\frac{\sinh^2\left(\frac{lr_+}{2}\right)}{\sinh\left(\frac{r_+}{2}(\Delta - l)\right)\sinh\left(\frac{r_+}{2}(2\pi - \Delta - l)\right)}\right) \\
I_4 &= 4\log\left(\frac{\sinh\left(\frac{lr_+}{2}\right)}{\cosh\left(\frac{r_+}{2}(\pi - \Delta)\right)}\right)
\end{aligned}
$$

with the mutual information equal to the largest of the four. The most interesting fourth phase is most likely to dominate when the separation Δ is at its largest possible value of π, and the lengths l of the intervals are $\pi/2$. In this case we have $0 = I_1 = I_3 > I_2$ and

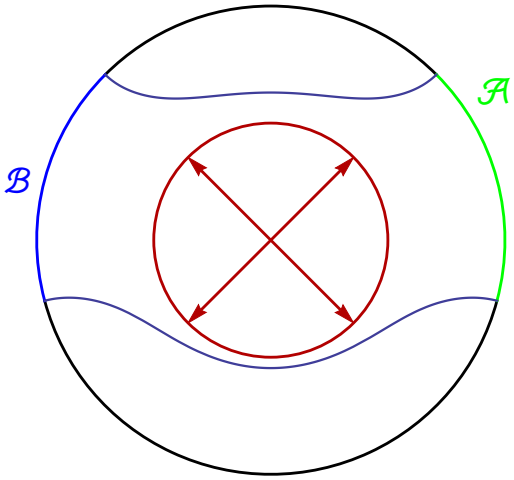
[8] Interpreting phases as different OPE channels in twist operator correlation functions, one might object to a geodesic connecting the left end of one interval to the right end of another, since nonvanishing charges of the twist operators do not allow the vacuum block to appear in the OPE expansion. But this is circumvented here, since the operators should be brought together through an antiholomorphic transition map, reversing the direction of the twist for one operator.



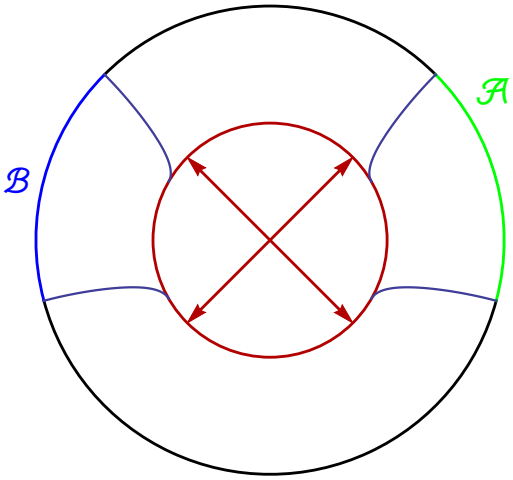
(a) Phase 1. Dominates for small l.



(b) Phase 2. Dominates for small Δ if $l+\Delta < \pi$.

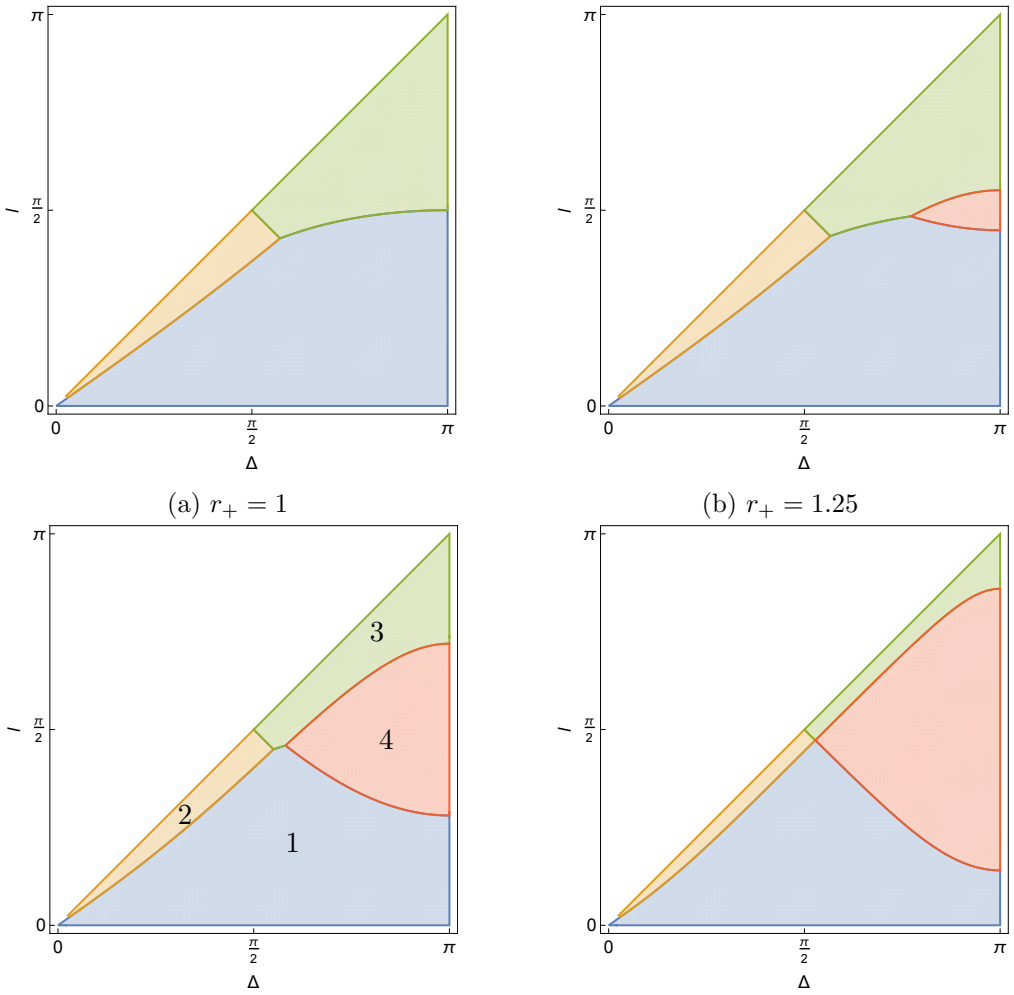


(c) Phase 3. Dominates for small Δ if $l+\Delta > \pi$.



(d) Phase 4. May dominate in some intermediate regime.

Figure 4.3: The choices of geodesics to compute $S(\mathcal{A} \cup \mathcal{B})$ in the $\mathbb{RP}^2$ geon. Options (a) and (b) would be admissible in BTZ, and (c) would additionally require inclusion of the event horizon. The geodesics in (d) pass through the crosscap, the arrows indicating that the antipodal points are identified.



(a) $r_+ = 1$ (b) $r_+ = 1.25$

(c) $r_+ = 2$ (d) $r_+ = 4$

Figure 4.4: Phases of mutual information of two intervals in the $\mathbb{RP}^2$ geon, of length l, plotted vertically, with centres separated by Δ, plotted horizontally, for various values of r_+. The blue, yellow, green and red regions indicate where phases 1,2,3 and 4 respectively dominate, as labelled in subfigure (c).

$I_4 = 4 \log \sinh\left(\frac{r_+ \pi}{4}\right)$, so the upshot is that I_4 can dominate when $r_+ > \frac{4}{\pi} \sinh^{-1} 1 \approx 1.12$. The regions in parameter space where the four phases are dominant are shown for various values of r_+ in fig. 4.4.

To understand the new phase, it helps to consider taking a large r_+ limit. For BTZ, this high temperature limit destroys correlations between all regions that are not parametrically close to the edge of parameter space, so that phase 1 dominates and the mutual information always vanishes. This is not the case in the $\mathbb{RP}^2$ geon, for which there are correlations for sufficiently large or well separated regions with $l + \Delta > \pi$, when the fourth phase dominates:

$$I \sim \max\{0, 2(l + \Delta - \pi) r_+\} \tag{4.35}$$

In particular, for small regions the mutual information is nonzero only when Δ is very close to π. This suggests that a region is most strongly entangled with regions furthest from it.

It would be interesting to understand how this arises from the state of the field theory.

A final exercise is to consider what happens to the mutual information under time evolution. This does not alter the first three phases, but the fourth phase changes to

$$I_4 = 2\log\left(\frac{\sinh^2\left(\frac{lr_+}{2}\right)}{\cosh\left(\frac{r_+}{2}(\pi-\Delta+2t)\right)\cosh\left(\frac{r_+}{2}(\pi-\Delta-2t)\right)}\right). \tag{4.36}$$

Despite the fact that the external region is static, the entanglement entropy is sensitive to the non-staticity behind the horizon. The results are indicative of a very special state at $t = 0$, which thermalizes after some order 1 time. Focussing on the intervals for which the fourth phase is most dominant, namely $\Delta = \pi$ and $l = \pi/2$, we see that the mutual information is positive when $\sinh\left(\frac{\pi r_+}{4}\right) > \cosh(r_+ t)$. The mutual information declines and saturates to zero at some time, which is in fact always bounded by $\pi/4$, for all r_+.

In the large r_+ limit, I_4 has a cross over at $|t| = (\pi - \Delta)/2$, when the geodesic moves from hugging the horizon to staying close to the singularity. Suggestively, this is exactly the time at which all right-movers from $\mathcal{A}$ have had a chance to meet some left-movers from *directly opposite* $\mathcal{B}$, the region $\mathcal{B}$ is most strongly entangled with according to the previous analysis. It begins to decrease linearly, as $2(l - 2|t|)r_+$ until saturating to either zero mutual information at time $l/2$, or to phase 3 at time $(\pi - l)/2$ for sufficiently large regions $l > \pi/2$. It would be interesting to attempt to understand whether this can be interpreted from a quasi-particle picture of the dynamics of entanglement, at least in this limit, and what this implies for the way the entanglement is distributed in the CFT state.

4.3.3 Three boundary wormhole

Before beginning with the cases of quotient groups generated by two elements, we describe a useful way to understand the $SL(2,\mathbb{R})$ group structure, by identifying elements with the tangent space at the identity via the exponential map. For simplicity, we restrict here to the case of nonrotating wormholes, for which the quotient group lies in the diagonal $PSL(2,\mathbb{R})$. The generalisation to the rotating case is relatively straightforward.

Write the generators as exponentials of $\mathfrak{sl}(2,\mathbb{R})$ Lie algebra elements, parameterised by

$$\xi = \frac{1}{2}\begin{pmatrix} z & x-t \\ x+t & -z \end{pmatrix}. \tag{4.37}$$

This three dimensional Lie algebra has a natural Lorentzian inner product from the Killing form, calculable from the determinant, and $PSL(2,\mathbb{R})$ isometries act by conjugation, giving the three dimensional Lorentz transformations. In the picture of $SL(2,\mathbb{R})$ as AdS_3, this Lorentzian structure is inherited from the tangent space of the origin.

The hyperbolic isometries are given by exponentials of spacelike elements. Thinking of the isometries like this as vectors in $\mathbb{R}^{2,1}$, with the coordinates (t, x, z), one generator of Γ may be boosted and rotated to lie along the x-axis. This leaves a single residual symmetry

of boosting in the z direction. Depending on the second generator, there are then two distinct cases depending on whether its z component or t component can be boosted to zero[9].

In the case when the x component can be set to zero, the spacetime that results is a wormhole with three asymptotic boundaries, all connected through a non-traversable bridge. We begin in this section with this class. In the opposing case when the t component may be set to zero, the result is a wormhole with a single exterior region and a torus behind the event horizon, which we move to in the next section (section 4.3.4).

There are three moduli to specify the spacetime, which can be picked in the Lorentzian Lie algebra language as the two lengths and the boost angle between them. A convenient choice of the generators is $g_1 = \exp \xi_1$ and $g_2 = \exp \xi_2$, where the $\mathfrak{sl}(2, \mathbb{R})$ Lie algebra elements are given by

$$\xi_1 = \frac{\ell_1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \xi_2 = \frac{\ell_2}{2} \begin{pmatrix} 0 & e^{\alpha} \\ e^{-\alpha} & 0 \end{pmatrix} \tag{4.38}$$

so that

$$g_1 = \begin{pmatrix} \cosh\left(\frac{\ell_1}{2}\right) & \sinh\left(\frac{\ell_1}{2}\right) \\ \sinh\left(\frac{\ell_1}{2}\right) & \cosh\left(\frac{\ell_1}{2}\right) \end{pmatrix}, \quad g_2 = \begin{pmatrix} \cosh\left(\frac{\ell_2}{2}\right) & e^{\alpha} \sinh\left(\frac{\ell_2}{2}\right) \\ e^{-\alpha} \sinh\left(\frac{\ell_2}{2}\right) & \cosh\left(\frac{\ell_2}{2}\right) \end{pmatrix}. \tag{4.39}$$

The ξs generate translations in the first and second asymptotic regions, and the geodesics connecting the fixed points of g_1 and g_2 lie along the event horizons of these regions. The corresponding element of Γ for the third region in this parametrization is $g_3 = -g_1 g_2^{-1}$. For the spacetime to be free of conical singularities, this generator must also be hyperbolic, which is equivalent to the condition that the boundaries of the fundamental region on the $t = 0$ slice do not meet. This is achieved for sufficiently large α, such that $e^{\alpha} > \coth\left(\frac{\ell_1}{4}\right) \coth\left(\frac{\ell_2}{4}\right)$, and when this is satisfied the length of the third event horizon can be found from (4.15):

$$\begin{aligned} e^{\alpha} = \operatorname{csch}\left(\frac{\ell_1}{2}\right) \operatorname{csch}\left(\frac{\ell_2}{2}\right) \Bigg[\cosh\left(\frac{\ell_3}{2}\right) + \cosh\left(\frac{\ell_1}{2}\right) \cosh\left(\frac{\ell_2}{2}\right) + \\ \left(2 \cosh\left(\frac{\ell_1}{2}\right) \cosh\left(\frac{\ell_2}{2}\right) \cosh\left(\frac{\ell_3}{2}\right) + \frac{\cosh(\ell_1) + \cosh(\ell_2) + \cosh(\ell_3) + 1}{2} \right)^{1/2} \Bigg] \end{aligned} \tag{4.40}$$

A symmetric choice of fundamental region on the $t = 0$ slice is shown in fig. 4.5, along with the geodesics that lie on the event horizons for each exterior region.

In the CFT, this quotient is (in some region of moduli space where this phase dominates) dual to a pure entangled state on three noninteracting circles, prepared by a path integral over a Riemann surface with three boundaries and no handles. Joining this to its reflection on the boundaries, to form the Schottky double as described in section 4.2.6, one obtains a

[9]The third possibility, that the residual vector is null, means that Γ contains parabolic, if not elliptic, elements. We will not consider it here.

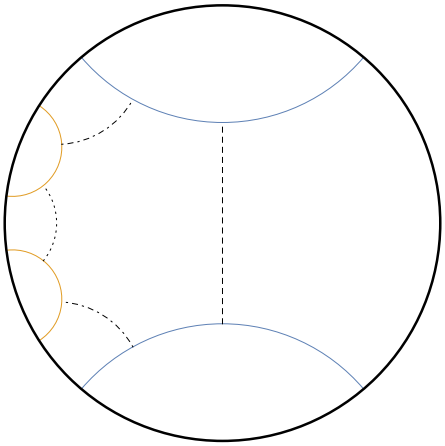


Figure 4.5: The $t = 0$ slice of AdS_3, showing a fundamental region and event horizons for the three boundary wormhole. The blue curves, identified by g_1, and the orange, identified by g_2, mark the edge of the fundamental domain. The dashed, dotted, and dot-dashed lines mark the event horizons of the three exterior BTZ regions.

closed Riemann surface of genus two.

Just as in the case of BTZ, this genus two surface is a boundary to many different on-shell Euclidean bulk spaces. There are five choices of handlebody that respect the time-reversal symmetry, and hence allow for translation back into Lorentzian language, and can be distinguished by connectedness of the Lorentzian section, exactly as for the thermal state. The physics will be described by the dominant phase with the smallest Euclidean action, which will depend on the 3 real moduli of the Riemann surface. One phase gives three disconnected copies of AdS, so there should be no correlations between the copies of the CFT in the classical limit, and 3 phases have one boundary disconnected and the others joined by a BTZ black hole. The final fifth phase is our wormhole geometry. It is an open question to determine the moduli for which each phase dominates, so excepting for certain symmetric situations and limits we do not know whether the wormhole is relevant. See [132] for a more detailed discussion. From here on, we will assume that we are always in the fully connected wormhole phase.

Here, we will look only at entanglement entropy for a single interval in one of the asymptotic regions. There are at least two phases that must inevitably dominate for some regions of the parameter space, analogous to the two phases for small and large intervals in the thermal state. In these cases, the geodesics remain outside the horizon in the region isometric to BTZ, and differ only in deciding which way to pass round the horizon. The closed geodesics required in the case of the 'large region' regime may live either on the horizon of the region in question, or on both the other two horizons if the sum of these lengths is smaller[10]. The more interesting possibility is whether there may be intermediate phases that ever dominate, with geodesics meeting the boundary and yet passing through

[10]There exist many more complicated closed geodesics that suffice to satisfy the homology constraint, but a thorough search as described later has shown them to always be longer than the horizons. This is unsurprising, though we do not know of any proof that this must always hold.

the interior region. It turns out that (at least ignoring the question of which bulk geometry phase dominates) this possibility is realised.

We begin again by choosing representative points in each boundary region. Due to the symmetric way in which we have picked our generators, this is very easy to do in the first and second asymptotic regions, though rather harder in the third. We will not reproduce the results for the third region here, though there will be very similar difficulties encountered in the next section, which will serve to illustrate how to proceed if required.

From fig. 4.5, it is clear that a simple choice of reference points in the first an second regions will be $\vec{u} = \vec{v} \propto (1,0)^t$ and $\vec{u} = \vec{v} \propto (0,1)^t$ respectively, lying at the far right and left of the disc. Normalising these, and translating with the appropriate Killing vectors, just as for the BTZ example, we obtain

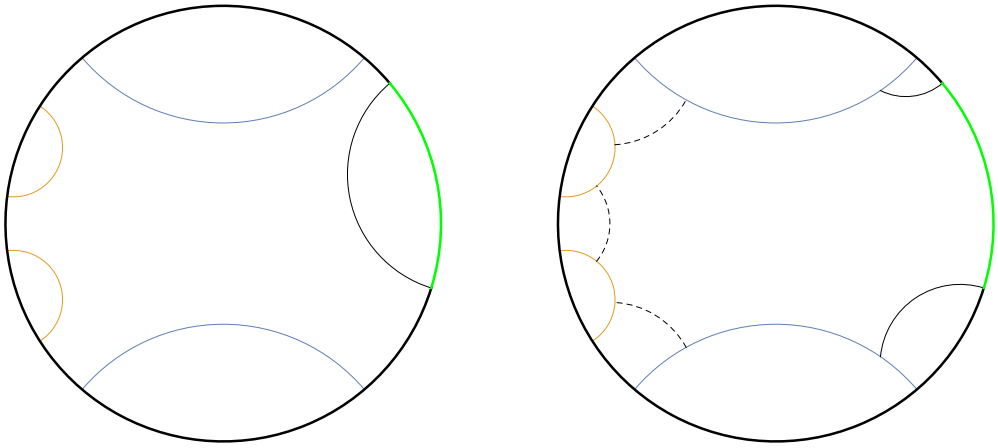
$$\vec{u}_1 = \vec{v}_1 = \sqrt{\frac{4\pi}{\ell_1}} \begin{pmatrix} \sinh\left(\frac{\ell_1\phi}{4\pi}\right) \\ \cosh\left(\frac{\ell_1\phi}{4\pi}\right) \end{pmatrix}, \quad \vec{u}_2 = \vec{v}_2 = \sqrt{\frac{4\pi}{\ell_2}} \begin{pmatrix} \cosh\left(\frac{\ell_2\phi}{4\pi}\right) \\ e^{-\alpha} \sinh\left(\frac{\ell_2\phi}{4\pi}\right) \end{pmatrix} \tag{4.41}$$

restricting to time $t = 0$ (the generalisation to include time dependence is exactly as for BTZ). We will use only the result for the first region here, to compute the entanglement entropy of an interval of length $\Delta\phi$, centred at angle ϕ_0, so between $\phi = \phi_0 \pm \frac{\Delta\phi}{2}$. By symmetry, we need only consider $\phi \in [0, \pi]$, and $\Delta\phi \in (0, 2\pi)$. The extremes of $\phi_0 = 0, \pi$ are when the interval is closest to the second and third asymptotic regions respectively.

The length results are quick to obtain from eq. (4.14) as before, though the exact forms are unrevealing so we will not reproduce them here. One advantage of having a systematic way of finding a relatively simple analytic answer is that it allows for computing a very large number of lengths for many different homotopy classes of geodesic. In the present case, we examined lengths of all geodesics with homotopy classes formed from words in the generators with up to eight letters, for a random choice of over one million sets of parameters, checking which classes dominated in any case, using `Mathematica` [137]. This was performed by first creating a list of all the elements of the group formed from some finite number of generators, and then calculating the appropriate matrix traces analytically. Each of these analytic expressions was turned into a C-compiled function of the parameters $(\ell_1, \ell_2, \ell_3, \Delta, \phi_0)$ for efficiency, and then evaluated on a randomly chosen set of these parameters repeatedly, checking the dominant geodesic each time. This gave a list of all realised possibilities in a few seconds, showing that there are only four phases of interest. An entirely similar calculation for closed geodesics (again checking conjugacy classes from words of up to eight letters in the generators) found that only the three event horizons are relevant.

The four phases, shown in fig. 4.6, correspond to the homotopy classes of the identity, g_1^{-1}, g_2^{-1}, and $g_3^{-1} = g_2 g_1^{-1}$, with the last three supplemented by the event horizons of exteriors 1, 2 and 3 (or a sum of the other two if shorter) respectively.

Each of the last three phases corresponds to saturation of some Araki-Lieb inequality



(a) Phase 1: 1

(b) Phase 2: g_1^{-1}

(c) Phase 3: g_2^{-1}

(d) Phase 4: $g_2 g_1^{-1}$

Figure 4.6: The geodesics giving the four possible phases of entanglement entropy of a single interval, in green, along with the event horizons added to satisfy the homology constraint, marked by dashed lines.

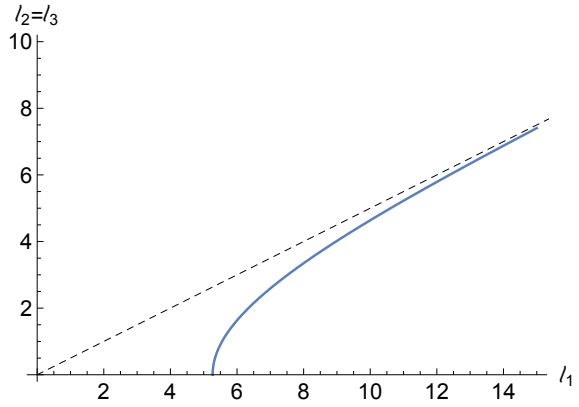


Figure 4.7: Phase diagram for single-interval entanglement as a function of the moduli of the spacetime, in the case when two of the horizon lengths are equal. Below and to the right of the solid line, one of phases 3 or 4 dominates in some region of moduli space, and $S(\mathcal{A})$ has nontrivial dependence on space and time. This is always below the dashed line $\ell_1 = \ell_2 + \ell_3$, where there is a phase transition associated to the closed geodesics.

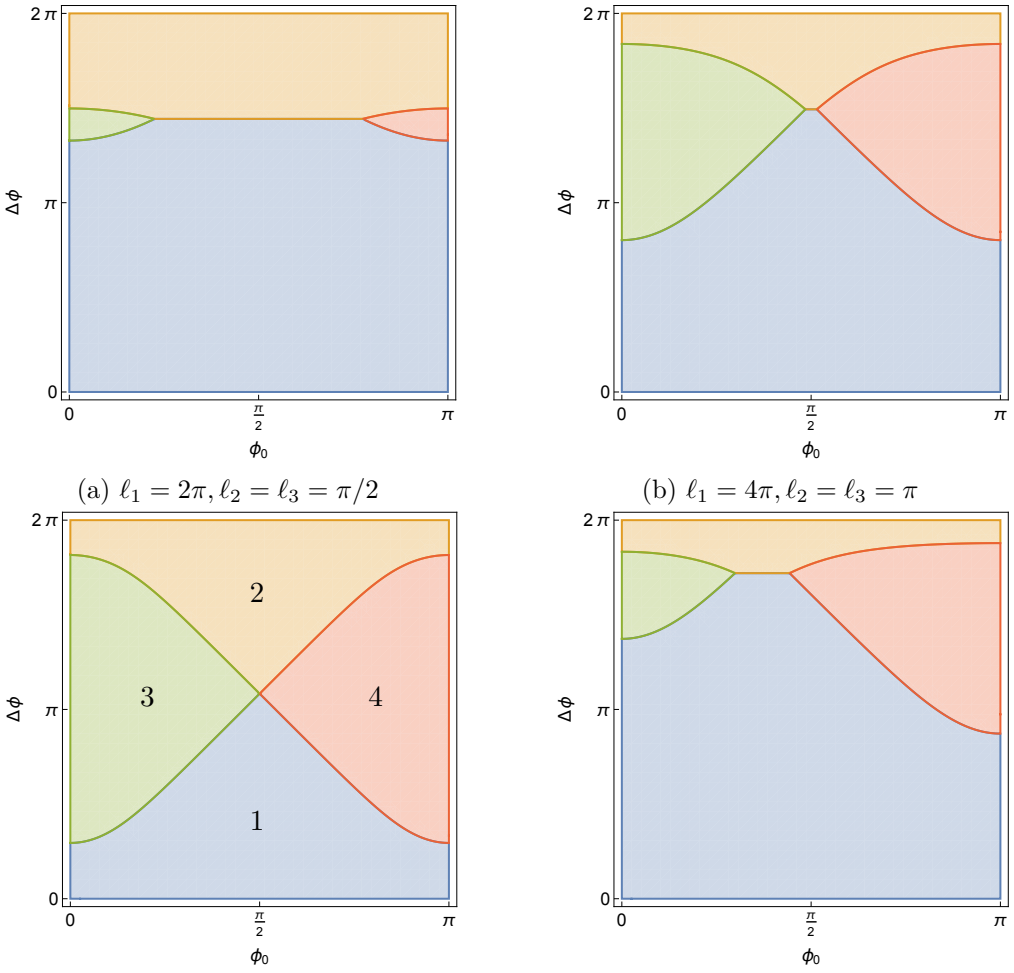
[10]

$$|S(\mathcal{A}) - S(\mathcal{B})| \leq S(\mathcal{A}\mathcal{B}), \text{ saturation when} S(\mathcal{A}) = S(\mathcal{B}) + S(\mathcal{A}\mathcal{B}) \tag{4.42}$$

where $\mathcal{A}$ is the interval in question, and $\mathcal{B}$ is either the complement of $\mathcal{A}$ in boundary 1, as happens in the thermal state [1], or the entirety of boundary 2 or 3. This has a very natural interpretation [138] in terms of the state on $\mathcal{A}\mathcal{B}$, that the Hilbert space of $\mathcal{A}$ can be split into two parts $\mathcal{A}_1$ and $\mathcal{A}_2$ in such a way that the state on $\mathcal{A}\mathcal{B}$ factorizes as a mixed state on $\mathcal{A}_1$ times a pure state on $\mathcal{A}_1\mathcal{B}$. This means in this case that $\mathcal{B}$ is only entangled with some subset of degrees of freedom in $\mathcal{A}$, and the remaining degrees of freedom in $\mathcal{A}$ are entangled only with the remainder of the system.

We find that the most interesting last two phases, where the geodesic passes behind the event horizon, may dominate as long as ℓ_1 is larger than some order one value, and also somewhat larger than ℓ_2 and ℓ_3. We will focus here on the symmetric case when $\ell_2 = \ell_3$, for which a phase diagram is plotted in fig. 4.7. The result here is that for ℓ_1 larger than some critical value, phases 2 and 3 dominate for some interval. For $\ell_{2,3}$ very small, the critical value of ℓ_3 is some order one number, which then increases with $\ell_{2,3}$, approaching $2\ell_{2,3}$ for large horizons but never exceeding it. This means that when the entropy of region 1 is computed from its own horizon radius, rather than the sum of the other two, all single-interval entanglement entropies are in exact agreement with the thermal state. When the other two horizons are small, on the other hand, the two systems are insufficiently entropic to purify a thermal state in the first region. The entanglement entropy of a sufficiently large region notices that the state is not thermal, and may saturate its entanglement with one or other of the boundaries.

For some specific values of the moduli, phase diagrams of varying interval size and length are shown in fig. 4.8. An interesting aspect of these is what they imply for the spatial distribution of entanglement. The Araki-Lieb inequality for the interval and one of



(a) $\ell_1 = 2\pi, \ell_2 = \ell_3 = \pi/2$

(b) $\ell_1 = 4\pi, \ell_2 = \ell_3 = \pi$

(c) $\ell_1 = 6\pi, \ell_2 = \ell_3 = \pi/4$

(d) $\ell_1 = 4\pi, \ell_2 = 2\pi, \ell_3 = \pi$

Figure 4.8: Phase diagrams of entanglement entropy for a single interval in region one, with the position of the centre of the interval plotted horizontally and its size vertically. The blue, yellow, green and red regions (also labelled in subfigure (c)) indicate where phases 1, 2, 3 and 4 respectively dominate, as defined in fig. 4.6.

the other boundaries is most likely to be saturated when they are closest to one another. Regarding the horizon lengths as proxies for the size of each system, when two boundaries are small and one is large, the small boundaries are not only entangled exclusively with the large one, but actually with a spatially localised region within it. This naïve idea of geometric closeness corresponding to entanglement may be a natural guess, but it is striking to see it so precisely realised.

The investigations here should not be viewed as exhaustive, but rather a demonstration of the results that can be quickly obtained from this method. There is much more that can be done in these geometries to understand the entanglement, particularly with measures that are not inherently bipartite, generalising the work in [132] to subintervals. One obvious quantity to look at is mutual information between intervals on different boundaries. Further generalisations would be to look at time dependence, allow the wormholes to spin, and to add more boundary regions. There are hints that there is some simple universal behaviour when the parameters ℓ become large (which is also a region of moduli space where this phase would be expected to dominate), and the techniques here are well suited to be used to understand this analytically. We leave all such investigations for future work.

4.3.4 Torus wormhole

The final example we discuss is a black hole with a single exterior, containing a torus hidden behind the event horizon. Once again, we will consider here only a nonrotating version; a spinning generalisation was discussed at length in [127], and what is done here can be straightforwardly extended to that case. This is a pure state on a single CFT, prepared by a path integral over a torus with a single boundary, a Riemann surface with three real moduli. The field theory interpretation of this state is quite mysterious, but in any case we expect it to be some atypical finely tuned excited state at $t = 0$, breaking translational symmetry in both space and time directions.

The partition function is that of a genus two surface, with a reflection symmetry fixing a single circle lying between the handles and splitting the surface into two parts. Apart from the bulk phase we discuss here, there is an infinite family of time-reflection symmetric Euclidean bulks, corresponding to filling in some combination of cycles on one handle of the surface, and the same (reflected) combination of cycles on the other. These are in correspondence with pairs of coprime integers defining the choice of cycle, much like the $SL(2,\mathbb{Z})$ family of Euclidean black holes generalising BTZ and Euclidean thermal AdS. For these bulks, the Lorentzian spacetime is always just pure AdS. Once again we will not worry about which saddle dominates the path integral, but work in the interesting phase with the caveat that the results apply only in certain regions of moduli space.

As discussed at the beginning of the last section, the geometry comes from a quotient by a free subgroup of $PSL(2,\mathbb{R})$ generated by two elements g and h. Considering these elements as exponentials of Lie algebra elements in a three dimensional Lorentzian space, they come from a pair of spacelike vectors which may now be boosted to simultaneously have vanishing timelike component. The moduli therefore are the two lengths of the vectors

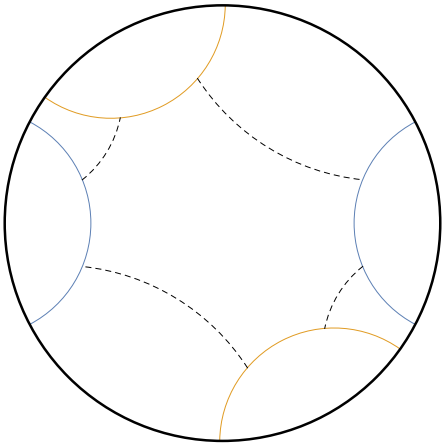


Figure 4.9: The $t = 0$ slice of AdS_3, showing a fundamental region and event horizon for the torus wormhole. The blue curves are identified by g, and the orange by h. The event horizon here is the set of dashed geodesics joining the fixed points of $ghg^{-1}h^{-1}$, $hg^{-1}h^{-1}g$, $g^{-1}h^{-1}gh$ and $h^{-1}ghg^{-1}$.

$\lambda > 0$ and $\mu > 0$, and the angle $\alpha \in (0, \pi/2]$ between them. Picking a convenient basis, Γ is generated by

$$g = \exp \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} = \begin{pmatrix} e^{\lambda} & 0 \\ 0 & e^{-\lambda} \end{pmatrix}, \tag{4.43}$$

$$h = \exp \begin{pmatrix} \mu \cos\alpha & \mu \sin\alpha \\ \mu \sin\alpha & -\mu \cos\alpha \end{pmatrix} = \begin{pmatrix} \cosh\mu + \cos\alpha \sinh\mu & \sin\alpha \sinh\mu \\ \sin\alpha \sinh\mu & \cosh\mu - \cos\alpha \sinh\mu \end{pmatrix}. \tag{4.44}$$

The resulting geometry is most easily understood once again by looking at the a fundamental region on the $t = 0$ Poincaré disc, as in fig. 4.9. This looks very similar to the three boundary case above, except that the identifications are not between neighbouring semicircles but between opposite sides. To see the topology, this is exactly like the identification of opposite edges of a square to form a torus, except with the corners cut off; this makes a torus with a single asymptotic boundary.

One computational difficulty that we encounter here is that the elements of Γ enacting the translation along the boundary by 2π are not so simple as before. Such 'horizon words' are here the conjugates of $ghg^{-1}h^{-1}$; any point on the boundary of the Poincaré disc will lie between the fixed points of one such element, which is the exponential of the generator of boundary translations there. This conjugacy class also defines the homotopy class of the event horizon, which is why we call the representatives horizon words. This is exactly like the third asymptotic region in the above three boundary wormhole. Note that the horizon word here is trivial when abelianised, so has trivial homology, since it is the boundary of everything behind the horizon. This is a manifestation of the purity of the state.

Calculating any such word and using equation (4.15) to find its length, we get the

horizon radius of the black hole

$$r_+ = \frac{1}{\pi}\cosh(2\sin^2\alpha\sinh^2\lambda\sinh^2\mu - 1) \tag{4.45}$$

from which it is apparent that we must take the angle between the generators large enough so that $\sin\alpha\sinh\lambda\sinh\mu > 1$. As with the three boundary wormhole, this can be understood from the requirement that the circles marking the edges of the fundamental region in fig. 4.9 do not overlap, and ensures that all elements of the group Γ are hyperbolic.

To find a nice regularisation to represent boundary points, as before it is simplest to first identify the local Killing vector ξ implementing translations. Since the horizon word is not in a simple form, this requires first inverting a matrix exponential, for example $e^{2\pi\xi} = -ghg^{-1}h^{-1}$, but for hyperbolic $SL(2,\mathbb{R})$ matrices, this is a straightforward procedure. We then must choose a reference point, which must be in the right part of the boundary of the Poincaré disc, lying on the proper side of the fixed points of ξ so that it is the correct local generator of translations. Since the fixed points move as the parameters vary, there is no universal choice to be made here, but it depends on the parameters. One practical possibility comes from examining the fixed points of ξ, being the eigenvectors, and taking the average of the two angles of the resulting vectors; this gives an especially simple answer in the symmetric case $\alpha = \frac{\pi}{2}$. With this initial point chosen, the others may be found by translating with the Killing vector, after which an overall scale can be fixed if desired, by matching to the BTZ result. Note that this will not give boundary representatives in the boundary fundamental region in fig. 4.9, but a connected fundamental region in the boundary covering space.

Apart from these small extra practical difficulties, the calculations proceed much as before, and we will not reproduce details here. We again examine the entanglement entropy for a single interval on the boundary. It turns out to be much harder in this case to systematically rule out the majority of geodesics built from long words in the generators as always being longer, as was possible for the three boundary wormhole, making a complete characterisation difficult. This is perhaps because the torus behind the horizon may get very 'twisted' relative to the generators g and h for some parameters, so that the shortest geodesics are more complicated words in Γ. Using a different choice of generators for Γ may then look more natural, and allow the geometry to be described by some different set of parameters in which the short geodesics are simple words. This is loosely analogous to describing a flat torus with upper half plane modular parameter τ, when short geodesics may look complicated if τ is not picked in a fundamental region, but instead related by some modular transformation.

In any case, to simplify things we now focus on a special case of the most symmetric geometry, by picking $\alpha = \pi/2$ and $\lambda = \mu$. The one remaining modulus can be tuned to give any value for the horizon radius. The geometry then goes from having no symmetry to a dihedral group, the symmetry group of the square, acting on the boundary as rotations by $\pi/2$ and four possible reflections of the circle.

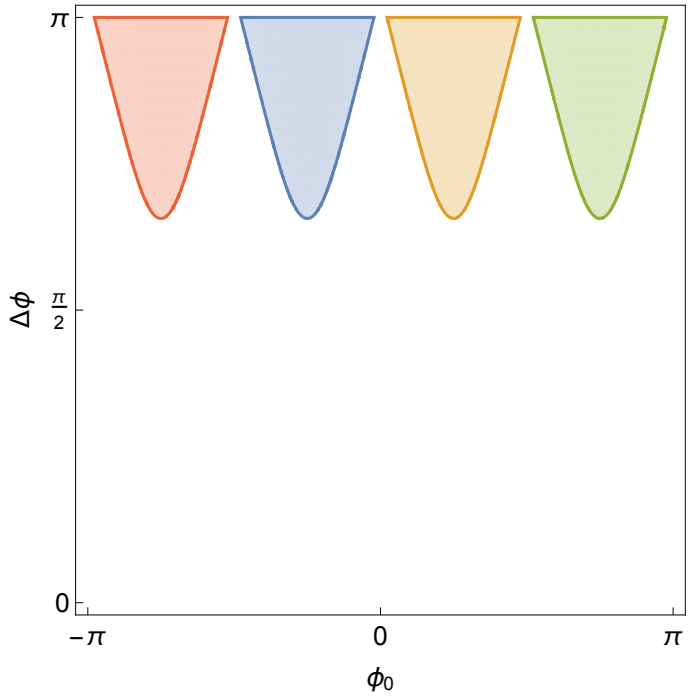


Figure 4.10: Phases of entanglement entropy for a single interval in the symmetric torus wormhole, with horizon radius $r_+ = 5$. The position of the centre of the interval is plotted horizontally and its size, up to π, plotted vertically. The four coloured regions correspond to dominance of four different geodesics passing behind the event horizon, all related by the discrete symmetries of the spacetime. The uncoloured region is where the trivial geodesic dominates, giving the BTZ result.

In this special case, a systematic study of the geodesics is more useful, showing that the more complicated geodesics do not dominate the entanglement entropy of a single interval. We followed a similar method to the three boundary case, of looking at words built from some maximal number of generators over a large sample of different parameters, including for closed geodesics. Restricting the interval to have its centre lying in a range of angles from 0 to $\pi/2$, as allowed by the discrete symmetries, and to have length up to π, as may be done since the state is pure, in fact gives just one nontrivial contributing geodesic, associated to the group element g. The three related by the discrete symmetries can be associated with h, $h^{-1}g^{-1}h$ and $gh^{-1}g^{-1}$. The entanglement entropy is computed including the closed geodesics in the conjugacy classes g^{-1}, h^{-1}, g or h respectively (all of equal length here) to satisfy the homology constraint.

These phases dominate when the horizon radius and entangling interval are sufficiently large, and further depend on the position of the interval, as shown in fig. 4.10. It would be interesting to see if relaxing the symmetry allows for more than these simple phases to dominate, or whether it simply distorts the shape of the regions of dominance.

We do not have any good field theory interpretation for this result. It would be interesting to study in more detail, relax the symmetry assumptions, and to understand what it implies for the state of the field theory. These preliminary results at least show

that the answer is not trivial, so entanglement entropy is a useful observable.

4.4 Relations between Euclidean and covariant entanglement entropy

4.4.1 Analytic continuation

In a Euclidean context, there is a derivation of the Ryu-Takayanagi formula from a Euclidean quantum gravity argument [36], by performing the replica trick in the bulk Euclidean spacetime. In time-reflection symmetric spacetimes, one may freely pass between Euclidean and Lorentzian descriptions (see section 4.2.6), by analytic continuation, or by a Hartle-Hawking procedure regarding the $t = 0$ slice as an initial Cauchy surface, which has zero extrinsic curvature due to the reflection symmetry, allowing for a well-defined Lorentzian evolution. From this, the argument can be regarded as proving the Ryu-Takayanagi prescription for regions lying at such a surface of time-reflection symmetry where the Euclidean description is available, of which static geometries are a special case.

This leaves at least two unanswered questions: firstly, how does the derivation generalise to states with no time-reflection symmetry? It is unclear how to pass to a Euclidean description in such a situation, if possible at all, and we will make no attempt to progress in this direction. The second is whether we can do something weaker, and generalise to time-reflection symmetric geometries, but where the region whose entanglement entropy we would like to compute does not lie at the point of time symmetry.

The natural thing to do here would be to begin with endpoints of an interval on the boundary at $t = 0$, which lives naturally in either Lorentzian or Euclidean spacetimes, and to evolve by some Euclidean time τ in the Euclidean boundary. The Euclidean quantum gravity argument may then be used to justify computing an entanglement entropy for the new interval from a geodesic length in the Euclidean bulk. Finally, to get the result for the real time evolution, assume analyticity and analytically continue the answer to $t = -i\tau$.

It is, however, not in general obvious how to carry out this procedure since the notion of Euclidean time evolution does not appear to be unambiguously defined. This works in BTZ, since there is a unique timelike Killing vector orthogonal to the translation symmetry, and many approaches requiring analytic continuation, such as [63], have relied on this (or equivalently, continuation of a conserved energy). But more complicated Riemann surfaces do not have any continuous symmetries, so this approach does not generalise. Further, it is not clear that the result thus obtained is in fact the length of a geodesic in the Lorentzian spacetime. The technology at hand solves both problems in the present context.

We begin by describing the relevant Euclidean time evolution in the context of a quotient by a Fuchsian group $\Gamma \subseteq PSL(2, \mathbb{R})$. Given a point on the boundary of the AdS covering space at $t = 0$, described by a singular matrix $p(0)$, there is a translation Killing field of the quotient $\xi \in \mathfrak{sl}(2, \mathbb{R})$ defined locally, being in the static case the horizon generator described in section 4.3.4. Any point on the boundary of the Poincaré disc lies between the

fixed points of some unique $\gamma \in \Gamma$, and $\gamma = e^{2\pi\xi}$ defines ξ. Translation (in either signature) is implemented by $p \mapsto e^{a\xi}\, p\, e^{a\xi^t}$; in Lorentzian signature the orthogonal translation implements time evolutions via $p(t) = e^{t\xi}\, p(0)\, e^{-t\xi^t}$. Motivated by this, continuing to Euclidean signature by $t = -i\tau$, the correct time translation must therefore be $p(\tau) = e^{-i\tau\xi}\, p(0)\, e^{i\tau\xi^t} = e^{-i\tau\xi}\, p(0)\, (e^{-i\tau\xi})^\dagger$; this is a sensible action on the Riemann sphere since the matrix on the right is the conjugate of the matrix on the left. It corresponds to following a geodesic on the Euclidean boundary with the flat metric, orthogonal to the $t = 0$ slice, where the metric matches with the flat Lorentzian metric. This is locally uniquely defined, but of course cannot be extended to the whole boundary (except for the torus) consistently with the quotient. An alternative way of characterising it is as the elliptic Killing field on the Riemann sphere with the same fixed points as the generator of translations along the boundary, with an appropriately normalised length.

Now with this notion of Euclidean time evolution, the algebraic expressions for the lengths of the geodesics (4.14) and (4.22) in the two signatures are manifestly analytic in t or τ, and follow from one another by continuation $t = -i\tau$. With the assumption that entanglement entropies should be appropriately analytic (saving for phase transitions between different geodesics), this could be interpreted as a proof that the covariant proposal [39] follows from the Euclidean calculation in this limited set of circumstances.

This works most straightforwardly in the case that the quotient is by a Fuchsian group Γ, with the initial points $p(0)$ at $t = 0$ lying on the real axis of the Riemann sphere. In this case the time-reflection invariant slice in the bulk is exactly the quotient of the upper half plane, with hyperbolic metric, by Γ, so the initial geodesic lies on the static slice, as do all the closed geodesics, so the analytically continued lengths all come from real geodesics. However, the Euclidean bulk geometries only follow from Fuchsian groups in a specific phase; most phases come from non-Fuchsian Schottky groups, in which case the Lorentzian bulk is disconnected. For each connected component of the bulk at $t = 0$ the Schottky group can by conjugation be put in a form such that the boundaries lie on the real axis. If the initial points lie in the same connected part, and further the initial geodesic can be lifted to covering space in such a way that its endpoints both lie on the real axis, the above argument still goes through, and the Lorentzian geodesic lengths still follow from the appropriately analytically continued Euclidean argument. This captures the fact that by HRT, when the spacetime is disconnected the entanglement entropy (at leading order in G_N at least) becomes disconnected also. Geodesics may only connect points in the same connected component, and the allowed homotopy classes of geodesics are those of the component of spacetime in which they live.

4.4.2 Euclidean geodesics without Lorentzian analogues

From a Euclidean point of view, these restrictions on geodesic endpoints and their allowed homotopy classes seem very much less natural. Even with initial points at $t = 0$ and without performing the analytic continuation, viewing the Lewkowycz-Maldacena computation from a purely Euclidean point of view, there is no reason to require that the geodesics (open or

closed) lie on the $t = 0$ slice. The argument requires the time-reflection symmetry only in order to translate to Euclidean signature; the bulk duals computing Rényi entropies may be $\mathbb{Z}_n$ replica symmetric without being time-reflection symmetric, with the reflection symmetry restored in the $n \to 1$ limit, leading to surfaces lying away from the time-symmetric slice. From the standpoint of the Lorentzian section, such saddle points would have no geometric interpretation, but would appear as complexified geodesics. This could, for example, allow for nonzero mutual information between intervals in a phase where they lie in disconnected components of the spacetime[11].

Complex entangling surfaces have been considered before [139], but these are not quite the same as what we consider here, which are real geodesics, hence with real length, but living in the Euclidean section of the spacetime. For an explicit illustration, we finish the section with a computation in the low temperature phase of the thermofield double state, to check whether the possibility is realised.

The protocol of section 4.2.5 can be employed to quickly compute the lengths of geodesics in Euclidean thermal AdS. The spacetime can be described as a quotient of H^3 by the group generated by the $PSL(2,\mathbb{R})$ transformation with associated Möbius map $z \mapsto e^{-\beta} z$. This is Fuchsian, but not with respect to a basis with the time reflection implemented by complex conjugation (time reflection here is instead $z \mapsto -1/\bar{z}$), so passing directly to Euclidean signature by putting it in the diagonal subgroup of AdS isometries does not achieve the correct thing here. The boundary circles can instead be chosen at $|z| = 1$ and $|z| = e^{\beta/2}$, represented by vectors in $\mathbb{C}^2$ by

$$\vec{u}_L = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\phi/2} \\ e^{-i\phi/2} \end{pmatrix}, \quad \vec{u}_R = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{\beta/4 + i\phi/2} \\ e^{-\beta/4 - i\phi/2} \end{pmatrix}. \tag{4.46}$$

From this it is immediate to apply (4.22) to get the lengths of geodesics between points on the same boundary, given by the vacuum $\log \sin^2\left(\frac{\Delta\phi}{2}\right)$ answer, and on opposite boundaries, joined through the torus to get

$$l = \log \left| \sin\left(\frac{\Delta\phi + i\beta/2}{2}\right) \right|^2 = \log \left[\cosh^2\left(\frac{\beta}{4}\right) \sin^2\left(\frac{\Delta\phi}{2}\right) + \sinh^2\left(\frac{\beta}{4}\right) \cos^2\left(\frac{\Delta\phi}{2}\right) \right]. \tag{4.47}$$

This later class of geodesics pass half way round the thermal circle to join boundary points through the Euclidean section, despite them being in disjoint components of the Lorentzian spacetime. They could give a dominant contribution to the mutual information if they are ever sufficiently short. This has the best chance of happening for large intervals of length π, at the same angular position on both sides, so there are geodesics passing from $\phi = 0$ on one side to $\phi = 0$ on the other, and the same between points at $\phi = \pi$. The usual geodesics for the entanglement entropy then have regularised lengths $\log \sin^2\left(\frac{\pi}{2}\right) = 0$. The inherently complex geodesic lengths then come from the last equation at $\Delta\phi = 0$ to

[11] We take the naïve point of view of assuming that the least-action saddle always dominates the path integral, but the possibility that it is not on the path of steepest descent should be borne in mind.

get $\log \sinh^2\left(\frac{\beta}{4}\right)$. The upshot is that these geodesics are short enough to dominate for $\beta < 4 \sinh^{-1} 1 \approx 3.5$. But if this is the case, we clearly have $\beta < 2\pi$, which implies that we are in fact in the high temperature phase where the dominant geometry is the black hole.

In this simple example, it is apparent that the geodesics in the Euclidean space with no Lorentzian counterpart are not the most relevant for computing entanglement entropies. However, it is interesting that this required the extra information about the dominant saddle point geometry, depending on the moduli of the boundary Riemann surface. This can be to some extent intuited from a very naïve picture of determining the correct phase, in which the g shortest cycles (for a genus g surface) are filled in. If moving along a geodesic round a nontrivial cycle in the Euclidean direction gives a short answer, then this may indicate that this cycle on the boundary is itself short (being the Euclidean time circle here), and should have been filled in to give a more dominant bulk action.

If the Ryu-Takayanagi prescription is to be regarded as following from the Euclidean quantum gravity replica computation, it is clearly difficult to rule out the possibility of complex geodesics for the Lorentzian prescription, in the precise sense used here. The techniques described here are useful for investigating this further in more involved examples. It would be interesting to know if this phenomenon can be ruled out entirely, or whether it has some physical relevance. If they could be important, and are not just an artefact of the special symmetric situation, then it is crucial to understand what 'complex geodesic' means more generally, in spacetimes with no time translation or reflection symmetry.

4.5 Discussion

The main result obtained is a completely algebraic description for regularised lengths of geodesics in solutions to pure gravity with negative cosmological constant, coming from a description of such solutions as quotients. This gives the ability to do such calculations without finding coordinate patches or their overlaps, and without solving a single differential equation, in situations where there are no available symmetries. The class of geometries this covers includes essentially all pure 3D gravity spacetimes, including states with nonzero angular momentum at the boundary, and nonorientable geometries.

We demonstrated the practical application of the formula for computing entanglement in various states of interest, in many of which an approach using more naïve techniques would have been intractable. The first novel results obtained were in the single-exterior $\mathbb{RP}^2$ geon. With access to a single interval covering less than half the boundary circle, there is no difference from the thermal state, though the entanglement entropy of an interval larger than half differs to due to the purity of the state. However, the mutual information between two intervals shows that the dual CFT state is far from being a typical microstate at $t = 0$, with highly tuned nonlocal correlations, any given region appearing to be most strongly entangled with regions furthest away from itself. Time evolution quickly destroys these correlations, and it would be interesting to understand better the dynamics of entanglement in this thermalisation.

The next, more involved, example was an entangled state of three noninteracting CFTs, described geometrically by three exterior BTZ regions all joined through a wormhole. Here we find that the entanglement entropy of a single interval is sensitive to the details of the geometry hidden behind the horizon, in particular showing dependence on space and time despite the exterior geometries being symmetric under these translations. There are circumstances where the answer saturates various Araki-Lieb inequalities, which show that entanglement is distributed between the three CFTs in a spatially ordered manner.

The spacetime in the final example is a single exterior black hole with a torus hidden behind the event horizon. Such states with nontrivial hidden topology are particularly mysterious from a field theory standpoint, and our preliminary investigations show that entanglement entropy is a practical and useful observable to analyse them. The results show that even in the most symmetric case, the entanglement entropy of a single interval of sufficient size knows about more than the thermal behaviour, in particular showing the breaking of translation symmetries in space and time.

The surface has only been scratched in the examples here, and there is much more to understand about the physics of entanglement in these geometries. The ability to compute entanglement entropy efficiently in states with no symmetry, with nonzero momentum, including dynamics, and over a range of moduli provides us more generally with a versatile 'laboratory' for expanding our knowledge of the area. We leave more detailed studies to future work.

The final chapter made inroads into obtaining an analytic continuation procedure to obtain HRT from a Euclidean quantum gravity derivation. It would be interesting to understand whether this can be generalised to a less prescriptive set of circumstances, allowing geometries with fields turned on and higher dimensions, as well as relaxing the time-reflection symmetry condition required to translate to Euclidean signature.

Finally, we discussed a very specific way in which HRT might fail, if such a Euclidean quantum gravity derivation is believed. This comes from a spontaneous breaking of the time reflection symmetry when computing the Rényi entropies, in such a way that the symmetry is restored in the $n \to 1$ limit. The consequence would be a geodesic in the Euclidean bulk section computing entanglement entropy, with no geometric analogue in the Lorentzian spacetime. It would, further, invalidate the geometric proofs of properties required for consistency, such as strong subadditivity [40] and causality [41]. We describe a simple example where such things turn out to be subdominant, for which the Euclidean version of the geodesic length calculation is well-suited, and it would be interesting to check dominance in more involved cases.

This is related to two other possible problems. One is the assumption of no spontaneous breaking of the cyclic $\mathbb{Z}_n$ part of the replica symmetry required in [36]. This is required to make any progress with the argument but it is difficult to rule out and hard to interpret. Perhaps a more basic problem comes from a similar symmetry assumption implicit in the interpretation of the bulk geometries themselves. The Hartle-Hawking procedure requires the time-reflection symmetric slice from which to evolve the Lorentzian geometry, but it

may be that the dominant Euclidean saddle point spontaneously breaks this symmetry, so obtaining the dominant Lorentzian geometry becomes impossible. It is known that this never occurs for the torus, but nothing is known at higher genus. A proof that such reflection symmetries are never spontaneously broken would solve all the above problems at a stroke (the dihedral replica symmetry group is generated by the reflections); a counterexample on the other hand would cast much into doubt.

On a more speculative note, the simple and universal form of the algebraic result obtained for entanglement entropy is very suggestive that there may be an alternative way to obtain it directly from conformal field theory. A natural way to attack this problem is in the spirit of previous work [34] in vacuum, where the Rényi entropies are obtained by correlation functions of twist operators on the sphere, matched to the gravity calculation [35]. The key ingredient is that at large central charge with few low lying operators, there is a universal answer coming from the vacuum block in a conformal block expansion of OPEs, with different geodesics coming from different OPE channels. The first challenge is therefore to classify the OPE channels of operators on higher genus Riemann surfaces, and to obtain a matching with the classification of both the choice of bulk saddle, as well as geodesics in the bulk, including the closed geodesics. This would in itself be very interesting, as it should, for example, give a microscopic interpretation of the homology constraint.

Chapter 5

Topological aspects of holographic entanglement entropy

This chapter is a reproduction of the paper *Topological aspects of generalized gravitational entropy* [4], written in collaboration with Felix Haehl, Tom Hartmann, Donald Marolf, and Mukund Rangamani.

The paper revisits the topological constraint in the Ryu-Takayanagi prescription, in the light of the Euclidean quantum gravity derivation [36]. The restriction on topologies of allowed minimal surfaces was, before this, an *ad hoc* prescription applied by hand for consistency, and it is unclear whether it follows from the Lewkowycz-Maldacena computation. We showed that the situation is quite subtle, since the simplest and strongest statement of topology is false, as illustrated by a collection of examples. We then described and proved a weaker sense in which the homology constraint follows from the construction.

The Lewkowycz-Maldacena argument uses the replica trick, described in section 1.1.3, constructing bulk geometries bounded by the replicated boundary Euclidean spacetime. Assuming the bulk geometry respects the cyclic symmetry of the boundary, one may then take a quotient by this symmetry; the expectation is then that there is a set of fixed points that, under the quotient, becomes a conical singularity. It is this singularity that is the central object of the argument, and becomes the surface $\mathcal{E}$, whose area computes the entanglement entropy.

Perhaps the most natural guess as to the origin of the topology in the prescription is that the conical singularities obtained by this procedure always have appropriate homology, but this turns out not to be true. However, if we demand that a conical singularity of a certain topology can be obtained for every value of the replica number q (with appropriate defect angle) by the construction, then the homology constraint follows.

5.1 Introduction

Holography provides an intriguing connection between quantum information and geometry, in part inspired by the geometrization of quantum entanglement by the Ryu-Takayanagi (RT) proposal [30, 31] and its covariant generalization [39] (HRT). These proposals identify a particular bulk codimension-2 surface as the geometric encoder of the entanglement structure in the dual field theory. Whilst the original proposals owed their origins to analogies with black hole entropy and covariant entropy bounds, we now have a derivation of the RT formula (in situations with time reflection symmetry) from a gravitational path integral courtesy of Lewkowycz and Maldacena (LM) [36].

This may at first seem like a complete story for the time-reflection symmetric states of a holographic QFT. Although it requires certain assumptions, the LM construction gives a first principles derivation of how the minimal surface of the RT proposal comes about from a gravitational path integral. But while the LM construction captures the dynamical part of the RT conjecture, to our knowledge it has yet to be established whether the quantum gravity path integral employed by LM is cognizant of the topological constraints that must be imposed on the RT surface. This is the primary question that will concern us in this paper.

To appreciate the issues involved, recall that RT (HRT) proposes that the holographic entanglement entropy of a given region $\mathcal{A}$ of the field theory is given by the area of a minimal (extremal) bulk surface $\mathcal{E}$ anchored on the boundary $\partial\mathcal{A}$ of $\mathcal{A}$, with $\mathcal{A}$, $\partial\mathcal{A}$ considered to lie in the boundary of the bulk spacetime. The extremal surface $\mathcal{E}$ is required to be homologous to the region $\mathcal{A}$ in question [32, 45]. The cleanest phrasing of this statement to our knowledge appears in [111], [41] for the RT and HRT proposals respectively. One requires that there be a spacelike codimension-1 interpolating homology surface $\mathcal{R}_{\mathcal{A}}$ whose only boundaries are $\mathcal{E}$ and $\mathcal{A}$. Without this constraint, the holographic formula would be at odds with known quantum mechanical properties of entanglement entropy, such as strong subadditivity; the topological prescription has been motivated as a natural, but *ad hoc* way of ensuring consistency. The simplest rationale for the homology constraint is that, in its absence, a subsystem $\mathcal{A}$ and its complement $\mathcal{A}^c$ can end up having the same entanglement even when the state of the system is impure (e.g., in a black hole geometry).

One can intuitively motivate this picture by realizing that one must introduce a cut for fields along $\mathcal{A}$ when computing matrix elements of the reduced density matrix $\rho_{\mathcal{A}}$ via the path integral. In particular, the replica construction for computing Rényi entropies requires that the operators of the QFT are cyclically permuted when one crosses the cut. Since local QFT operators are the boundary values of bulk fields, one expects the cut to extend into the bulk as well. The boundary of the bulk branching surface $\mathcal{R}_{\mathcal{A}}$ is formed from $\mathcal{E}$ and $\mathcal{A}$ so that the homology constraint is satisfied; see Fig. 5.1 for an illustration. Existence of the homology surface $\mathcal{R}_{\mathcal{A}}$ ensures that fields are appropriately branched and one can treat the bulk geometry itself as a branched cover over some fundamental domain.

Based on the above arguments, one might imagine that the converse holds true trivially

at the level of topology. That is, given a replica-symmetric bulk saddle with the correct boundary conditions, might we be guaranteed to find a codimension-2 defect and a codimension-1 interpolating surface implementing the homology constraint? Surprisingly, this turns out not to be true. One can construct bulk geometries with the requisite replica symmetry but which nevertheless do not admit an appropriate branching surface $\mathcal{R}_{\mathcal{A}}$. While it is plausible that such geometries are never dominant in the bulk path integral, their presence begs the question "under what conditions does the LM construction give rise to the homology constraint?" We will address this question in some detail below. A simplified version of our final statement is that, as long as one has a family of geometries parameterized by a real parameter q which for all positive integer values of q admits a branched cover description, then we recover the homology constraint from the LM construction.

Before beginning, we pause to dispel a potential confusion. The reader may well ask whether the above condition coincides with the assumptions actually made by Lewkowycz and Maldacena in their original paper [36]. The potential confusion arises from the fact that the phrasing of the assumptions in [36] is subject to at least two different interpretations, which are naturally termed 'global' and 'local'. Indeed, the present authors do not agree among themselves as to which interpretation best fits the words written in [36]. Under the global interpretation, LM assumes that the replica geometries may be analytically continued to $q \sim 1$ with global topology outside the conical singularity given by a trivial $\mathbf{S}^1$ bundle $\mathbf{S}^1 \times X$ for some X. This global assumption coincides with the intuitive discussion above and immediately implies the homology constraint, as the homology surface is given by any global section of this trivial bundle.

However, the global assumption is rather stronger than one might like. As mentioned above, there is no a-priori reason why a (q-fold quotient of a) given replica-symmetric solution at integer q should have this structure. Furthermore, as may be readily seen by noting that there is no obstacle to constructing bulk solutions with small conical singularities on homology-violating surfaces, the global product structure is not required by the assumption that the $q \to 1$ limit of the bulk solutions be described by small conical singularities. Such considerations motivates us to focus on the alternate and weaker local interpretation of the LM assumptions, which does not determine the global structure and imposes the $\mathbf{S}^1 \times X$ structure only locally along the conical singularity. The implications are two-fold. First, as is necessary to allow LM to work in topologically nontrivial examples (such as computing the entropy of an interval on a torus), the $\mathbf{S}^1$ fiber need not be defined far from the singularity. Second, even the region near-but-outside the singularity is allowed to be a non-trivial $\mathbf{S}^1$ bundle over some X. Our main result can now be restated as saying that this local interpretation of the LM assumptions *also* implies the homology constraint so long as the family of geometries exists at all $q \geq 1$ and is an appropriate $\mathbb{Z}_q$ quotient of a smooth q-fold replica at all integer q.

The outline of the paper is as follows: in §5.2 we give a rapid overview of the concepts we need from QFT and holography vis-à-vis entanglement considerations. We also formulate

the above issue concerning the homology constraint in detail. In §5.3 we present counterexamples to the naïve intuition that the homology constraint is an automatic consequence of existence of a branched cover. In §5.4 we formulate an essential topological consistency condition for branched covers as constructed by LM. This condition is a statement about consistency of the field theory replica trick with the existence of bulk branched covers at all integer values of q. We then illustrate and prove in §5.5 that this topological consistency condition is in fact equivalent to the homology constraint on the RT prescription. We conclude with a discussion in §5.6. Further technical details and a review of algebraic topology relevant for our proof can be found in the appendices.

5.2 Review of holographic entanglement

To set the stage for our discussion, let us quickly review the salient features of the RT/HRT and LM constructions that will play a role below. Consider a holographic d-dimensional QFT with a bulk gravity dual in asymptotically $d+1$ dimensional AdS spacetime. We will limit our discussion to situations where the QFT is planar, strongly coupled, and has a suitable gap in its spectrum, so that we can regard the bulk as a two derivative effective field theory which we take to be Einstein-Hilbert gravity coupled to matter degrees of freedom.[1]

5.2.1 The RT and LM constructions

The QFT resides on a background geometry $\mathcal{B}$ foliated by Cauchy surfaces Σ_t.[2] The dual gravitational background to this field theory is $\mathcal{M}$ and has $\mathcal{B}$ as its conformal boundary. Specifically, we will take $\mathcal{M}$ (and all other bulk regions) to denote the conformal compactification, so that $\mathcal{M}$ is a manifold with boundary: $\partial\mathcal{M} = \mathcal{B}$. See Table 5.1 for an overview of our notation.

We are interested in computing the entanglement entropy in such a holographic QFT for a region $\mathcal{A} \subset \Sigma_t$ of the boundary geometry. The HRT prescription for computing holographic entanglement entropy requires us to consider a bulk codimension-2 extremal surface $\mathcal{E} \subset \mathcal{M}$ anchored on $\partial\mathcal{A}$ and takes the area of $\mathcal{E}$ to give the boundary entaglement, i.e.,

$$S_{\mathcal{A}} = \frac{\text{Area}(\mathcal{E})}{4\,G_N}\,, \qquad \partial\mathcal{E} = \mathcal{E} \cap \mathcal{B} = \partial\mathcal{A}\,. \tag{5.1}$$

Crucially the bulk extremal surface is required to satisfy a *homology constraint* originally motivated in [32]. Usually this is stated as the requirement that the extremal surface must be homologous to the region $\mathcal{A}$. More precisely, following e.g., [41, 111] we will take this

[1] The discussion generalizes to other planar QFTs dual to (classical) higher derivative gravitational theories following [37, 38], for which we would be evaluating some other local functional on a codimension-2 surface $\mathcal{E}$. As we primarily focus in on topological aspects, much of what we describe later will go through unmodified.

[2] We often use language appropriate for Lorentzian spacetimes despite focusing on Euclidean geometries. We will be interested in surfaces $\Sigma_{t=0}$ which sit at a moment of $\mathbb{Z}_2$ time-reflection symmetry (allowing thereby translation between the two cases by a suitable Wick rotation).

Boundary regions		
Symbol	Description	Dimension
$\mathcal{B}$	full boundary manifold	d
Σ_t	fixed-time slice	$d-1$
$\mathcal{A}$	subregion of Σ_t	$d-1$
$\partial\mathcal{A}$	entangling surface	$d-2$
$\tilde{\mathcal{B}}_q$	q-fold branched cover of $\mathcal{B}$ used in the replica trick	d
Bulk regions		
Symbol	Description	Dimension
$\mathcal{M}$	full bulk manifold (with $\partial\mathcal{M}=\mathcal{B}$)	$d+1$
$\mathcal{E}$	RT minimal surface (with $\partial\mathcal{E}=\partial\mathcal{A}$)	$d-1$
$\mathcal{R}_\mathcal{A}$	homology surface interpolating between $\mathcal{E}$ and $\mathcal{A}$	d
$\tilde{\mathcal{M}}_q$	smooth bulk replica manifold (with singular $\partial\tilde{\mathcal{M}}_q=\tilde{\mathcal{B}}_q$)	$d+1$
$\mathcal{M}_q$	fundamental domain of $\tilde{\mathcal{M}}_q$ when $\tilde{\mathcal{M}}_q$ is a branched cover (i.e. $\tilde{\mathcal{M}}_q/\mathbb{Z}_q$)	$d+1$
$\mathbf{e}_q$	branching surface of the branched cover $\tilde{\mathcal{M}}_q\to\mathcal{M}_q$	varies
$\mathbf{r}_q$	homology surface interpolating between $\mathbf{e}_q$ and $\mathcal{A}$ (if it exists)	d

Table 5.1: Definition of boundary and bulk regions that we consider in the course of our discussion.

to mean that there exists a bulk codimension-1 spacelike surface $\mathcal{R}_\mathcal{A}$ which interpolates between the extremal surface and the boundary region of interest. To wit,

$$\exists\,\mathcal{R}_\mathcal{A}\subset\mathcal{M}:\quad \partial\mathcal{R}_\mathcal{A}=\mathcal{E}\cup\mathcal{A}\,. \tag{5.2}$$

In the RT/HRT constructions, taking $\mathcal{A}$ to be spacelike is an additional restriction on the allowed minimal/extremal surfaces, though it is naturally incorporated in the maximin proposal of [40]. We will have much more to say about this constraint below.

We need one more ingredient to make contact with the LM path integral derivation. This ingredient is the replica trick for computing powers of the reduced density matrix $\rho_\mathcal{A}$ whose von Neumann entropy is the entanglement entropy under discussion. We define

$$\rho_\mathcal{A}=r_-Tr_{\mathcal{A}^c}\rho\,,\qquad S_\mathcal{A}=-\operatorname{Tr}\left(\rho_\mathcal{A}\log\rho_\mathcal{A}\right)\,, \tag{5.3}$$

where ρ is the total density matrix on $\Sigma_t=\mathcal{A}\cup\mathcal{A}^c$. After setting $t=0$ as in footnote 2 and passing to Euclidean signature, ρ can be viewed as a state prepared by some path integral over $\mathcal{B}$ (now denoting a Euclidean boundary) cut along $\Sigma_{t=0}$ where boundary conditions are imposed on fields at $\Sigma_{t=0}$ to compute particular matrix elements of ρ. Up to a normalizing factor, the trace over $\mathcal{A}^c$ to obtain $\rho_\mathcal{A}$ is implemented by sewing up the part of the cut along $\mathcal{A}^c$, leaving $\mathcal{B}$ with a cut only along $\mathcal{A}$. The replica construction in the QFT then proceeds by sewing q copies together cyclically along the cuts at $\mathcal{A}$ to construct a singular manifold $\tilde{\mathcal{B}}_q$ whose partition function computes $\operatorname{Tr}(\rho_\mathcal{A}^q)$. This then allows us to

recover the q^{r-th} Rényi entropy of the QFT via:

$$S_{\mathcal{A}}^{(q)} = \frac{1}{1-q} \log \mathrm{Tr}(\rho_{\mathcal{A}}^{q}) = \frac{1}{1-q} \log \frac{Z_q}{Z_1^q}\,, \tag{5.4}$$

where Z_q is the partition function of the QFT on $\tilde{\mathcal{B}}_q$ and Z_1 that on $\mathcal{B}_1 \equiv \mathcal{B}$. The entanglement entropy $S_{\mathcal{A}}$ of (5.3) is recovered in the limit $q \to 1$.

The LM construction first implements this computation of Rényi entropies holographically by extending the replica trick into the bulk. It then extracts the entanglement entropy as above by giving a geometric implementation of the continuation to non-integer q.

To compute Z_q, one proceeds by obtaining a bulk manifold $\tilde{\mathcal{M}}_q$ with boundary $\tilde{\mathcal{B}}_q$ (for some explicit examples see [35, 42]); as always, the partition function is simply given by the on-shell gravitational action computed on this geometry. This bulk computation follows the usual rules of Euclidean quantum gravity and the LM saddle point analysis remains valid when we analytically continue $q \in \mathbb{Z}_+ \mapsto q \in \mathbb{R}$, as long as $(q-1)\,\ell_{AdS}/\ell_{Planck} \gg 1$.

To get from the replica spacetimes $\tilde{\mathcal{M}}_q$ of LM to the RT minimal surface one proceeds as follows. Since the boundary geometry $\tilde{\mathcal{B}}_q$ is a q-fold cover over $\mathcal{B}$ branched at $\partial\mathcal{A}$ with cyclic $\mathbb{Z}_q$ symmetry, we can restrict attention to a single 'fundamental domain' by focusing on the quotient spacetime $\tilde{\mathcal{B}}_q/\mathbb{Z}_q$; this is just a copy of $\mathcal{B}$ itself. Assuming the bulk saddle point geometry $\tilde{\mathcal{M}}_q$ to respect replica symmetry,[3] we may similarly consider the bulk quotient $\mathcal{M}_q = \tilde{\mathcal{M}}_q/\mathbb{Z}_q$. LM focus on the case where the action of $\mathbb{Z}_q$ on $\tilde{\mathcal{M}}_q$ has a codimension-2 fixed point set $\mathbf{e}_q$ with boundary $\partial\mathcal{A}$. This $\mathbf{e}_q$ is to be identified as the progenitor of the extremal surface $\mathcal{E}$. It is assumed to result in a conical defect of angle $\frac{2\pi}{q}$ in $\mathcal{M}_q$. The desired QFT partition function Z_q on $\tilde{\mathcal{B}}_q$ is then q times the bulk action on $\mathcal{M}_q$, computed without a contribution from the conical defect.

The point of considering this quotient is that it allows continuation to arbitrary real values of q. The protocol is to find a geometry $\mathcal{M}_q$ with boundary $\mathcal{B}$ and from which a 'singular' codimension-2 surface $\mathbf{e}_q$ ending at $\partial\mathcal{A}$ has been excised. One then imposes as a further boundary condition that $\mathbf{e}_q$ is a conical defect of opening angle $\frac{2\pi}{q}$. The geometry is fixed by minimizing the action subject to this requirement, with no contribution to the action from the singularity. In the $q \to 1$ limit we require $\mathcal{M}_q \to \mathcal{M}$. The defect surface $\mathbf{e}_q$ then becomes the minimal area surface $\mathcal{E}$ in the Euclidean geometry. One can furthermore argue that the contribution to Z_q localizes on this surface, giving a correction to the action proportional to the area, in such a way that the area of the extremal surface computes the entanglement entropy.

[3] The $\mathbb{Z}_2$ time-reflection symmetry about $t = 0$ of the state ρ intertwines with the cyclic $\mathbb{Z}_q$ symmetry of the replica construction, to give a larger dihedral symmetry group $\mathbb{D}_q$; see [140] for its relevance in computing Rényi entropies. We refrain from utilizing the full dihedral symmetry, allowing for the possibility that the LM construction gives a surface that does not lie at $t = 0$ in the bulk. Therefore, in what follows, replica symmetry will always refer to the cyclic $\mathbb{Z}_q$ group.

5.2.2 A question of homology

The LM construction shows that the computation of Rényi entropies for arbitrary positive real q can be performed by finding geometries $\mathcal{M}_q$ with boundary $\mathcal{B}$, but with a conical defect of angle $\frac{2\pi}{q}$. At integer values of q, via $\mathcal{M}_q = \tilde{\mathcal{M}}_q/\mathbb{Z}_q$ this should be related to a nonsingular replicated bulk $\tilde{\mathcal{M}}_q$ whose boundary is the replica $\tilde{\mathcal{B}}_q$ (on which this $\mathbb{Z}_q$ acts as the replica symmetry). When this is the case, we say that $\mathcal{M}_q$ lifts to a q-fold branched cover. Of primary interest to us is the relationship between the homology constraint on the one hand, and this lifting of the singular bulk geometry to a nonsingular replicated bulk on the other.

The bulk conical defect, coming from the fixed point set of the replica symmetry, is the codimension-2 surface $\mathbf{e}_q$ anchored on the boundary at $\partial\mathcal{A}$. On the boundary $\tilde{\mathcal{B}}_q$, upon traversing a small loop around $\partial\mathcal{A}$, one passes through $\mathcal{A}$, and goes from one copy of $\mathcal{B}$ to the next. It is tempting to imagine a natural picture of the full geometry that arises from continuing this reasoning into the bulk as follows: At the level of topology, $\tilde{\mathcal{M}}_q$ is formed by sewing together q copies of $\mathcal{M}_q$ along some $\mathbf{e}_q$ such that traversing a small loop around $\mathbf{e}_q$ also results in a change of sheet in the cover $\tilde{\mathcal{M}}_q$ of $\mathcal{M}_q$. We shall investigate the correctness of this picture below.

This picture is straightforward in cases where there is a codimension-1 interpolating surface $\mathbf{r}_q$ bounded by $\mathbf{e}_q \cup \mathcal{A}$: in other words, if the homology constraint is satisfied by the conical defect. Cutting along $\mathbf{r}_q$, and gluing together q copies cyclically, just as on the boundary, builds the covering space $\tilde{\mathcal{M}}_q$ at the level of topology. Passing through $\mathbf{r}_q$ will cause a change of sheet in the cover, just as passing through $\mathcal{A}$ changes sheets on the boundary. So any $\mathcal{E}$ obeying the homology constraint[4] naturally leads to a family of bulk saddles $\tilde{\mathcal{M}}_q$ obeying the correct boundary conditions. We may say that $\mathcal{M}$ lifts to $\tilde{\mathcal{M}}_q$, with $\mathcal{E}$ lifting to $\mathbf{e}_q$.[5] See Fig. 5.1 for an illustration of this scenario.

However, the converse, or at least the strongest converse one might propose, fails to be true. More specifically, the following three statements will be demonstrated in sections §5.3, §5.4, and §5.5.

(i). The existence of a two-sided interpolating codimension-1 surface $\mathbf{r}_q$ implies that $\mathcal{M}_q$ lifts to a branched cover (homology $\implies$ lift). This is what we have informally argued above.

(ii). There are branched covers with the correct boundary conditions which cannot be

[4]The natural homology constraint to assume is the one in Lorentz signature. We will take the Lorentz signature spacetime to be time-orientable. It then admits a nowhere-vanishing vector field along which $\mathcal{R}_\mathcal{A}$ can be deformed until it lies in the moment of time symmetry. It follows that the homology constraint also holds in Euclidean signature. For future reference, we also note that Lorentz-signature time-orientability makes $\mathcal{R}_\mathcal{A}$ two-sided, i.e., it has a continuous unit normal. In case of an orientable bulk, two-sidedness is equivalent to orientability of $\mathcal{R}_\mathcal{A}$.

[5]When all surfaces of interest lie on the $t = 0$ slice, $\mathcal{R}_\mathcal{A}$ is uniquely determined by $\mathcal{E}$ and $\mathcal{A}$ as long as that slice has no closed, boundaryless components. More generally, $\mathcal{R}_\mathcal{A}$ is unique if $H_d(\mathcal{M}) = 0$, which we expect to hold in most physical situations (see, for example [141]). If $H_d(\mathcal{M}) \neq 0$ then there is an ambiguity, though it will not make a difference to the entanglement entropy in the classical limit (but it will matter for Rényi entropies and for quantum corrections).

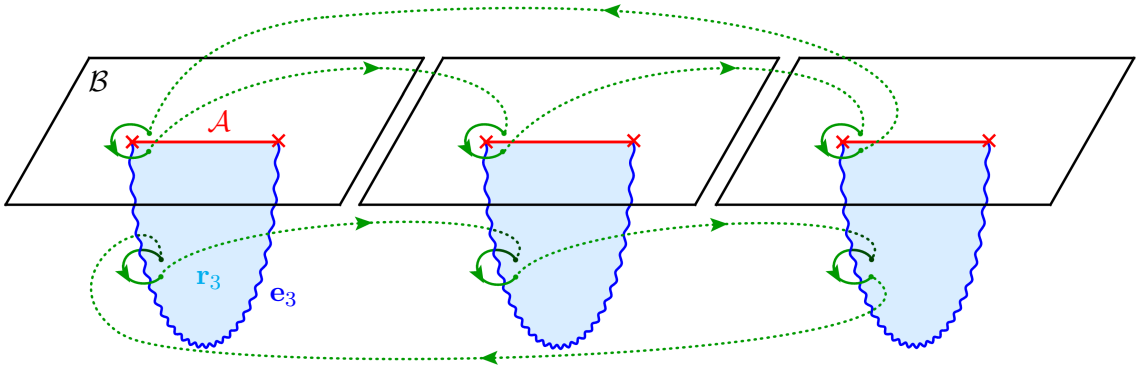


Figure 5.1: Replica construction in the boundary and bulk for $q = 3$. The replica symmetric q copies of the field theory on $\mathcal{B}$, form a q-fold branched cover $\tilde{\mathcal{B}}_q$ which fixes the asymptotic data for the bulk problem. The bulk covering spacetime $\tilde{\mathcal{M}}_q$ has a $\mathbb{Z}_q$ symmetry with fixed point locus $\mathbf{e}_q$ (shown as the wavy lines) anchored on $\partial\mathcal{A}$. Typically one also encounters via this construction a bulk interpolating surface $\mathbf{r}_q$ (the light blue branching surface) in the bulk whose boundaries are $\mathbf{e}_q$ and $\mathcal{A}$. Conventional intuition dictates that the bulk spacetimes are all covers over a single fundamental domain (one of the components in the picture) branched over the codimension-1 surface $\mathbf{r}_q$. Passing through this surface cycles through the sheets of the bulk in a fashion identical to passage through $\mathcal{A}$. The homology condition posits that such an $\mathbf{r}_q$ exists. We argue that this picture is accurate as long as we are suitably careful with the notion of allowed branched covers. As $q \to 1$, $\mathbf{r}_q \to \mathcal{R}_\mathcal{A}$ and $\mathbf{e}_q \to \mathcal{E}$.

realized in this way. That is, for given $q \in \mathbb{Z}_+$, there can exist an $\mathcal{M}_q$ formed from a quotient of a branched cover which does not admit an interpolating surface $\mathbf{r}_q$ between $\mathcal{A}$ and the fixed point set $\mathbf{e}_q$ (lift $\not\Longrightarrow$ homology).

(iii). However, given a continuous family of bulk geometries $\mathcal{M}_q$ parameterized by real q which for every $q \in \mathbb{Z}_+$ lifts to a q-fold branched cover, we will show that each $\mathcal{M}_q$ admits an interpolating surface $\mathbf{r}_q$. Taking $q \to 1$ then shows that $\mathcal{E}$ satisfies the homology constraint as desired (lift $\forall\, q \implies$ homology).

We also note that, on top of this, the fixed point set arising from a $\mathbb{Z}_q$ quotient may give rise to something other than a $\frac{2\pi}{q}$ conical defect. It is possible to generate fixed point sets with the wrong codimension, and also to engineer situations wherein the codimension-2 fixed point set has an incorrect defect angle. In §5.3 we give examples where both these scenarios can be realized.

Below we generally confine ourselves to topological arguments. In particular, we refrain from employing dynamical information from the path integral to constrain $\mathcal{M}_q$. This is in part due to the fact that classification of all replica invariant saddles with given boundary geometry is a notoriously hard problem (even for $d = 2$).[6] In the same spirit, the notion

[6] The kinematical aspect of our analysis is reminiscent of the first attempt to prove the RT proposal in [45] where the homology condition was introduced. But while [45] argued that the bulk should be a branched cover satisfying homology, we would instead like to ascertain the conditions under which the homology condition becomes automatic. We thank Matt Headrick for a discussion on this issue.

of replica symmetry should always be understood in a topological sense. Formally, the bulk will be a multi-sheeted surface – we refrain from specifying the detailed geometry on the sheets, but do keep track of what the boundary conditions of the gravitational path integral imply for moving across the sheets (see §5.4).

5.3 Exemplifying replica symmetric homology violation

We begin by considering some simple examples, focusing only on the topology, which illustrates some of the unexpected features that may be encountered.

5.3.1 A torus with a crosscap

Suppose we wish to compute the entropy of the thermal state of a two-dimensional field theory living on a circle (or equivalently the entanglement entropy between the two halves of the thermofield double). The state is defined by a path integral over a cylinder of length $\beta = 1/T$. The q^{r-th} Rényi entropy can be computed by the partition function on q copies of the cylinder glued cyclically, which gives a torus $\tilde{\mathcal{B}}_q$ of length $q\beta$. The replica symmetry is implemented by rotating through β in the Euclidean time direction. The region $\mathcal{A}$ of interest becomes a spatial circle at fixed Euclidean time, with empty boundary.

To compute the partition function holographically, one must choose how the torus is to be 'filled in' to obtain a bulk $\tilde{\mathcal{M}}_q$. One option is to fill the spatial circle with a disk, like thermal AdS, in which case there are no fixed points under the replica symmetry, so the quotient $\mathcal{M}_q$ is smooth, and the entanglement entropy obtained would vanish. The region $\mathcal{A}$ is contractible in the bulk, so this is consistent with the homology constraint.

If, on the other hand, the Euclidean time circle is filled in with a disk, as in the BTZ black hole geometry, $\mathcal{A}$ is not contractible. But the centre of the disk is fixed under the rotation implementing replica symmetry. So it gives rise to a circle $\mathbf{e}_q$ of fixed points and a conical defect in the quotient $\mathcal{M}_q$. This defect becomes the bifurcation circle of the event horizon in the $q \to 1$ limit. There is an obvious interpolating surface joining it to $\mathcal{A}$ on the boundary, so again the homology constraint is obeyed.

However, ignoring for now the equations of motion, there are more topologies that might in principle be allowed. One is to fill the Euclidean time circle not with a disk, but using a cross-cap. (In what follows, the spatial circle will play little role, so we focus on some 2-dimensional slice corresponding to a point on this circle.) This means that we first fill it with an annulus, taking the outside edge of the annulus to be the boundary, but we then identify antipodal points of the inside edge so that passing through the inside edge causes one to jump to the point directly opposite. The replica symmetry can be implemented by rotation of the cross-cap in the obvious manner. This gives a space $\tilde{\mathcal{M}}_q$ with the topology of a Möbius strip times the spatial circle. The (single) edge of the Möbius strip is the boundary Euclidean time circle.

There are now two cases, depending on whether q is odd or even, both of which are illustrated in Fig. 5.2. We begin with q odd. In this case, there are no fixed points of the

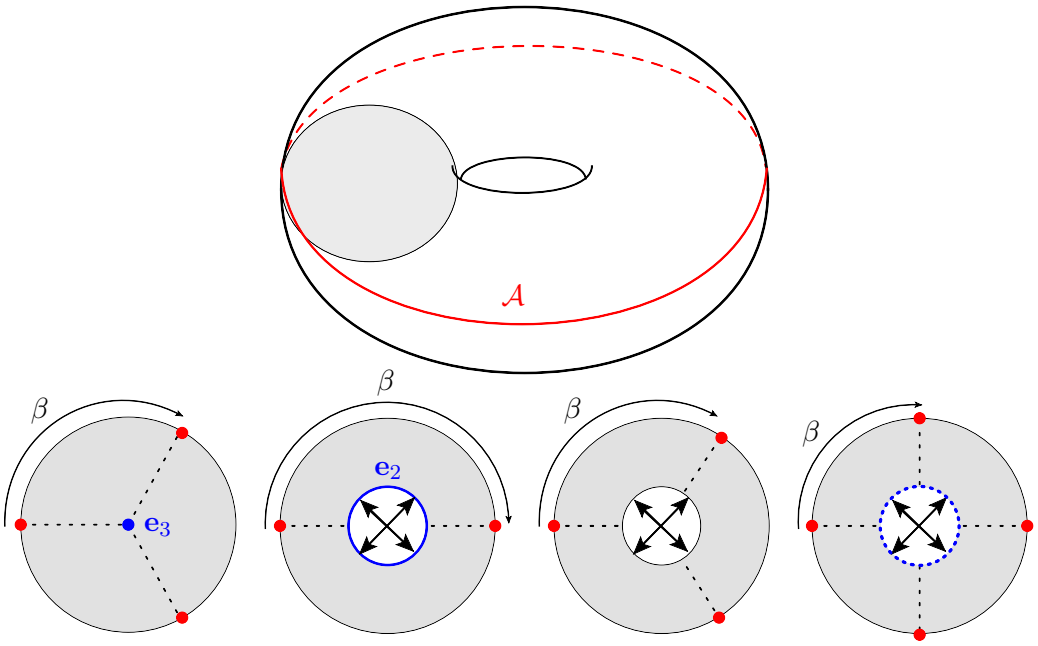


Figure 5.2: Different ways of filling the boundary torus. The replica $\mathbb{Z}_q$ symmetry acts as a rotation along the Euclidean time direction by β. Case (a) shows a slice of $\tilde{\mathcal{M}}_3$ after filling the boundary time circle with a disk; the replica fixed point set $\mathbf{e}_3$ (blue) is a codimension-2 surface in the centre of the circle. Cases (b), (c) and (d) show slices of $\tilde{\mathcal{M}}_q$ for $q = 2, 3, 4$, respectively, after filling the boundary torus with a cross cap. For $q = 2$ the fixed point set $\mathbf{e}_2$ is a codimension-1 orbifold plane wrapping the cross cap. For $q \geq 3$ there is no fixed point set under $\mathbb{Z}_q$ (i.e. $\mathbf{e}_q$ is empty) and the homology constraint is violated. However, we indicate in (d) that for even q the cross cap itself is still a fixed point set under the subgroup $\mathbb{Z}_2$ of rotations by $\frac{q}{2}\beta$, resulting in an orbifold plane under the quotient.

replica symmetry, and the quotient $\mathcal{M}_q$ is a smooth geometry with the same topology as $\tilde{\mathcal{M}}_q$. But $\mathcal{A}$ has non-trivial homology in $\mathcal{M}_q$: there is no surface in the bulk whose only boundary is a spatial circle. In this way, the homology constraint is violated because the fixed point set of the replica symmetry (being empty here) is not homologous to the boundary region.

The second case occurs for even q. While no surface is fixed by every element of the replica group, the $q/2$ replica symmetry now rotates half way round the time circle. On the inner edge of the annulus where points are identified with their antipodes, this symmetry thus takes every point to itself. In the quotient $\mathcal{M}_q$, the resulting singular set is the inside circle of the annulus times the spatial circle. This set is not codimension-2, but instead codimension-1; it is a $\mathbb{Z}_2$ orbifold plane. Thus, for even q there at least exists a fixed point set under a $\mathbb{Z}_2$ subgroup of $\mathbb{Z}_q$, despite the fixed points of the full replica group $\mathbb{Z}_q$ being empty for $q > 2$.

There are several objections that one might raise to this example. The most obvious is that a metric is never constructed, and that there can never be any genuine saddle points with the given topology. The clearest refutation of this is to give an example sharing all the same qualitative features, but with a metric. One can in fact give such a construction for pure three dimensional gravity. To explain the idea, consider the example given instead by starting from a geometry with two disjoint torus asymptotic boundaries, and then taking a $\mathbb{Z}_2$ quotient by swapping the two tori and simultaneously rotating half way round the time circle. This gives a cross-cap in the time circle as described here. Now generalize it to start with not tori, but higher genus Riemann surfaces with negative Euler characteristic. There is an easy way to put a constant negative curvature metric on these geometries, as in the Maldacena-Maoz wormhole [142]. If the Riemann surface has a fixed-point free involutive isometry, the $\mathbb{Z}_2$ quotient of this combined with swapping the two boundaries gives the relevant example.

A second objection is that this fails to satisfy the homology constraint in a very particular way, by having no fixed points under the replica symmetry. This is impossible if the region $\mathcal{A}$ has a nonempty boundary, since there must be a set of fixed points extending from $\partial\mathcal{A}$ into the bulk. We will show later in this section that it is possible to circumvent this objection and construct an example with $\mathcal{E}$ nonempty, but not homologous to $\mathcal{A}$.

Finally, these geometries might never be dominant saddles in the path integral. While this is plausible, it is very difficult to see how their dominance can be ruled out in general. For this reason, we would like to make arguments that apply using only the topology.

5.3.2 Implications for the homology constraint

We believe this example has a real lesson to teach us: the homology constraint follows from the holographic replica trick only when the defect arises from a $\mathbb{Z}_q$ quotient *at every positive integer* q. In the context of the LM argument, one may start with a bulk geometry $\mathcal{M}$ with boundary $\mathcal{B}$, and pick some extremal surface $\mathcal{E}$ ending on $\partial\mathcal{A}$. Now introduce a small conical defect along $\mathcal{E}$, and adjust the geometry so it remains on-shell away from

the defect, which we now call $\mathbf{e}_q$. The angle of the conical singularity can be dialed to $\frac{2\pi}{q}$ for real positive values of q, to give a family of singular bulks $\mathcal{M}_q$, returning to $\mathcal{M}$ when $q \to 1$. We assume that this can be done without changing the topology of the bulk, or the defect within it. When q hits an integer, we can make a connection with the replica trick, but only if $\mathcal{M}_q$ can be lifted to a branched cover. That is, we require that there exists some $\tilde{\mathcal{M}}_q$, with boundary $\tilde{\mathcal{B}}_q$, and with a $\mathbb{Z}_q$ symmetry such that the quotient yields $\mathcal{M}_q$. Under this condition the action evaluated on $\tilde{\mathcal{M}}_q$ will legitimately give the q^{r-th} Rényi entropy (assuming $\tilde{\mathcal{M}}_q$ is the dominant saddle).

The examples tell us that this may happen at some q despite $\mathcal{E}$ violating the homology constraint. But taking $\mathcal{M}$ as the cross-cap geometry above, with $\mathcal{E}$ chosen to be empty, this lift to $\tilde{\mathcal{M}}_q$ can be constructed only at odd q, and does not exist for even q. This is because in the lift we must choose which copy of the boundary to end up on after passing through the cross-cap, say by rotating through k copies. But going through the cross-cap twice is homotopic to passing round the boundary circle, so must take us through one copy. This means we require $1 = 2k \mod q$, which has a solution if and only if q is odd, so no choice of k gives the correct cover on the boundary. We argue below that demanding the stronger condition that the lift must exist for all q so that $\mathcal{M}_q$ can be used to compute $S_{\mathcal{A}}^{(q)}$ at every positive integer, is equivalent to the homology constraint on $\mathcal{E}$.

This example generalizes to allow nonempty $\mathcal{E}$ as follows. Take the bulk $\mathcal{M}$ to be the same geometry as above, a solid torus with a neighborhood of a loop round the torus cut out and replaced with a Möbius strip times a circle. Consider now the region $\mathcal{A}$ as not a whole boundary circle, but an interval. Choose the surface $\mathcal{E}$ to be a curve joining the endpoints of $\mathcal{A}$, but passing round the non-contractible spatial circle on the side of the torus opposite of $\mathcal{A}$, and hence not homologous to it. Nonetheless, again at odd values of q this lifts to a smooth replica-symmetric bulk $\tilde{\mathcal{M}}_q$ in much the same way. When going round the surface $\mathcal{E}$, move from sheet to sheet as usual, but when passing through the crosscap pass to the $(\frac{q+1}{2})^{\text{th}}$ sheet relative to where you begin. In essence, the crosscap acts to replace the closed geodesic that would otherwise be required, but without the associated fixed point set or conical defect.

5.3.3 Further examples

There are several aspects of the above examples that may appear to be required to get the sort of homology violation we see: for example, a non-orientable spacetime, a distinction only between odd and even q, or the specific way a topological feature replaces a part of the spacetime where a fixed point 'should be'. We now briefly describe an example to show that none of these are necessary. Take $\mathcal{M}$ as a 3-dimensional ball, with boundary $\mathcal{B} = S^2$, cut out a ball at the center, and then identify antipodal points of the spherical edge of the resulting hole. This results in an $\mathbb{RP}^3$ with a boundary, an orientable spacetime. Take $\mathcal{A}$ as an interval on the boundary sphere, running between the poles perhaps, and choose $\mathcal{E}$ to pass through the nontrivial topology we have introduced in the centre. This can be

lifted to a replica-symmetric cover[7] to compute Rényi entropies for odd q as above, but not for even q. A generalization identifies the points on the sphere bounding the cut out region as in the construction of a Lens space $L(p;p')$ from identifications of the boundary of a 3-ball (the special case of antipodal points being $p=2$). The result of this is that a lift will exist for q co-prime to p and not otherwise.

We conclude this section with a different sort of example, which shows that it is possible to obtain a codimension-2 defect with deficit angle different from $\frac{2\pi}{q}$ from the quotient construction. The simplest example generalizes the computation of thermal entropy from BTZ; again we will not explicitly mention the trivial dependence on the spatial direction and focus on a constant spatial slice. The usual geometry then picks $\mathcal{E}$ as a point in the centre of this slice, and passing anticlockwise round this point takes you up one sheet in the cover. To generalize this, choose $\mathcal{E}$ as a collection of several points, and for each point choose some integer number of sheets to change by when passing round them anticlockwise. To be consistent with the boundary covering space, the only requirement is that these integers sum to one. But if the integer chosen for a point, say n is anything other than ± 1, the quotient of the covers thus constructed will result in a variety of conical deficit angles, depending on the greatest common factor r of n and q. Specifically, the covering space will contain r copies of the point, related by the $\mathbb{Z}_r$ subgroup of the $\mathbb{Z}_q$ replica symmetry, and the resulting defect angle after the quotient will be $\frac{2\pi r}{q}$.

For a specific example, take two points, choosing to increase by two sheets when passing round the first, and to decrease by one when passing round the second. The $q=4$ covering space has the topology of a torus with a single boundary (being the usual four glued copies of the boundary circle). This can be understood as a regular octagon, with top and bottom edges identified, as well as left and right edges; the remaining diagonal edges form the boundary circle. The replica symmetry acts by rotations by $\frac{\pi}{2}$. This leaves the obvious fixed point at the centre, which becomes the usual $\frac{\pi}{2}$ deficit after the quotient. The replica symmetry rotating by π has an additional two fixed points, being the centre of the top or bottom edge (identified with one another), and the centre of the left or right edge. Under the quotient, these two points become identified, and the result is a conical defect of angle π.

These sorts of examples can be thought of as many degenerate extremal surfaces lying on top of one another, which is very physically natural when viewing the defects as cosmic strings, and can be recovered as a limit of several of the usual defects coalescing. We would not expect these configurations to dominate in simple examples, particularly in the $q \to 1$ limit, but we can not rule out the possibility that these surfaces with multiplicity may be favourable enough to dominate in some more complicated geometry.

[7] An easy way to see this is to notice that $\mathcal{M} - \mathcal{E}$ deformation retracts to a circle, and going twice round this circle (or p times in the generalization) is homotopic to a boundary loop passing through $\mathcal{A}$ once.

5.4 A topological condition on q-Rényi saddles

In light of the example described in §5.3, we now wish to formulate a refined requirement regarding existence of covering spaces consistent with replica symmetry, which will be sufficiently strong to impose the expected topological constraint on the branching surface.

5.4.1 Boundary conditions for branched covers

The replica trick on the boundary involves defining the CFT on a branched cover $\tilde{\mathcal{B}}_q$. To compute the semiclassical Rényi entropy, one should evaluate the gravitational action on every bulk manifold $\tilde{\mathcal{M}}_q$ satisfying the equations of motion and the boundary condition $\partial\tilde{\mathcal{M}}_q = \tilde{\mathcal{B}}_q$, and choose the dominant saddle. This is prohibitively difficult, so we ignore the question of which saddle dominates, and furthermore we restrict to $\tilde{\mathcal{M}}_q$ which can be realized as a branched cover over the original space $\mathcal{M}$. That is, we assume that $\mathcal{M}_q = \tilde{\mathcal{M}}_q/\mathbb{Z}_q$ is homeomorphic to the original manifold $\mathcal{M}$. The boundary condition $\partial\tilde{\mathcal{M}}_q = \tilde{\mathcal{B}}_q$ imposes a topological condition on $\mathcal{M}_q - \mathbf{e}_q$, which we have seen above is *not* the homology constraint. First we need to understand this condition on $\mathcal{M}_q - \mathbf{e}_q$ more precisely.

Let $\mathcal{B}$ be the boundary, and $\mathcal{A}$ the region whose entanglement entropy we want to compute. Consider arbitrary closed loops in $\mathcal{B} - \partial\mathcal{A}$, i.e., loops which may intersect $\mathcal{A}$, but not $\partial\mathcal{A}$. Then there is a homomorphism $\phi : \pi_1(\mathcal{B} - \partial\mathcal{A}) \to \mathbb{Z}$ which counts the number of times the loop passes through $\mathcal{A}$. This map is defined to take into account the orientation of the loop relative to $\mathcal{A}$, so it computes the signed intersection number. Hence for any given q there is a map $\phi_q : \pi_1(\mathcal{B} - \partial\mathcal{A}) \to \mathbb{Z}_q$ which is just the previous map modulo q. The replicated boundary can be found by taking the cover $\tilde{\mathcal{B}}_q - \partial\mathcal{A}$ of $\mathcal{B} - \partial\mathcal{A}$, defined so that loops in $\ker\phi_q$ lift to closed loops in the covering space. This is just saying that ϕ_q counts which sheet we are on in the replicated CFT, and loops which intersect q times with $\mathcal{A}$ are to be identified with closed loops in the q-fold covering space in accordance with the replica trick.

Replica symmetry and the boundary condition $\partial\tilde{\mathcal{M}}_q = \tilde{\mathcal{B}}_q$ imply a similar construction for the bulk. First, it implies that there exists a map ψ_q keeping track of movements between sheets of the cover in the bulk. Second, it requires that this restricts to ϕ_q on the boundary $\partial(\mathcal{M}_q - \mathbf{e}_q) = \mathcal{B} - \partial\mathcal{A}$ to give the correct boundary conditions. This can be summarized as the existence of a bulk sheet-counting homomorphism ψ_q, factoring ϕ_q so that the following diagram commutes:

$$\begin{array}{ccc} \pi_1(\mathcal{B} - \partial\mathcal{A}) & \xrightarrow{\phi_q} & \mathbb{Z}_q \\ \downarrow{\scriptstyle i_*} & \nearrow{\scriptstyle \psi_q} & \\ \pi_1(\mathcal{M}_q - \mathbf{e}_q) & & \end{array} \tag{5.5}$$

Here i_* is the pushforward induced by the inclusion i of the boundary into the bulk. Below, we show more formally that the existence of such a ψ_q is equivalent to the existence of a

replica-symmetric covering space of the bulk.

5.4.2 Stronger criterion from the cosmic brane construction

The condition (5.5) is true of any branched cover satisfying the boundary conditions required by AdS/CFT. The example in §5.3 illustrates that with this condition alone, the singularity $\mathbf{e}_q$ is not necessarily homologous to $\mathcal{A}$.

But now, let us we restrict to branched covers of the original manifold $\mathcal{M}$ constructed following LM: choose a codimension-2 surface $\mathcal{E}$, and introduce a conical defect $\frac{2\pi}{q}$ at this surface. Can this be lifted to a branched cover $\tilde{\mathcal{M}}_q$ obeying the correct boundary conditions? In general, this is possible only if the boundary condition (5.5) holds for $\mathcal{M}-\mathcal{E}$. This criterion must be applied at each q separately, as in general it may be possible to find ψ_q for some values of q but not for others, as in the crosscap example in §5.3. If ψ_q exists lifting ϕ_q for all q, then there is a lift[8] $\psi : \pi_1(\mathcal{M}-\mathcal{E}) \to \mathbb{Z}$ of ϕ, meaning that the following diagram commutes:

$$\begin{array}{ccc} \pi_1(\mathcal{B}-\partial\mathcal{A}) & \xrightarrow{\phi} & \mathbb{Z} \\ \downarrow{\scriptstyle i_*} & \nearrow{\scriptstyle \psi} & \\ \pi_1(\mathcal{M}-\mathcal{E}) & & \end{array} \tag{5.6}$$

We will prove that (5.6) is equivalent to the homology condition for the RT surface.

To illustrate this, let us return to the crosscap example of §5.3 in this language. In this case, the starting point $\mathcal{M}$ is the Möbius strip times a circle. We choose the defect $\mathcal{E}$ to be the empty set, and try to construct branched covers at integer q. This can be done in a way satisfying the boundary condition (5.5) for odd q, as illustrated in Fig. 5.2, but for even q, there is no way to construct the branched cover. This can be seen very explicitly: we have $\phi_q(\Gamma) = 1$, where Γ is the loop going once around the boundary time circle. Assume there exists a lift ψ_q and require $\psi_q(\Gamma) = 1$. Now try to find a consistent way of choosing a value under ψ_q for the bulk loop γ that goes once through the cross-cap: since Γ is homotopic to 2γ, we require $1 \,(\mathrm{mod}\, q) = \psi_q(\Gamma) = \psi_q(2\gamma) = 2\,\psi_q(\gamma)$ which has a solution in $\mathbb{Z}_q$ if and only if q is odd. At the level of the maps ϕ and ψ of (5.6), this contradiction for even values of q manifests itself as follows: both fundamental groups in (5.6) are $\mathbb{Z}$ (ignoring the spatial circle), but one of them is generated by Γ, the other is generated by γ. This means ϕ is the identity (in particular $\phi(\Gamma) = 1$) but i_* is multiplication by 2, so clearly ψ does not exist.

5.4.3 Another look at the boundary condition

Above we motivated (5.5) from the AdS/CFT boundary conditions. We now discuss more formally how to show that, for a codimension-2 defect, this condition is equivalent to the existence of a topologically replica-symmetric covering space of the bulk.

[8] This is guaranteed if the relevant groups are finitely generated, which holds for compact manifolds with boundary.

We take here the perspective that a branched cover is defined as a covering space in the usual sense, where any point in the base space has a neighborhood lifting to q homeomorphic copies of itself, but after removing the branching surface.[9] Given this perspective, a replica-symmetric q-fold cover of $\mathcal{M}_q$, branched at $\mathbf{e}_q$, restricting correctly on the boundary, is defined by the following diagrams:

$$\begin{array}{ccc} \tilde{\mathcal{B}}_q - \partial\mathcal{A} & \xrightarrow{p} & \mathcal{B} - \partial\mathcal{A} \\ \downarrow{\scriptstyle \tilde{i}} & & \downarrow{\scriptstyle i} \\ \tilde{\mathcal{M}}_q - \mathbf{e}_q & \xrightarrow{P} & \mathcal{M}_q - \mathbf{e}_q \end{array} \qquad \begin{array}{ccccccc} \pi_1(\tilde{\mathcal{B}}_q - \partial\mathcal{A}) & \xrightarrow{p_*} & \pi_1(\mathcal{B} - \partial\mathcal{A}) & \xrightarrow{\phi_q} & \mathbb{Z}_q & \longrightarrow & 1 \\ \downarrow{\scriptstyle \tilde{i}_*} & & \downarrow{\scriptstyle i_*} & & & & \\ \pi_1(\tilde{\mathcal{M}}_q - \mathbf{e}_q) & \xrightarrow{P_*} & \pi_1(\mathcal{M}_q - \mathbf{e}_q) & \xrightarrow{\psi_q} & \mathbb{Z}_q & \longrightarrow & 1 \end{array} \tag{5.7}$$

The left hand diagram gives the covering space for boundary (top row) and bulk (bottom row), with respective covering maps p and P. Then i and $\tilde{i}$ are the injective inclusion maps of boundary into bulk, and the diagram commutes to implement the inclusions consistently. This implies immediately existence and commutativity of the left square of the diagram on the right. The additional maps are required to make the rows exact; the existence of ϕ_q and ψ_q is then equivalent to replica symmetry. This is because we require the q-fold cover to possess the replica symmetry, acting via a $\mathbb{Z}_q$ deck transformation group, acting cyclically. This implies that the covering space is *normal*, so that $P_*(\pi_1(\tilde{\mathcal{M}}_q - \mathbf{e}_q))$ is a normal subgroup of $\pi_1(\mathcal{M}_q - \mathbf{e}_q)$, and taking the quotient by that subgroup gives the $\mathbb{Z}_q$ deck transformation group. This shows that the map ψ exists, being the quotient map. The same argument also applies on the boundary. In particular $\ker(\phi_q) = \mathrm{im}(p_*)$ and $\ker(\psi_q) = \mathrm{im}(P_*)$. This shows that exactness is just a formal restatement of replica symmetry (in a topological sense) both at the boundary and in the bulk.

We prove in Appendix 5.A that the holographically natural consistency condition (5.5) holds if and only if $\mathcal{M}_q - \mathbf{e}_q$ is such that the diagrams (5.7) exist as described: existence of the map ψ_q to factor ϕ_q as $\psi_q \circ i_*$ implies that a cover can be constructed, and conversely existence of a cover implies that there is such a ψ_q.

5.5 Relation between topological consistency and homology constraint

Having dealt with the topological consistency condition required for the LM construction, we now turn to the central thesis of this work: "when is the homology condition satisfied?" We will give an overview of our general strategy first and exemplify it with the BTZ spacetime. In §5.5.3 we prove Theorem 1 which posits that we are guaranteed the homology constraint provided the topological consistency condition is met (and vice-versa).

[9] This works when the branching surface is codimension-2 as expected, since then there is a uniquely specified way to put back the branching surface in the cover. It fails when the branching surface is codimension-1, since extra information on what happens when passing through the surface must be imposed, but as this can only happen for even q in any case, this will not alter our main conclusions. Similarly, there are examples with fixed point sets of codimension greater than 2, but the local topology of such a quotient depends on q, so that there is no obvious way to fit them into a family at all values of q.

5.5.1 General strategy

The extension of the replica trick into the bulk [36] constructs a bulk geometry with boundary conditions given by the replicated field theory. At integer q this leads to a geometry which is a smooth q-fold covering space $\tilde{\mathcal{M}}_q$ of the original bulk geometry. The quotient $\mathcal{M}_q = \tilde{\mathcal{M}}_q/\mathbb{Z}_q$ is then just the original bulk geometry with a conical defect inserted along the codimension-2 fixed point set of $\mathbb{Z}_q$. By construction $\mathcal{M}_q$ lifts to a q-fold branched cover of $\mathcal{M}$ at every integer value of q and the topological consistency condition (5.6) holds. This topological consistency condition is the essential feature of the LM construction on which we will focus for the remainder of this section. In particular, we will no longer concern ourselves with covering spaces and simply restrict attention to topological properties of $\mathcal{M}$ and $\mathcal{E}$.

For concreteness, consider the following setup. We start with a bulk geometry $\mathcal{M}$, a boundary region $\mathcal{A}$ and a candidate extremal surface $\mathcal{E}$ to be used in computing entanglement entropy of $\mathcal{A}$. The goal of the present section is to first illustrate and then prove the following statement: *The topological consistency condition is satisfied for all q if and only if $\mathcal{A}$ is homologous (in the sense of* (5.2)*) to the surface $\mathcal{E}$.*[10] In order to decide whether $\mathcal{M} - \mathcal{E}$ satisfies the topological consistency condition for all q (and thus whether inserting a conical defect $\frac{2\pi}{q}$ along $\mathcal{E}$ would lift to a q-fold branched cover in the sense of LM at all integer q), consider the following maps:

- A *boundary sheet counting map* $\phi \in H^1(\mathcal{B} - \partial\mathcal{A})$: this is the map from (5.6) which counts how many times a boundary loop passes through $\mathcal{A}$. Although ϕ was originally defined on the homotopy group $\pi_1(\mathcal{B} - \partial\mathcal{A})$, the fact that it maps to an abelian group ($\mathbb{Z}$) makes its action on the first homology group well-defined. It is an element of the first cohomology group $H^1(\mathcal{B} - \partial\mathcal{A})$ of $\mathcal{B} - \partial\mathcal{A}$ with integer coefficients; such elements are just homomorphisms from boundary loops into $\mathbb{Z}$.

- A *local intersection map* $u \in H^2(\mathcal{M}, \mathcal{M} - \mathcal{E})$. By $H^2(\mathcal{M}, \mathcal{M} - \mathcal{E})$ we mean the second cohomology group of $\mathcal{M}$ relative to $\mathcal{M} - \mathcal{E}$ with integer coefficients. This u is defined on 2-dimensional surfaces (2-chains) in a neighborhood of $\mathcal{E}$ whose boundary is not a part of $\mathcal{E}$. Given such a 2-surface $\mathcal{D}$, the map u counts the (signed) number of intersections of $\mathcal{E}$ with $\mathcal{D}$.[11] The sign of an intersection is given by the relative orientation of $\mathcal{D}$ and $\mathcal{E}$.

Intuitively, ϕ (being defined by $\mathcal{A}$) encodes the boundary conditions for bulk geometries that can be lifted to replica symmetric covers. On the other hand, u (being defined by $\mathcal{E}$) carries topological information about the way such a bulk covering space would be branched. The topological consistency condition relates these two objects: namely, we can

[10] We postpone subtleties concerning orientation to §5.5.4.

[11] Note that this is only true in the absence of torsion cycles. If the bulk spacetime has torsion (in the topological sense), then the second cohomology group is not isomorphic to intersection counting homomorphisms. To illustrate the essence of our argument, we refrain from considering these subtleties at the present. We will however account for torsion in our proof of Theorem 1 in §5.5.3.

translate the topological consistency condition into a certain consistency of ϕ with respect to u. To this end, take any 2-surface $\mathcal{D}$ in the bulk, which is anchored outside of $\partial\mathcal{A}$ on the boundary, i.e., $\partial\mathcal{D} \subset \mathcal{B} - \partial\mathcal{A}$. Now compare the following two properties of $\mathcal{D}$:

- Use ϕ to count how many times the loop $\partial\mathcal{D}$ passes through $\mathcal{A}$ on the boundary (taking into account orientations).
- On the other hand, consider the parts of $\mathcal{D}$ which lie in a tubular neighborhood of $\mathcal{E}$ and compute their (signed) intersection number with $\mathcal{E}$ using u.

Our main statements can be summarized as follows:

1. The topological consistency condition means that these two countings have to agree for any 2-surface $\mathcal{D}$ anchored at the boundary.
2. The two countings agree for all 2-surfaces $\mathcal{D}$ anchored at the boundary if and only if $\mathcal{A}$ and $\mathcal{E}$ are homologous.

Illustration and proof of these points will be our main task. The rationale behind the first point, is roughly the following; it is a reformulation of the consistency condition at the level of the unreplicated bulk geometry. In fact, the topological consistency condition as formulated in (5.6) means exactly that ϕ can be lifted to a global sheet counting map $\psi \in H^1(\mathcal{M} - \mathcal{E})$ for bulk loops in a way that is consistent with u on loops that bound 2-surfaces which intersect $\mathcal{E}$. Roughly speaking, 2-surfaces intersecting $\mathcal{E}$ encode the same information about local branching near the intersection point as do small loops going around the intersection point. We thus require that u is inherited from the global sheet counting map ψ; technically we need $u = \delta\psi$ where δ is the coboundary map as defined in Appendix 5.B. This idea will be made precise in §5.5.3.

In the rest of this section we proceed as follows. We first illustrate the above ideas using the example of the BTZ black hole. Here we refrain from rigor and just give qualitative arguments for the validity of the argument. §5.5.3 then presents a rigorous and general proof that the topological consistency condition is satisfied if and only if $\mathcal{A}$ is homologous to $\mathcal{E}$ in an appropriate sense.

5.5.2 Example: BTZ black hole

Let us consider the single interval entanglement entropy of a CFT_2 on a spatial circle and at finite temperature, in the deconfined phase above the Hawking-Page transition. The Euclidean bulk $\mathcal{M}$ is topologically a solid torus. For illustration, consider the various extremal surfaces sketched in Fig. 5.3. In each case we want to study the relation between our topological consistency condition (5.6) and the homological properties of $\mathcal{E}$ compared to $\mathcal{A}$. The geodesics $\mathcal{E}^{(a)}$ and $\mathcal{E}^{(c)}$ are both homologous to $\mathcal{A}$ (i.e., in each case there exists an interpolating surface whose only boundaries are $\mathcal{A}$ and $\mathcal{E}^{(a,c)}$) and should therefore be allowed saddles in the RT formula. As we will see, both of them satisfy the topological consistency condition for all q and are thus consistent with the LM construction. The

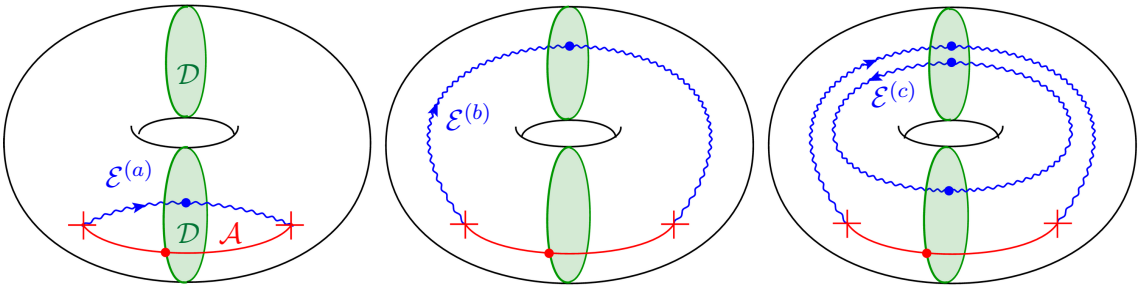


Figure 5.3: Three different candidates for bulk geodesics (blue) whose length may be considered to compute the entanglement entropy of the spatial boundary interval $\mathcal{A}$ (red). The green surfaces $\mathcal{D}$ illustrate the connection between boundary sheet counting (intersections with $\mathcal{A}$) and bulk intersections. Thick dots illustrate intersections. The bulk surfaces $\mathcal{E}^{(a)}$ and $\mathcal{E}^{(c)}$ are homologous to $\mathcal{A}$; which one of them is dominant depends on the length of $\mathcal{A}$. The surface $\mathcal{E}^{(b)}$ is not homologous to $\mathcal{A}$ and is also forbidden by our topological consistency condition. In this latter case, boundary intersections do not match bulk intersection numbers.

extremal surface $\mathcal{E}^{(b)}$, on the other hand, is not homologous to $\mathcal{A}$ which will manifest itself as a breakdown of the topological consistency condition (meaning that $\mathcal{E}^{(b)}$ can never arise as the fixed point set of a LM-type argument).

By inspection, it is evident from the examples in Fig. 5.3 how homology of $\mathcal{A}$ and $\mathcal{E}$ are tied to consistency of the boundary sheet counting map ϕ and the bulk intersection map u. For instance, in case (a) we have drawn two choices of surfaces $\mathcal{D}$ with boundary on $\mathcal{B}-\partial\mathcal{A}$ and clearly for both of them the number of intersections of $\partial\mathcal{D}$ with $\mathcal{A}$ matches the local intersection number of $\mathcal{D}$ with $\mathcal{E}^{(a)}$ in the bulk. The reader can easily verify that this matching indeed holds for any choice of 2-surface $\mathcal{D}$ anchored on $\mathcal{B}-\partial\mathcal{A}$. This consistency of ϕ and u allows for a global sheet counting map $\psi \in H^1(\mathcal{M}-\mathcal{E})$ which reduces to ϕ at the boundary.

Similar reasoning holds for case (c): The only novelty here is the fact that orientations need to be taken into account to properly count intersections. For instance, the intersection number of the rear disk with $\mathcal{E}^{(c)}$ is zero due to opposite orientations of $\mathcal{E}^{(c)}$ relative to $\mathcal{D}$ at the two intersection points. This matches the fact that the boundary of the rear disk does not intersect with $\mathcal{A}$.

Now consider case (b) where $\mathcal{E}^{(b)}$ indeed violates the homology constraint. As we can see, there now exist 2-surfaces $\mathcal{D}$ on which ϕ (intersections of $\partial\mathcal{D}$ with $\mathcal{A}$) takes a different value than u (intersection number of $\mathcal{D}$ with $\mathcal{E}^{(b)}$). Therefore the boundary sheet counting map ϕ does not lift to a global sheet counting map ψ consistent with u on loops which are boundaries of 2-surfaces that intersect $\mathcal{E}^{(b)}$. Such a lift not being possible means that covers of the bulk branched along $\mathcal{E}^{(b)}$ do not have the correct boundary conditions. This situation can therefore never arise from a LM construction.

At a pictorial level, this demonstrates our claims for the BTZ geometry: the homology constraint on $\mathcal{A}$ and $\mathcal{E}$ comes in conjunction with consistency of bulk intersection number

with $\mathcal{E}$ on the one hand, and boundary intersection number with $\mathcal{A}$ on the other hand. This consistency is equivalent to the topological consistency condition (5.6). We will now turn to a rigorous proof of this idea.

5.5.3 Topological consistency is equivalent to homology constraint

We begin with a reminder of the notation. The bulk $\mathcal{M}$ is a $d+1$ dimensional orientable manifold with boundary $\mathcal{B}$, $\mathcal{A}$ is a $d-1$ dimensional submanifold of $\mathcal{B}$ with boundary $\partial\mathcal{A}$, and $\mathcal{E}$ is a $d-1$ dimensional submanifold of $\mathcal{M}$, also with boundary $\partial\mathcal{A}$. It should be borne in mind that either or both of $\mathcal{A}$ and $\mathcal{E}$ may be disconnected. For this proof, we will focus on the case when the bulk is orientable; we will return to the extension to the non-orientable case in the next subsection. A review of the algebraic topology required for this section can be found in Appendix 5.B.

To construct the cover on the boundary, we have a map from $\pi_1(\mathcal{B}-\partial\mathcal{A})$ to $\mathbb{Z}$, counting signed intersections with $\mathcal{A}$. This is equivalent to a map $H_1(\mathcal{B}-\partial\mathcal{A})\to\mathbb{Z}$, since this homology group[12] is the abelianization of the fundamental group by the Hurewicz theorem. In turn, by the universal coefficient theorem, this homomorphism is equivalent to a cohomology class $\phi\in H^1(\mathcal{B}-\partial\mathcal{A})$.

This sheet counting map is equivalent to a homological description of the region $\mathcal{A}$. Regard $\mathcal{B}-\partial\mathcal{A}$ as the boundary minus a tubular open neighbourhood of $\partial\mathcal{A}$, so it becomes a compact manifold with boundary. We may then use Poincaré-Lefschetz duality (5.13) to send $\phi\in H^1(\mathcal{B}-\partial\mathcal{A})$ to $a\in H_{d-1}(\mathcal{B}-\partial\mathcal{A},\partial(\mathcal{B}-\partial\mathcal{A}))\approx H_{d-1}(\mathcal{B},\partial\mathcal{A})$. This is a homology class represented by an appropriately chosen orientation of $\mathcal{A}$.

The arguments above show that the bulk lifts to an appropriate branched cover at all q whenever there exists some extension of ϕ into the bulk. So we require some $\psi\in H^1(\mathcal{M}-\mathcal{E})$ satisfying $i^*\psi=\phi$, where $i:\mathcal{B}-\partial\mathcal{A}\to\mathcal{M}-\mathcal{E}$ is the inclusion of boundary into bulk, so that the restriction of ψ to the boundary is ϕ.

In addition to this, we need to know what happens when we traverse a small loop passing round the branching surface $\mathcal{E}$. We make this precise by considering the normal bundle of $\mathcal{E}$, which can be embedded in $\mathcal{M}$ as some tubular neighbourhood of $\mathcal{E}$. The fibres F of the bundle are copies of $\mathbb{R}^2$, with $\mathcal{E}$ lying at the origin, so the action of ψ passing round the surface is described by its restriction to $F-\{0\}$ on each fibre, $\psi_F\in H^1(F-\{0\})\approx\mathbb{Z}$. We would normally like this restriction to each fibre to be a generator of $H^1(F-\{0\})$, which implies that after taking the q-fold cover and the $\mathbb{Z}_q$ quotient there will be a $\frac{2\pi}{q}$ conical deect at every q. This need not hold in general, as examplified in §5.3, and the conical defect could be of a different angle depending on q in a more complicated way. This can be accounted for in the end by counting such a surface with the appropriate multiplicity in the homology computation, and by allowing this multiplicity we may assume that ψ_F is always a generator.

[12] All homology and cohomology groups are taken with coefficients in $\mathbb{Z}$, so we will not explicitly indicate this dependence until we later generalize to the non-orientable case.

So we would like necessary and sufficient conditions for the existence of some $\psi \in H^1(\mathcal{M} - \mathcal{E})$, such that

1. The restriction $i^*\psi$ to the boundary is $\phi \in H^1(\mathcal{B} - \partial\mathcal{A})$, and
2. The restriction ψ_F to the embedding into $\mathcal{M}$ of the fibres F of the normal bundle N of $\mathcal{E}$ is a generator of $H^1(F - \{0\})$.

The choice of ψ_F will in the course of the proof supply us with an appropriate orientation of $\mathcal{E}$, and thus an element $e \in H_{d-1}(\mathcal{E}, \partial\mathcal{A})$. We now have enough to state our main result.

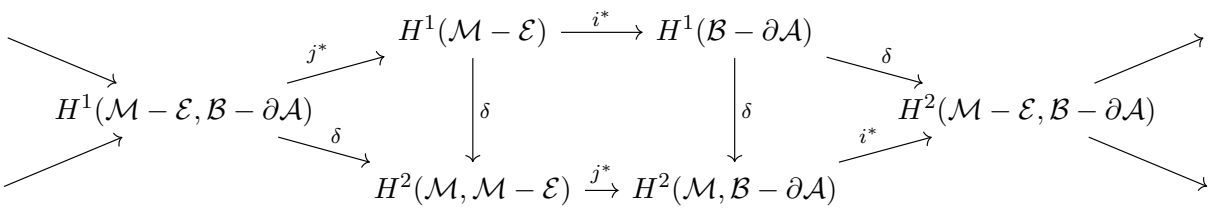
Theorem 1. *There exists some $\psi \in H^1(\mathcal{M} - \mathcal{E})$ which restricts to ϕ on the boundary, and which restricts to ψ_F on the fibres of N, if and only if $\mathcal{A}$ is homologous to $\mathcal{E}$, in the sense that the inclusions of e and a into $H_{d-1}(\mathcal{M}, \partial\mathcal{A})$ are equal.*

Proof. The choice of generators of $H^1(F-\{0\})$ on each fibre is equivalent to an orientation of each fibre, which can be regarded equivalently as a smooth choice of generators of the relative cohomology groups $H^2(F, F - \{0\})$, via a coboundary map. In each connected component of $\mathcal{E}$, making a choice on one fibre leaves no freedom on the others, which are determined uniquely from continuity. From this we can get a well-known object in the topology of vector bundles, a Thom class $u \in H^2(N, N - \mathcal{E})$, where $\mathcal{E}$ here sits at the zero section. This follows since if $\mathcal{E}$ is connected, then the restrictions $H^2(N, N - \mathcal{E}) \to H^2(F, F - \{0\})$ are in fact all isomorphisms. This is clear when restricted to a local trivialization of the bundle, and can be extended to the whole of the connected component by a standard procedure of gluing trivializations together one by one using a Mayer-Vietoris sequence. So our choices of ψ_F can be summarized by a choice of generator $u \in H^2(N, N - \mathcal{E})$ for each connected component of $\mathcal{E}$. See section 4.D of [143] for a discussion of the Thom class.

Now regarding N once again as a tubular neighbourhood[13] of $\mathcal{E}$ in $\mathcal{M}$, we can use an excision theorem to give an isomorphism $H^2(N, N - \mathcal{E}) \approx H^2(\mathcal{M}, \mathcal{M} - \mathcal{E})$ (excising $\mathcal{M} - \mathcal{E}$), so we may regard u instead as an element of $H^2(\mathcal{M}, \mathcal{M} - \mathcal{E})$. Roughly speaking, we may think of u as a map from two-dimensional chains with boundaries away from $\mathcal{E}$, which counts, with signs, the number of intersections with $\mathcal{E}$.

Now if $\mathcal{M}$ is oriented, the orientation of the normal bundle of $\mathcal{E}$ induces also an orientation of the tangent bundle of $\mathcal{E}$. This can be characterized by a relative homology class $e \in H_{d-1}(\mathcal{E}, \partial\mathcal{A})$.

Consider now the two cohomological long exact sequences, of the pair $(\mathcal{M} - \mathcal{E}, \mathcal{B} - \partial\mathcal{A})$, and of the triple $(\mathcal{M}, \mathcal{M} - \mathcal{E}, \mathcal{B} - \partial\mathcal{A})$, as explained in (5.10), which fit into a diagram:



[13] The existence of such a neighborhood is in fact a technical assumption, though this can be relaxed [144].

where the i^* and j^* maps are the relevant restrictions and extensions (dual to inclusion and quotient) respectively, and all the δs are various relative coboundary maps. The top and bottom rows are the usual exact sequences of relative cohomology, and it is straightforward to show that the diagram commutes at the level of cochains. It is then an easy exercise[14] to show that this induces another long exact sequence, the crucial part of which is

$$H^1(\mathcal{M}-\mathcal{E}) \xrightarrow{(i^*,\delta)} H^1(\mathcal{B}-\partial\mathcal{A}) \oplus H^2(\mathcal{M},\mathcal{M}-\mathcal{E}) \xrightarrow{\delta - j^*} H^2(\mathcal{M},\mathcal{B}-\partial\mathcal{A})$$

where the maps are the obvious ones from the diagram above, except that one factor in the second map has a minus sign.

Now an element $\psi \in H^1(\mathcal{M}-\partial\mathcal{A})$ restricting to ϕ on the boundary means that $i^*\psi = \phi$, and restricting to ψ_F on the fibres is equivalent to $\delta\psi = u$. So ψ with the desired properties exists iff $\phi \oplus u$ is in the image of the first map, which by exactness is the kernel of the second, so this is equivalent to $\delta\phi = j^*u$.

Finally, there is a generalization of Poincaré-Lefschetz duality (5.13) that by splitting $\mathcal{B}$ into a tubular neighborhood of $\partial\mathcal{A}$ and its complement gives us an isomorphism $H^2(\mathcal{M},\mathcal{B}-\partial\mathcal{A}) \approx H_{d-1}(\mathcal{M},\partial\mathcal{A})$, via cap product with the fundamental class $\mu \in H_{d+1}(\mathcal{M},\mathcal{B})$. The theorem follows by dualizing the equality $\delta\phi = j^*u$ under this, and showing that the duals of $\delta\phi$ and j^*u are repectively the inclusions of a and e into $H_{d-1}(\mathcal{M},\mathcal{B}-\partial\mathcal{A})$.

The dual of $\delta\phi$ is $\mu \frown \delta\phi$, which equals the inclusion of $\partial\mu \frown \phi$ into the bulk. This is indeed the correct thing, since $\partial\mu \frown \phi$ is just the boundary Poincaré-Lefschetz duality we originally used to relate a and ϕ.

The final part, computing the dual of j^*u, generalizes a well-known result in the case of closed manifolds [144]. The dual of j^*u is $\mu \frown j^*u = j_*\mu \frown u$, where $j_*\mu \in H_{d+1}(\mathcal{M},(\mathcal{M}-\partial\mathcal{A})\cup\mathcal{B})$. By restricting this cap product to a tubular neighborhood N of $\partial\mathcal{A}$, with inclusion map $k : N \to \mathcal{M}$, it is clear that the result must be in the image of k_*. Choosing N such that it retracts onto $\mathcal{E}$ by a retraction $r : N \to \mathcal{E}$ homotopic to the identity, it further follows that $k_* = i_*r_*$, so we get a cycle in the image of i_*. Working separately in each connected component of $\mathcal{E}$, this must be $n\,i_*e$ for some integer n. A local analysis in a trivialization of N, where the computations are simply in $\mathbb{R}^{d+1}$, shows that $n = 1$, so the dual of j^*u must in fact equal i_*e as required.

□

5.5.4 Non-orientable manifolds

In the case when the bulk is not orientable, the cohomological parts of the above proof go through unchanged, but the dualities used to make statements in terms of homology are not applicable. In fact, since $\mathcal{E}$, for instance, need not be orientable, there may not even be a homology class in $H_{d-1}(\mathcal{M},\partial\mathcal{A};\mathbb{Z})$ that represents it, so it becomes less clear how to state the result. This difficulty can be overcome by passing to $\mathbb{Z}_2$ coefficients, but since the correct cohomological argument requires integer coefficients, this gives an insufficiently

[14] Exercise 38 in chapter 2.2 of [143].

strong constraint on allowed $\mathcal{E}$. The correct thing is instead to use homology with local (or twisted) coefficients (see [143], section 3.H).

The generalization to local coefficients is performed by, roughly speaking, allowing the coefficients of the chain complexes to live in some module of the fundamental group of the spacetime. In our case, this will be the orientation class, being the module with action given by ± 1 depending on whether traversing a loop preserves or reverses orientation.

There is a generalized Poincaré-Lefschetz duality that holds for non-orientable spacetimes, between integer cohomology, and homology with coefficients in this module twisted by the orientation class. Since the cohomology we use is ambivalent to the presence or absence of orientation, nothing is altered until the very end, when we take this duality, and as such the homology statement will be in terms of local coefficients.

5.6 Discussion

The generalized gravitational entropy construction of LM, which is inspired by the replica trick in field theory, provides a derivation of the RT prescription for holographic entanglement entropy. We examined the conditions under which the LM construction guarantees, at the level of topology, that the extremal surface computing entanglement entropy satisfies the homology constraint in the bulk.

Suppose we have a bulk Euclidean geometry $\mathcal{M}$, a boundary region $\mathcal{A}$ and a minimal surface $\mathcal{E}$, a candidate for computing the entanglement entropy of $\mathcal{A}$. The local, dynamical part of the LM argument relates this to the replica trick as follows. Introduce a conical deficit at $\mathcal{E}$, with defect angle $\frac{2\pi}{q}$, increasing q away from unity, whilst changing the geometry to keep it on shell away from $\mathcal{E}$. Locally, near $\mathcal{E}$, we may choose a 'cigar' where there are polar coordinates parametrising a distance from $\mathcal{E}$ and an angle around it, with $\mathcal{E}$ itself at the origin. When q is an integer, this cigar with a conical deficit can be unwrapped by allowing the period of the angle to become q times larger, so we obtain a smooth space. The resulting boundary is *locally* the correct space with which to compute the q^{r-th} Rényi entropy, which makes contact with the replica trick.

It is less obvious that this picture works globally, since there is usually not a globally valid 'cigar', with a circle parameterised by a global 'Euclidean time' shrinking to zero size at $\mathcal{E}$. There is no obvious obstruction to introducing a conical defect at any minimal surface to give a continuous family of geometries, but at integer q, when the defects are locally unwrapped, it may not be possible to consistently extend the local pictures to the whole spacetime, or doing say may produce the Rényi entropy for the wrong region. For example, on a torus, the replicated geometries for an interval and its complement are locally the same, but differ globally by what happens on traversing the nontrivial cycles of the torus. This global picture, left implicit in LM, is crucial for understanding which choices for $\mathcal{E}$ give rise to the correct Rényi entropies.

Our main result addresses this global aspect, and can be phrased as follows: *A conical deficit introduced at a given extremal surface $\mathcal{E}$ comes from a quotient of a geometry whose*

boundary is the replicated space relevant for the qth Rényi entropy of region $\mathcal{A}$, for every positive integer q, if and only if $\mathcal{E}$ is homologous to $\mathcal{A}$. In particular, using the terminology explained in the introduction, the local interpretation of the LM assumptions suffices to enforce the homology constraint so long as it holds at all q and is consistent with the replica trick at integer q.

In this statement it is crucial that the construction of the bulk covering space works at all integer values of q. We gave an explicit example to show that branching along homology violating extremal surfaces can occur when the replica symmetric covering space exists for some value of q, but not for others. For the cosmic brane construction of LM [36] which starts from a singular spacetime $\mathcal{M}_q$, the homology constraint is implemented in the $q \to 1$ limit, iff for all integer q, $\mathcal{M}_q$ lifts to a non-singular q-fold replica symmetric branched cover (asymptoting to the replicated boundary geometry).

As such our analysis requires that the entanglement entropy is computed by an extremal surface which arises within a family of cosmic brane configurations as a limiting case. We do not require that the Rényi entropies themselves for integer values of q are computed by the elements of the same family. They could be computed (at large central charge) by other families, which would allow for Rényi phase transitions as in [50], or even isolated configurations. This also implies that the logarithmic negativity, which is obtained for bi-partitioning of a pure state by analytic continuation of the even-q Rényi entropies [145], could arise from a completely different family, for we only need a family of cosmic brane solutions for even integral q.

We emphasize here that the arguments were purely topological, with no recourse to dynamical information of the bulk gravitational theory. As such they hold for any choice of the boundary region $\mathcal{A}$ including disconnected ones. The statements regarding topological consistency, covering spaces, and the proof of Theorem 1 straightforwardly carry over to such situations. Of course, being agnostic of the dynamics has drawbacks, in that we will be unable to decide which of the homology respecting extremal surfaces actually computes the entanglement entropy. On the other hand, this gives us the advantage of being able to allow breaking replica symmetry away from integral values of q, as perhaps may be necessary for non Einstein-Hilbert gravitational dynamics [146].

We have assumed the bulk spacetime to be orientable in our discussion. For non-orientable spacetimes as indicated in §5.5.4 the natural generalization involves homology groups with twisted coefficients as opposed to integral homology used in our proof. This is relevant, for example, in the case of the $\mathbb{RP}^2$-geon spacetime [134], which is a pure state of a two-dimensional CFT (on a Klein bottle). In the bulk the spacetime has a horizon but a single asymptotic boundary (hence single-exterior black hole). In this case the homology condition is necessary to argue that the entanglement entropy for a region and its complement (on a spatial circle of the boundary) are the same as required by purity; see [3] for an analysis of the entanglement structure in the $\mathbb{RP}^2$-geon (on the $t = 0$ slice the $\mathbb{Z}_2$-valued homology group will suffice). It would be interesting to flesh out the details of the twisted homology constraint for more general physical examples.

Finally, let us also note that our analysis takes seriously the Euclidean LM construction. We only require the boundary region $\mathcal{A}$ to lie on the time-reflection symmetric surface. We are not a-priori guaranteed then that the extremal surface $\mathcal{E}$, and consequentially the homology surface $\mathcal{R}_{\mathcal{A}}$, lie at the moment of time-reflection symmetry in the bulk (see [3] for an example in thermal AdS where a sub-dominant extremal surface of this kind was described). If such surfaces give the dominant contribution to the holographic entanglement entropy, then various results in the holographic entanglement entropy literature such as strong-subadditivity [32, 40] and causality [40, 41] would have to be revisited.

5.A Equivalence of topological consistency conditions

We here prove two theorems which will make precise the equivalence between the existence of a replica symmetric covering space as in (5.7), and the existence of a sheet counting map in the bulk factoring the boundary sheet counting map by the inclusion (5.5).

To make the structure of the arguments clearer, we will prove something more general. We will take X to be a subspace of Y, and look for covering spaces $\tilde{X}$ and $\tilde{Y}$, with an inclusion map consistent with the inclusion of the base spaces. We additionally require the covers to possess certain symmetries, specifically that they are normal with deck transformation groups G and H respectively. The existence of such a cover will be shown to be equivalent to the existence of a commutative square of homomorphisms between $\pi_1(X)$, $\pi_1(Y)$, G and H. See section 1.3 of [143] for the necessary mathematical material.

To apply this in the context we require it, identify X and Y with boundary and bulk geometries respectively, both with the branching surfaces removed. Then G and H are both taken as $\mathbb{Z}_q$, so that we are requiring the covers to be replica symmetric. The square of homomorphisms will then be equivalent to (5.5).

The first part of the proof, Theorem 2, constructs the covering spaces given the appropriate homomorphisms. Theorem 3 shows the converse, that the appropriate homomorphisms exist starting from the covering spaces.

Theorem 2. *Let X and Y be path-connected, locally path-connected, semi-locally simply-connected spaces, and $i : X \to Y$ an injective inclusion map. Let $\phi : \pi_1(X) \to G$ and $\psi : \pi_1(Y) \to H$ be surjective homomorphisms, and $\rho : G \to H$ an injective homomorphism such that*

$$\begin{array}{ccc} \pi_1(X) & \xrightarrow{\phi} & G \\ \downarrow{\scriptstyle i_*} & & \downarrow{\scriptstyle \rho} \\ \pi_1(Y) & \xrightarrow{\psi} & H \end{array}$$

commutes. Then there are covering spaces $p : \tilde{X} \to X$, and $P : \tilde{Y} \to Y$, with deck

transformation groups G, H and an injective inclusion map $\tilde{i} : \tilde{X} \to \tilde{Y}$ such that

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{p} & X \\ \downarrow{\scriptstyle \tilde{i}} & & \downarrow{\scriptstyle i} \\ \tilde{Y} & \xrightarrow{P} & Y \end{array}$$

commutes.

In the case when $G = H (= \mathbb{Z}_q$, say), and the groups are finite, ρ is an isomorphism so we can identify the two groups. This is the way we use the theorem above. A similar remark applies for the second theorem.

Proof. Denote homotopy equivalence classes of loops by $[\cdot]$, and reversal of curves by $\bar{\cdot}$.

We construct $\tilde{X}$ as the set of equivalence classes $[\gamma]_{\tilde{X}}$ of curves $\gamma : [0,1] \to X$ starting at the basepoint in X, where $\gamma \sim_{\tilde{X}} \eta$ if (i) $\gamma(1) = \eta(1)$ and (ii) $[\gamma\bar{\eta}] \in \ker(\phi)$. The projection map is $p : [\gamma]_{\tilde{X}} \mapsto \gamma(1)$, which is clearly well-defined. This is the standard construction of the covering space with fundamental group $\ker(\phi)$. The deck transformation group is $\pi_1(X)/\ker(\phi) \cong G$, since ϕ is surjective. Similar statements apply to construct $\tilde{Y}$.

It remains only to construct the inclusion map $\tilde{i}$, by $\tilde{i} : [\gamma]_{\tilde{X}} \mapsto [i\gamma]_{\tilde{Y}}$. This is well defined, since if $\gamma \sim_{\tilde{X}} \eta$, then (i) $i\gamma(1) = i\eta(1)$, and (ii) $\psi[i\gamma\overline{i\eta}] = \psi i_*[\gamma\bar{\eta}] = \rho\phi[\gamma\bar{\eta}] = 1$ since $[\gamma\bar{\eta}] \in \ker\phi$, so $[i\gamma\overline{i\eta}] \in \ker\psi$. Finally, $\tilde{i}$ is injective, since if $[i\gamma]_{\tilde{Y}} = [i\eta]_{\tilde{Y}}$, then (i) $i\gamma(1) = i\eta(1) \Rightarrow \gamma(1) = \eta(1)$ by injectivity of i, and (ii) $1 = \psi[i\gamma\overline{i\eta}] = \psi i_*[\gamma\bar{\eta}] = \rho\phi[\gamma\bar{\eta}] \Rightarrow [\gamma\bar{\eta}] \in \ker(\phi)$, where we have used injectivity of ρ in the last step. □

Theorem 3. *Suppose that there are normal covering spaces $p : \tilde{X} \to X$, and $P : \tilde{Y} \to Y$, with deck transformation groups G, H and injective inclusion maps $i : X \to Y$ and $\tilde{i} : \tilde{X} \to \tilde{Y}$ such that*

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{p} & X \\ \downarrow{\scriptstyle \tilde{i}} & & \downarrow{\scriptstyle i} \\ \tilde{Y} & \xrightarrow{P} & Y \end{array}$$

commutes. Then an injective homomorphism $\rho : G \to H$ exists, such that the following diagram commutes:

$$\begin{array}{ccccccc} \pi_1(\tilde{X}) & \xrightarrow{p_*} & \pi_1(X) & \xrightarrow{\phi} & G & \longrightarrow & 1 \\ \downarrow{\scriptstyle \tilde{i}_*} & & \downarrow{\scriptstyle i_*} & & \vdots{\scriptstyle \rho} & & \\ \pi_1(\tilde{Y}) & \xrightarrow{P_*} & \pi_1(Y) & \xrightarrow{\psi} & H & \longrightarrow & 1 \end{array}$$

The maps ϕ and ψ here are the quotient maps by the fundamental groups of the covering spaces (using the fact that the coverings are normal), which implies that the rows are exact.

Proof. Let $g \in G$. Since ϕ is surjective, there is some loop γ in X so that $\phi([\gamma]) = g$. Define $\rho(g) = \psi i_*[\gamma]$. We must check that this is well defined, so suppose that η is another loop in X with $\phi([\eta]) = g$. Then $\phi([\gamma\bar{\eta}]) = \phi([\gamma][\eta]^{-1}) = 1$. This means that $\gamma\bar{\eta}$ lifts to a loop $\tilde{\gamma}$ in $\tilde{X}$, by exactness of the top row. Now $i_*[\gamma\bar{\eta}] = i_*p_*[\tilde{\gamma}] = P_*\tilde{i}_*[\tilde{\gamma}] \in \ker(\psi)$ by

exactness of the bottom row, so $\psi i_*[\gamma] = \psi i_*[\eta]$. It's also clear from the definition that $\psi i_* = \rho\phi$.

Finally, we check injectivity of ρ, so let $g \in G$ be distinct from the identity. Pick a loop γ with $\phi[\gamma] = g$, and lift it to a curve $\tilde{\gamma}$ in $\tilde{X}$. Since g is not the identity, this is not a loop, and since $\tilde{i}$ is injective, the curve $\tilde{i}\tilde{\gamma}$ in $\tilde{Y}$ is also not a loop. By the commutativity, this is the same curve as the lift of $i\gamma$, which implies that $[i\gamma] \notin \mathrm{im}(P_*) = \ker(\psi)$. So $\rho(g) = \psi([i\gamma])$ is not the identity, and ρ is injective. □

5.B Review of algebraic topology

This appendix reviews the algebraic topology required for the main result of the paper. See [143] for more detailed discussions.

The space of singular n-chains $C_n(X)$ of a topological space X is the free abelian group with basis singular n-simplices, which are maps from the standard n-simplex $\Delta^n = \{(t_0, \ldots, t_n) \in \mathbb{R}^{n+1} | \sum_i t_i = 1, t_i \geq 0\}$ to X. The boundary maps $\partial_n : C_n(X) \to C_{n-1}(X)$ act on chains σ by a sum of restrictions of σ to its $n+1$ faces, with apropriate signs. A cycle is defined as a chain with zero boundary, in the kernel of ∂, and a boundary is a chain in the image of ∂. Since $\partial_n \partial_{n+1} = 0$, all boundaries are cycles, and this means we can define the singular homology groups as cycles modulo boundaries: $H_n(X) = \ker \partial_n / \operatorname{im} \partial_{n+1}$.

A generalization of this is a relative homology group, where we ignore what goes on in some part of the space. Let A be a subset of the space X; the space of relative chains is the space of chains in X modulo chains in A, $C_n(X, A) = C_n(X)/C_n(A)$. The usual boundary map continues to be well-defined between relative chains, so we may again take cycles modulo boundaries to get the relative homology groups $H_n(X, A)$. Cycles here may therefore have a boundary in the subspace A, and it is trivial if it is homologous to a cycle completely within A. Often the space A will be the complement of some set, $A = X - B$, when $H_n(X, X - B)$ is sometimes referred to as 'homology at B', since it ignores what happens to chains except at the subspace B.

Since the relative homology ignores what goes on in some subspace, nothing is lost by removing some part of that subspace. This is made precise by the excision theorem, which states that if $Z \subseteq A \subseteq X$, and the closure of Z is contained in the interior of A, then $H_n(X, A)$ is isomorphic to $H_n(X - Z, A - Z)$.

Of central importance is the fact that relative homology groups fit into a long exact sequence:

$$\cdots \xrightarrow{\partial} H_n(A) \xrightarrow{i_*} H_n(X) \xrightarrow{j_*} H_n(X, A) \xrightarrow{\partial} H_{n-1}(A) \xrightarrow{i_*} \cdots \tag{5.8}$$

The maps i_* here are inclusions of cycles in A into X, and the maps j_* are the quotients by chains in A to get to relative homology. Finally, the maps ∂ are boundary maps: any relative n-cycle must have boundary contained in A by definition, and this gives a homology class in $H_{n-1}(A)$. Exactness is geometrically very intuitive: for example, a cycle in $H_n(A)$ gives zero when it is included into $H_n(X)$ if and only if it is the boundary of some chain

in X, which is the case if and only if it is in the image of the boundary map ∂ from the relative homology group. So $\ker i_* = \operatorname{im} \partial$.

This can be slightly generalized to the long exact sequence of the triple (X, A, B), where $B \subseteq A \subseteq X$, by doing everything relative to the smallest subspace B:

$$\cdots \xrightarrow{\partial} H_n(A, B) \xrightarrow{i_*} H_n(X, B) \xrightarrow{j_*} H_n(X, A) \xrightarrow{\partial} H_{n-1}(A, B) \xrightarrow{i_*} \cdots \tag{5.9}$$

Now most of the work we will do will be not in terms of homology, but cohomology, to which all of the above carries over. To pass to cohomology, all the relevant spaces and maps should be dualized: we consider spaces of cochains $C^n(X) = \hom(C_n(X), \mathbb{Z})$, joined by coboundary maps $\delta_n : C^n(X) \to C^{n+1}(X)$ dual to the boundary maps, so $\delta f(x) = f(\partial x)$. The kernel of δ gives the cochains, and the image the coboundaries; the cohomology groups are cochains mod coboundaries $H^n(X) = \ker \delta_n / \operatorname{im} \delta_{n-1}$.

The constructions for relative homology groups carry over to cohomology, by dualizing the relative chain complex. In particular the excision theorem is the same in cohomology, and the long exact sequences of relative homology groups dualize, for example for the triple (X, A, B) we have

$$\cdots \xrightarrow{i^*} H^{n-1}(A, B) \xrightarrow{\delta} H^n(X, A) \xrightarrow{j^*} H^n(X, B) \xrightarrow{i^*} H^n(A, B) \xrightarrow{\delta} \cdots \tag{5.10}$$

While most of the argument is phrased in terms of cohomology, in the end we want to translate the result into homology language. The crucial tool to do this is a generalization of Poincaré duality. Geometrically, these dualities can be roughly thought of as taking a submanifold to a function on submanifolds of complementary dimension, counting with signs the number of intersections between the submanifolds.

This requires the cap product between chains and cochains, defined by

$$\begin{aligned} \frown\colon C_k(X) \times C^l(X) &\longrightarrow C_{k-l}(X) \\ \sigma \frown \varphi &= \varphi(\sigma_{[v_0,\ldots,v_l]}) \cdot \sigma_{[v_l,\ldots v_k]} \end{aligned} \tag{5.11}$$

where $k \geq l$ and $\sigma_{[v_0,\ldots,v_l]}$ denotes the restriction of σ to the simplex spanned by the vertices $v_0, \ldots, v_l$. This map induces a cap product between homology and cohomology classes.

A connected closed orientable n-manifold X has top homology group $H_n(X)$ isomorphic to $\mathbb{Z}$, and an orientation of X is equivalent to a generator of this group, a *fundamental class* $[X]$. The Poincaré duality map is given by the cap product with the fundamental class $[X] \frown : H^k(X) \to H_{n-k}(X)$, which is in fact an isomorphism between complementary homology and cohomology groups.

This duality generalizes to manifolds with boundary. Crucial to this is a relative version of the cap product. Using the same definitions as before, it can be checked that the product on chains and cochains induces a product on relative (co)homology groups

$$\frown\colon H_k(X, A \cup B) \times H^l(X, A) \longrightarrow H_{k-l}(X, B) \tag{5.12}$$

where A and B are open sets in X. The duality theorem we need applies this in the case of an oriented n-manifold X with boundary ∂X, where the boundary is decomposed as $\partial X = A \cup B$, where A and B are $(n-1)$-dimensional manifolds with common boundary $A \cap B = \partial A = \partial B$. An orientation of X is defined in this case by a generator of $H_n(X, \partial X)$, which as before gives us isomorphisms

$$[X] \frown : H^k(X, A) \xrightarrow{\approx} H_{n-k}(X, B). \tag{5.13}$$

This includes the special cases where $A = \emptyset$ and $B = \partial X$, or vice-versa, known as Poincaré-Lefschetz duality.

Chapter 6

Outlook

As is often the case with progress towards understanding, the results of many of these investigations have thrown up more questions. With this in mind, I will conclude this thesis with some of the outstanding problems offering promise for fresh understanding, some questions that now seem within reach, and some more ambitious ideas for the future.

The first major aim is a complete understanding of the emergence of entropy as area from holography. Whilst the derivation of the Ryu-Takayanagi formula by Lewkowycz and Maldacena [36] may seem to provide a complete explanation in time-reflection symmetric cases, it leaves several outstanding subtleties.

The arguments required to reproduce the homology constraint in chapter 5 make some of these subtleties clear. The replica trick relies on computing Rényi entropies for integers q, and in the classical, saddle point approximation the Euclidean path integral can be evaluated from the on-shell action of the appropriate geometry. If there is one such possible saddle point for each q, it seems very natural that the answer can be analytically continued to find the entanglement entropy. Even if there are several saddle point geometries at each q, which can be fitted into distinct families, we can still proceed by analytically continuing each family separately, and taking the dominant answer. This remains true if different families dominate for different q: presumably this indicates a first order phase transition for Rényi entropies at some q, which would be smoothed out at finite N by quantum corrections. This could even happen at some q between 1 and 2, so that the integer Rényis are never computed by the same family of saddle points as the entanglement entropy.

What is less clear is how to interpret a case where a saddle point exists at some values of q, with no natural counterpart at others. We have seen some examples of how this may occur in chapter 5. A different mechanism for such a phenomenon would be to spontaneously break the cyclic $\mathbb{Z}_q$ replica symmetry in the bulk. There are two distinct ways this could manifest. Firstly, it could happen dynamically by the values of some fields breaking the symmetry; it is possible that in this case the analytic continuation could be fixed by some protocol, such as suggested in [146]. The second, more dramatic way, is for the topology to fail to respect the symmetry. It would appear in this case to be impossible to fit such saddle points into a family in such a way that an analytic continuation is

admissible.

It may be the case that all such 'nasty' saddle points are always subdominant, in which case maybe it is legitimate to simply ignore them. But if we are to take seriously the possibility that a continuation of a saddle point, though subdominant at every integer $q \geq 2$, may give the leading contribution at $q = 1$, this is still perhaps too fast. It is much worse if such a saddle may dominate: the Rényi entropies should at least be continuous in q, even in the classical limit, so how is one to evaluate them in the neighbourhood of an integer, when there is no way to continue the answer away from that value? And even if such cases are isolated, if we can't trust the analysis for these values of q, why should we believe that we nonetheless obtain the correct answer when we take $q \to 1$?

A related issue is the rôle of the time-reflection symmetry. The only crucial part this plays in the Lewkowycz-Maldacena argument is for the translation to Euclidean signature, and is unnecessary for what follows. As described in section 4.4, this would allow for minimal surfaces that spontaneously break the reflection symmetry of the Euclidean spacetime, and if such surfaces have the least area, and one takes at face value the Euclidean quantum gravity argument, there is nothing to stop them giving the dominant contribution to the path integral. From a purely Euclidean point of view, this is not a problem, but it is more troubling if one would like to translate the result back to the Lorentzian spacetime, since there will be no extremal surface which gives the correct entanglement entropy there. This raises the question of 'complexified extremal surfaces' more generally for HRT, where it is unclear even how one would go about defining such an object in general. If these are ever important, the existing geometric proofs of strong subadditivity, and other consistency checks such as causality, become invalid.

The situation becomes yet more problematic if we would like to consider quantum corrections to the entanglement entropy. One might expect subdominant saddle points to contribute nonperturbative 'instanton' corrections to the Rényi entropies, but given the difficulty in general of the analytic continuation it is unclear how this works. Perhaps more dangerously, it would appear that when considering the contributions of larger extremal surfaces, the $q \to 1$ limit to retrieve the entanglement entropy does not commute with the classical limit, leading to an apparent contradiction.

It may even be that the classical limit is not valid for q different from an integer at all. If the Lewkowycz-Maldacena argument is taken completely seriously, it is natural to propose that the Rényi entropies are computable from actions of Euclidean geometries with conical defects. But it is not clear whether the saddle points obtained with these boundary conditions have bounded curvature as the defect is approached. If not, it is possible that quantum (and stringy) corrections are not suppressed, even in the classical (and strong coupling) limit.

If we believe that the relationship between geometry and area underlies some deep truth, then these sorts of phenomena should never be important, and may simply be technical artefacts appearing from the specific methods of calculation. It would be useful to check some more examples, to see whether the naïve answer remains valid, and if it

is robust, perhaps it can be proven (though it seems very difficult) that the dominant saddles are always 'simple' in an appropriate sense. All of this provides clear motivation for some alternative derivation of the Ryu-Takayanagi prescription, avoiding the replica trick, or perhaps performing an analytic continuation in some other way, that avoids all these subtleties.

Perhaps it is premature to attack these questions when there is a more pressing issue outstanding, of generalising the proof of RT in time-reflection symmetric cases to HRT in more general dynamical geometries. One natural approach to this, in some states, is to use a mixed Euclidean and Lorentzian path integral, Euclidean pieces to prepare a state, and Lorentzian to evolve it in real time, to generalise the replica trick. The aim is then to use a Hartle-Hawking type procedure to extend this to a bulk spacetime, where arguments similar to Lewkowycz-Maldacena can be applied. If this can be made sense of, perhaps this will elucidate the rôle of causality, particularly in the topological constraints on extremal surfaces.

Three dimensional gravity offers the best testing ground for many of these ideas, and the most hope for progress. Already, the matching [34, 35] between field theory and gravity in vacuum gives a microscopic picture for the CFT origin of different geodesics, related to OPE channels of correlation functions of twist operators. This was extended to certain phases[1] in the thermal state in [42], including some quantum corrections. The simple algebraic answer given in chapter 4 for lengths of geodesics in more general pure gravity geometries is suggestive that it may be possible to generalise to higher genus, and perhaps to get a microscopic interpretation for the homology constraint from this.

A more practical question is to find robust, efficient and generally applicable techniques for finding extremal surfaces relevant for HRT. This would be very useful for investigating the dynamics of entanglement in holographic theories in more than two spacetime dimensions, in contexts with insufficient symmetries to reduce the extremal surface equations to ODEs. Most, if not all, methods for finding minimal surfaces explicitly rely on the fact that the surface minimises some functional, and are unstable if applied to the covariant, extremal case. One idea is to use 'mountain pass', or 'minimax' algorithms, closely related to the maximin characterisation of extremal surfaces [40]. It would be particularly nice if this could be done in a geometrically natural way, perhaps analogous to the mean curvature flow of section 2.A.

A related connection between geometry entanglement is embodied by the 'ER=EPR' proposal [147], which posits that quantum entanglement is essentially synonymous with connectedness of spacetime. It is very natural that two black holes joined by a nontraversable wormhole are entangled (so 'ER $\implies$ EPR'), but it is less clear that the converse, that entangled black holes are necessarily joined by an Einstein-Rosen bridge, is a useful model. The tools of RT and HRT, in particular in wormhole geometries, are certainly useful for exploring this idea further.

More generally, a fuller understanding of the relationship between area and entropy

[1] It does not include the 'plateau' phase of chapter 2

seems necessary to make sense of how a dynamical geometry emerges from a quantum theory, and how that notion of geometry should be modified away from a classical limit. One might hope that this may offer clues to universal features of quantum gravity, without requiring input from a complete, high energy, nonperturbative description. Perhaps the connections between entanglement and geometry can begin to unlock the long-standing mysteries of quantum gravity.

Bibliography

[1] V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, *Holographic entanglement plateaux*, *JHEP* **1308** (2013) 092, [arXiv:1306.4004].

[2] V. E. Hubeny and H. Maxfield, *Holographic probes of collapsing black holes*, *JHEP* **1403** (2014) 097, [arXiv:1312.6887].

[3] H. Maxfield, *Entanglement entropy in three dimensional gravity*, arXiv:1412.0687.

[4] F. M. Haehl, T. Hartman, D. Marolf, H. Maxfield, and M. Rangamani, *Topological aspects of generalized gravitational entropy*, arXiv:1412.7561.

[5] J. Bell, *On the Einstein-Podolsky-Rosen paradox*, *Physics* **1** (1964) 195–200.

[6] T. Ritz, P. Thalau, J. B. Phillips, R. Wiltschko, and W. Wiltschko, *Resonance effects indicate a radical-pair mechanism for avian magnetic compass*, *Nature* **429** (2004), no. 6988 177–180.

[7] E. Schrödinger, *Discussion of probability relations between separated systems*, in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, pp. 555–563, Cambridge Univ Press, 1935.

[8] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum mechanical description of physical reality be considered complete?*, *Phys.Rev.* **47** (1935) 777–780.

[9] D. N. Page, *Average entropy of a subsystem*, *Phys.Rev.Lett.* **71** (1993) 1291–1294, [gr-qc/9305007].

[10] H. Araki and E. Lieb, *Entropy inequalities*, *Commun.Math.Phys.* **18** (1970) 160–170.

[11] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, *Area laws in quantum systems: mutual information and correlations*, *Physical review letters* **100** (2008), no. 7 070502.

[12] M. Srednicki, *Entropy and area*, *Phys.Rev.Lett.* **71** (1993) 666–669, [hep-th/9303048].

[13] H. Liu and M. Mezei, *A Refinement of entanglement entropy and the number of degrees of freedom*, *JHEP* **1304** (2013) 162, [arXiv:1202.2070].

[14] A. Kitaev and J. Preskill, *Topological entanglement entropy*, *Phys.Rev.Lett.* **96** (2006) 110404, [hep-th/0510092].

[15] M. Levin and X.-G. Wen, *Detecting Topological Order in a Ground State Wave Function*, *Phys.Rev.Lett.* **96** (2006) 110405.

[16] H. Casini and M. Huerta, *A Finite entanglement entropy and the c-theorem*, *Phys.Lett.* **B600** (2004) 142–150, [hep-th/0405111].

[17] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *J.Phys.* **A42** (2009) 504005, [arXiv:0905.4013].

[18] H. Casini, M. Huerta, and R. C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **1105** (2011) 036, [arXiv:1102.0440].

[19] J. Bekenstein, *Black holes and the second law*, *Lett.Nuovo Cim.* **4** (1972) 737–740.

[20] S. Hawking, *Gravitational radiation from colliding black holes*, *Phys.Rev.Lett.* **26** (1971) 1344–1346.

[21] S. Hawking, *Particle Creation by Black Holes*, *Commun.Math.Phys.* **43** (1975) 199–220.

[22] R. Bousso, *The Holographic principle*, *Rev.Mod.Phys.* **74** (2002) 825–874, [hep-th/0203101].

[23] D. Bigatti and L. Susskind, *TASI lectures on the holographic principle*, hep-th/0002044.

[24] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int.J.Theor.Phys.* **38** (1999) 1113–1133, [hep-th/9711200].

[25] S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys.Lett.* **B428** (1998) 105–114, [hep-th/9802109].

[26] E. Witten, *Anti-de Sitter space, thermal phase transition, and confinement in gauge theories*, *Adv.Theor.Math.Phys.* **2** (1998) 505–532, [hep-th/9803131].

[27] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl.Phys.* **B72** (1974) 461.

[28] G. T. Horowitz and J. Polchinski, *Gauge/gravity duality*, gr-qc/0602037.

[29] J. McGreevy, *Holographic duality with a view toward many-body physics*, *Adv.High Energy Phys.* **2010** (2010) 723105, [arXiv:0909.0518].

[30] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys.Rev.Lett.* **96** (2006) 181602, [hep-th/0603001].

[31] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **0608** (2006) 045, [hep-th/0605073].

[32] M. Headrick and T. Takayanagi, *A Holographic proof of the strong subadditivity of entanglement entropy*, *Phys.Rev.* **D76** (2007) 106013, [arXiv:0704.3719].

[33] P. Hayden, M. Headrick, and A. Maloney, *Holographic Mutual Information is Monogamous*, *Phys.Rev.* **D87** (2013), no. 4 046003, [arXiv:1107.2940].

[34] T. Hartman, *Entanglement Entropy at Large Central Charge*, arXiv:1303.6955.

[35] T. Faulkner, *The Entanglement Renyi Entropies of Disjoint Intervals in AdS/CFT*, arXiv:1303.7221.

[36] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **1308** (2013) 090, [arXiv:1304.4926].

[37] X. Dong, *Holographic Entanglement Entropy for General Higher Derivative Gravity*, *JHEP* **1401** (2014) 044, [arXiv:1310.5713].

[38] J. Camps, *Generalized entropy and higher derivative Gravity*, *JHEP* **1403** (2014) 070, [arXiv:1310.6659].

[39] V. E. Hubeny, M. Rangamani, and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **0707** (2007) 062, [arXiv:0705.0016].

[40] A. C. Wall, *Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy*, *Class.Quant.Grav.* **31** (2014), no. 22 225007, [arXiv:1211.3494].

[41] M. Headrick, V. E. Hubeny, A. Lawrence, and M. Rangamani, *Causality & holographic entanglement entropy*, arXiv:1408.6300.

[42] T. Barrella, X. Dong, S. A. Hartnoll, and V. L. Martin, *Holographic entanglement beyond classical gravity*, *JHEP* **1309** (2013) 109, [arXiv:1306.4682].

[43] T. Faulkner, A. Lewkowycz, and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **1311** (2013) 074, [arXiv:1307.2892].

[44] N. Engelhardt and A. C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **1501** (2015) 073, [arXiv:1408.3203].

[45] D. V. Fursaev, *Proof of the holographic formula for entanglement entropy*, *JHEP* **0609** (2006) 018, [hep-th/0606184].

[46] M. Headrick, *Entanglement Renyi entropies in holographic theories*, *Phys.Rev.* **D82** (2010) 126010, [arXiv:1006.0047].

[47] T. Nishioka and T. Takayanagi, *AdS Bubbles, Entropy and Closed String Tachyons*, *JHEP* **0701** (2007) 090, [hep-th/0611035].

[48] I. R. Klebanov, D. Kutasov, and A. Murugan, *Entanglement as a probe of confinement*, *Nucl.Phys.* **B796** (2008) 274–293, [arXiv:0709.2140].

[49] T. Albash and C. V. Johnson, *Holographic Studies of Entanglement Entropy in Superconductors*, *JHEP* **1205** (2012) 079, [arXiv:1202.2605].

[50] A. Belin, A. Maloney, and S. Matsuura, *Holographic Phases of Renyi Entropies*, *JHEP* **1312** (2013) 050, [arXiv:1306.2640].

[51] T. Azeyanagi, T. Nishioka, and T. Takayanagi, *Near Extremal Black Hole Entropy as Entanglement Entropy via AdS(2)/CFT(1)*, *Phys.Rev.* **D77** (2008) 064005, [arXiv:0710.2956].

[52] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, *Relative Entropy and Holography*, *JHEP* **1308** (2013) 060, [arXiv:1305.3182].

[53] B. Freivogel, V. E. Hubeny, A. Maloney, R. C. Myers, M. Rangamani, et al., *Inflation in AdS/CFT*, *JHEP* **0603** (2006) 007, [hep-th/0510046].

[54] V. E. Hubeny and M. Rangamani, *Causal Holographic Information*, *JHEP* **1206** (2012) 114, [arXiv:1204.1698].

[55] V. E. Hubeny, M. Rangamani, and E. Tonni, *Global properties of causal wedges in asymptotically AdS spacetimes (to appear)*, .

[56] P. Calabrese and J. L. Cardy, *Entanglement entropy and quantum field theory*, *J.Stat.Mech.* **0406** (2004) P06002, [hep-th/0405152].

[57] N. Ogawa, T. Takayanagi, and T. Ugajin, *Holographic Fermi Surfaces and Entanglement Entropy*, *JHEP* **1201** (2012) 125, [arXiv:1111.1023].

[58] C. P. Herzog and M. Spillane, *Tracing Through Scalar Entanglement*, *Phys.Rev.* **D87** (2013) 025012, [arXiv:1209.6368].

[59] M. M. Caldarelli, O. J. Dias, R. Emparan, and D. Klemm, *Black Holes as Lumps of Fluid*, *JHEP* **0904** (2009) 024, [arXiv:0811.2381].

[60] R. Emparan, R. Suzuki, and K. Tanabe, *The large D limit of General Relativity*, *JHEP* **1306** (2013) 009, [arXiv:1302.6382].

[61] V. E. Hubeny, *Extremal surfaces as bulk probes in AdS/CFT*, *JHEP* **1207** (2012) 093, [arXiv:1203.1044].

[62] J. Abajo-Arrastia, J. Aparicio, and E. Lopez, *Holographic Evolution of Entanglement Entropy*, *JHEP* **1011** (2010) 149, [arXiv:1006.4090].

[63] T. Hartman and J. Maldacena, *Time Evolution of Entanglement Entropy from Black Hole Interiors*, *JHEP* **1305** (2013) 014, [arXiv:1303.1080].

[64] H. Liu and S. J. Suh, *Entanglement Tsunami: Universal Scaling in Holographic Thermalization*, *Phys.Rev.Lett.* **112** (2014) 011601, [arXiv:1305.7244].

[65] S. H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **1403** (2014) 067, [arXiv:1306.0622].

[66] B. Czech, J. L. Karczmarek, F. Nogueira, and M. Van Raamsdonk, *The Gravity Dual of a Density Matrix*, *Class.Quant.Grav.* **29** (2012) 155009, [arXiv:1204.1330].

[67] D. Marolf, *Black Holes, AdS, and CFTs*, *Gen.Rel.Grav.* **41** (2009) 903–917, [arXiv:0810.4886].

[68] B. Sundborg, *The Hagedorn transition, deconfinement and N=4 SYM theory*, *Nucl.Phys.* **B573** (2000) 349–363, [hep-th/9908001].

[69] O. Aharony, J. Marsano, S. Minwalla, K. Papadodimas, and M. Van Raamsdonk, *The Hagedorn - deconfinement phase transition in weakly coupled large N gauge theories*, *Adv.Theor.Math.Phys.* **8** (2004) 603–696, [hep-th/0310285].

[70] I. A. Morrison and M. M. Roberts, *Mutual information between thermo-field doubles and disconnected holographic boundaries*, *JHEP* **1307** (2013) 081, [arXiv:1211.2887].

[71] Y. Zhang, F. Pan, Y.-X. Liu, and J. Draayer, *Critical point symmetries in deformed odd-A nuclei*, *Phys.Rev.* **C84** (2011) 054319.

[72] V. E. Hubeny, M. Rangamani, and E. Tonni, *Thermalization of Causal Holographic Information*, *JHEP* **1305** (2013) 136, [arXiv:1302.0853].

[73] T. Takayanagi and T. Ugajin, *Measuring Black Hole Formations by Entanglement Entropy via Coarse-Graining*, *JHEP* **1011** (2010) 054, [arXiv:1008.3439].

[74] T. Banks, M. R. Douglas, G. T. Horowitz, and E. J. Martinec, *AdS dynamics from conformal field theory*, hep-th/9808016.

[75] J. Polchinski, *S matrices from AdS space-time*, hep-th/9901076.

[76] L. Susskind, *Holography in the flat space limit*, *AIP Conf.Proc.* **493** (1999) 98–112, [hep-th/9901079].

[77] S. B. Giddings, *Flat space scattering and bulk locality in the AdS / CFT correspondence*, *Phys.Rev.* **D61** (2000) 106008, [hep-th/9907129].

[78] G. T. Horowitz and V. E. Hubeny, *CFT description of small objects in AdS*, *JHEP* **0010** (2000) 027, [hep-th/0009051].

[79] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, *Local bulk operators in AdS/CFT: A Boundary view of horizons and locality*, *Phys.Rev.* **D73** (2006) 086003, [hep-th/0506118].

[80] M. Gary and S. B. Giddings, *The Flat space S-matrix from the AdS/CFT correspondence?*, *Phys.Rev.* **D80** (2009) 046008, [arXiv:0904.3544].

[81] I. Heemskerk, J. Penedones, J. Polchinski, and J. Sully, *Holography from Conformal Field Theory*, *JHEP* **0910** (2009) 079, [arXiv:0907.0151].

[82] J. Penedones, *Writing CFT correlation functions as AdS scattering amplitudes*, *JHEP* **1103** (2011) 025, [arXiv:1011.1485].

[83] D. Kabat, G. Lifschytz, and D. A. Lowe, *Constructing local bulk observables in interacting AdS/CFT*, *Phys.Rev.* **D83** (2011) 106009, [arXiv:1102.2910].

[84] A. L. Fitzpatrick and J. Kaplan, *Scattering States in AdS/CFT*, arXiv:1104.2597.

[85] I. Heemskerk, D. Marolf, J. Polchinski, and J. Sully, *Bulk and Transhorizon Measurements in AdS/CFT*, *JHEP* **1210** (2012) 165, [arXiv:1201.3664].

[86] S. Raju, *New Recursion Relations and a Flat Space Limit for AdS/CFT Correlators*, *Phys.Rev.* **D85** (2012) 126009, [arXiv:1201.6449].

[87] S. Raju, *Four Point Functions of the Stress Tensor and Conserved Currents in AdS_4/CFT_3*, *Phys.Rev.* **D85** (2012) 126008, [arXiv:1201.6452].

[88] V. E. Hubeny, H. Liu, and M. Rangamani, *Bulk-cone singularities and signatures of horizon formation in AdS/CFT*, *JHEP* **0701** (2007) 009, [hep-th/0610041].

[89] J. Hammersley, *Extracting the bulk metric from boundary information in asymptotically AdS spacetimes*, *JHEP* **0612** (2006) 047, [hep-th/0609202].

[90] S. Bilson, *Extracting spacetimes using the AdS/CFT conjecture*, *JHEP* **0808** (2008) 073, [arXiv:0807.3695].

[91] J. M. Maldacena, *Wilson loops in large N field theories*, *Phys.Rev.Lett.* **80** (1998) 4859–4862, [hep-th/9803002].

[92] S.-J. Rey and J.-T. Yee, *Macroscopic strings as heavy quarks in large N gauge theory and anti-de Sitter supergravity*, *Eur.Phys.J.* **C22** (2001) 379–394, [hep-th/9803001].

[93] V. E. Hubeny, *Precursors see inside black holes*, *Int.J.Mod.Phys.* **D12** (2003) 1693–1698, [hep-th/0208047].

[94] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **1302** (2013) 062, [arXiv:1207.3123].

[95] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, *An Apologia for Firewalls*, *JHEP* **1309** (2013) 018, [arXiv:1304.6483].

[96] J. Aparicio and E. Lopez, *Evolution of Two-Point Functions from Holography*, *JHEP* **1112** (2011) 082, [arXiv:1109.3571].

[97] T. Albash and C. V. Johnson, *Evolution of Holographic Entanglement Entropy after Thermal and Electromagnetic Quenches*, *New J.Phys.* **13** (2011) 045017, [arXiv:1008.3027].

[98] V. Balasubramanian, A. Bernamonti, J. de Boer, N. Copland, B. Craps, et al., *Thermalization of Strongly Coupled Field Theories*, *Phys.Rev.Lett.* **106** (2011) 191601, [arXiv:1012.4753].

[99] V. Balasubramanian, A. Bernamonti, J. de Boer, N. Copland, B. Craps, et al., *Holographic Thermalization*, *Phys.Rev.* **D84** (2011) 026010, [arXiv:1103.2683].

[100] H. Liu and S. J. Suh, *Entanglement growth during thermalization in holographic systems*, *Phys.Rev.* **D89** (2014), no. 6 066012, [arXiv:1311.1200].

[101] D. A. Lowe and S. Roy, *Holographic description of asymptotically AdS(2) collapse geometries*, *Phys.Rev.* **D78** (2008) 124017, [arXiv:0810.1750].

[102] T. Albash and C. V. Johnson, *Holographic Entanglement Entropy and Renormalization Group Flow*, *JHEP* **1202** (2012) 095, [arXiv:1110.1074].

[103] W. Baron, D. Galante, and M. Schvellinger, *Dynamics of holographic thermalization*, *JHEP* **1303** (2013) 070, [arXiv:1212.5234].

[104] E. Caceres and A. Kundu, *Holographic Thermalization with Chemical Potential*, *JHEP* **1209** (2012) 055, [arXiv:1205.2354].

[105] D. Galante and M. Schvellinger, *Thermalization with a chemical potential from AdS spaces*, *JHEP* **1207** (2012) 096, [arXiv:1205.1548].

[106] W. Fischler, S. Kundu, and J. F. Pedraza, *Entanglement and out-of-equilibrium dynamics in holographic models of de Sitter QFTs*, *JHEP* **1407** (2014) 021, [arXiv:1311.5519].

[107] V. E. Hubeny and M. Rangamani, *A Holographic view on physics out of equilibrium*, *Adv.High Energy Phys.* **2010** (2010) 297916, [arXiv:1006.3675].

[108] L. Fidkowski, V. Hubeny, M. Kleban, and S. Shenker, *The Black hole singularity in AdS / CFT*, *JHEP* **0402** (2004) 014, [hep-th/0306170].

[109] P. Kraus, H. Ooguri, and S. Shenker, *Inside the horizon with AdS / CFT*, *Phys.Rev.* **D67** (2003) 124022, [hep-th/0212277].

[110] P. Bizo and J. Jamuna, *Globally regular instability of AdS_3*, *Phys.Rev.Lett.* **111** (2013), no. 4 041102, [arXiv:1306.0317].

[111] M. Headrick, *General properties of holographic entanglement entropy*, *JHEP* **1403** (2014) 085, [arXiv:1312.6717].

[112] V. Balasubramanian and S. F. Ross, *Holographic particle detection*, *Phys.Rev.* **D61** (2000) 044007, [hep-th/9906226].

[113] J. Louko, D. Marolf, and S. F. Ross, *On geodesic propagators and black hole holography*, *Phys.Rev.* **D62** (2000) 044041, [hep-th/0002111].

[114] J. Kaplan, *Extracting data from behind horizons with the AdS / CFT correspondence*, hep-th/0402066.

[115] G. Festuccia and H. Liu, *Excursions beyond the horizon: Black hole singularities in Yang-Mills theories. I.*, *JHEP* **0604** (2006) 044, [hep-th/0506202].

[116] S. H. Shenker and D. Stanford, *Multiple Shocks*, *JHEP* **1412** (2014) 046, [arXiv:1312.3296].

[117] T. Andrade, S. Fischetti, D. Marolf, S. F. Ross, and M. Rozali, *Entanglement and correlations near extremality: CFTs dual to Reissner-Nordstrm AdS_5*, *JHEP* **1404** (2014) 023, [arXiv:1312.2839].

[118] E. Caceres, A. Kundu, J. F. Pedraza, and W. Tangarife, *Strong Subadditivity, Null Energy Condition and Charged Black Holes*, *JHEP* **1401** (2014) 084, [arXiv:1304.3398].

[119] N. Engelhardt and A. C. Wall, *Extremal Surface Barriers*, *JHEP* **1403** (2014) 068, [arXiv:1312.3699].

[120] F. Nogueira, *Extremal Surfaces in Asymptotically AdS Charged Boson Stars Backgrounds*, *Phys.Rev.* **D87** (2013), no. 10 106006, [arXiv:1301.4316].

[121] S. A. Gentle and M. Rangamani, *Holographic entanglement and causal information in coherent states*, *JHEP* **1401** (2014) 120, [arXiv:1311.0015].

[122] 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].

[123] 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].

[124] D. R. Brill, *Multi - black hole geometries in (2+1)-dimensional gravity*, *Phys.Rev.* **D53** (1996) 4133–4176, [gr-qc/9511022].

[125] S. Aminneborg, I. Bengtsson, D. Brill, S. Holst, and P. Peldan, *Black holes and wormholes in (2+1)-dimensions*, *Class.Quant.Grav.* **15** (1998) 627–644, [gr-qc/9707036].

[126] D. Brill, *Black holes and wormholes in (2+1)-dimensions*, *Lect.Notes Phys.* **537** (2000) 143, [gr-qc/9904083].

[127] S. Aminneborg, I. Bengtsson, and S. Holst, *A Spinning anti-de Sitter wormhole*, *Class.Quant.Grav.* **16** (1999) 363–382, [gr-qc/9805028].

[128] X. Yin, *On Non-handlebody Instantons in 3D Gravity*, *JHEP* **0809** (2008) 120, [arXiv:0711.2803].

[129] K. Skenderis and B. C. van Rees, *Holography and wormholes in 2+1 dimensions*, *Commun.Math.Phys.* **301** (2011) 583–626, [arXiv:0912.2090].

[130] K. Krasnov, *Holography and Riemann surfaces*, *Adv.Theor.Math.Phys.* **4** (2000) 929–979, [hep-th/0005106].

[131] J. M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **0304** (2003) 021, [hep-th/0106112].

[132] V. Balasubramanian, P. Hayden, A. Maloney, D. Marolf, and S. F. Ross, *Multiboundary Wormholes and Holographic Entanglement*, *Class.Quant.Grav.* **31** (2014) 185015, [arXiv:1406.2663].

[133] N. Iizuka and N. Ogawa, *On the Entanglement of Multiple CFTs via Rotating Black Hole Interior*, arXiv:1402.4548.

[134] J. Louko and D. Marolf, *Single exterior black holes and the AdS / CFT conjecture*, *Phys.Rev.* **D59** (1999) 066002, [hep-th/9808081].

[135] D. Giulini, *3-manifolds in canonical quantum gravity*. PhD thesis, University of Cambridge, 1989.

[136] J. L. Friedman, K. Schleich, and D. M. Witt, *Topological censorship*, *Phys.Rev.Lett.* **71** (1993) 1486–1489, [gr-qc/9305017].

[137] Wolfram Research, Inc., *Mathematica*. Wolfram Research, Inc., Champaign, Illinois, version 10.0 ed., 2014.

[138] Z. Xi, X.-M. Lu, X. Wang, and Y. Li, *Necessary and sufficient condition for saturating the upper bound of quantum discord*, arXiv:1111.3837.

[139] S. Fischetti and D. Marolf, *Complex Entangling Surfaces for AdS and Lifshitz Black Holes?*, *Class.Quant.Grav.* **31** (2014), no. 21 214005, [arXiv:1407.2900].

[140] M. Headrick, A. Lawrence, and M. Roberts, *Bose-Fermi duality and entanglement entropies*, *J.Stat.Mech.* **1302** (2013) P02022, [arXiv:1209.2428].

[141] E. Witten and S.-T. Yau, *Connectedness of the boundary in the AdS / CFT correspondence*, *Adv.Theor.Math.Phys.* **3** (1999) 1635–1655, [hep-th/9910245].

[142] J. M. Maldacena and L. Maoz, *Wormholes in AdS*, *JHEP* **0402** (2004) 053, [hep-th/0401024].

[143] A. Hatcher, *Algebraic Topology*. Cambridge University Press, 2002.

[144] H. Samelson, *On the Thom class of a submanifold*, *Michigan Math. J.* **12** (09, 1965) 257–261.

[145] P. Calabrese, J. Cardy, and E. Tonni, *Entanglement negativity in extended systems: A field theoretical approach*, *J.Stat.Mech.* **1302** (2013) P02008, [arXiv:1210.5359].

[146] J. Camps and W. R. Kelly, *Generalized gravitational entropy without replica symmetry*, arXiv:1412.4093.

[147] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch.Phys.* **61** (2013) 781–811, [arXiv:1306.0533].