Utrecht University

MASTER THESIS

Entanglement entropy, holography and gravity

Author:
L.R. DE RUITER BSc.
Theoretical Physics

Supervisor:
prof. dr. S.J.G. VANDOREN
Utrecht University

June 2018

Abstract

In this thesis we review how entanglement entropy is related to the gravitational equations of motion by means of the AdS/CFT correspondence. We consider the entanglement entropy of a ball-shaped spatial region in the CFT vacuum and derive a holographic relation between the relative entropy and the equations of motion in the dual bulk spacetime. For small perturbations around the CFT vacuum, we show that the first law of entanglement entropy requires the dual metric to satisfy the linearized Einstein equation when the entanglement entropy is given by the area of an extremal surface in the bulk. We discuss a recent paper [1] in which the complete Einstein equation is derived and explain why we believe this result is incomplete. This thesis serves both as a commentary on this paper, as well as a detailed review of the relationship between entanglement and gravity.

Contents

Chapter 1

Introduction

One of today's greatest challenges in physics is finding a consistent theory of quantum gravity. Quantum mechanics, in its relativistic form called quantum field theory, accurately describes the world we live in on a microscopic length scale. Driven by a desire to unify all models within a single framework, physicists started searching for a quantum description of Einstein's theory of general relativity. It turns out, however, that if GR is quantized canonically it cannot be renormalized. Consequently all attempts so far have failed.

The question why physicists want to unify all of physics within one theory is legitimate. Both general relativity and quantum field theory provide predictions which agree with experiments to high precision. On first sight it seems a matter of elegance, rather than necessity. However, general relativity contains singularities, which are points of infinite curvature. At these points the classical description of spacetime breaks down. A well-known example of singularities are black holes. Due to the extreme conditions in a black hole, it is thought that quantum mechanics should play a role in its description. Therefore, understanding black holes lies at the heart of understanding quantum gravity.

Over the years progress was made. By thinking about black holes as thermodynamical systems, it was concluded they must have non-vanishing entropy. The argument is very simple: if black holes were to carry no entropy, throwing in matter would remove entropy from the universe, which is a violation of the second law of thermodynamics. Motivated by the idea that more could be learned from black hole entropy, Hawking and Bekenstein derived a formula to calculate it [2, 3, 4]. The entropy is related to the area of the black hole horizon. This was puzzling because thermodynamic entropy is an extensive quantity, such that it is expected to scale with the volume of the black hole instead.

More recently, the AdS/CFT correspondence was conjectured [5]. Motivated by string theory, a duality between 4d $\mathcal{N} = 4$ super Yang-Mills theory and 10d supergravity on $\mathrm{AdS}_5 \times S^5$ was found by studying D3-branes. This correspondence suggests that certain quantum field theories can be re-expressed in terms of quantum theories of gravity. The quantum field theories of interest are conformal field theories. A CFT quantum state then has a gravitational dual description in asymptotically anti-de Sitter (AdS) spacetime. The geometry at the AdS boundary corresponds to the spacetime on which the quantum state is defined. In the bulk AdS spacetime, the metric is dynamical and should therefore satisfy gravitational equations of motion. In the classical regime, these will be the Einstein equations.

In the light of quantum gravity, some questions regarding AdS/CFT naturally pop up. What quantum mechanical object describes the geometry of the AdS space? And how are the gravitational equations of motion encoded within the quantum field theory? Understanding these questions might be our best lead to learn more about quantum gravity. Somewhat ironically, it turns out that gravity is intimately related to quantum entanglement. Entangled particles are in a quantum superposition state. They are correlated, even when separated. This phenomenon is inherent to quantum mechanics, and absent in classical theories such as general relativity. The amount of entanglement between two systems is measured by the entanglement entropy. This quantity can be thought of as the quantum generalization of thermal entropy. It scales as the area of the surface that separates the two systems. The black hole entropy shares this property. The event horizon separates the spacetime into the black hole interior and exterior. According to Hawking and Bekenstein, the black hole entropy is given by the area of the horizon. This indicates that entanglement plays a role here too.

In 2006, Ryu and Takayanagi proposed a method to compute the entanglement entropy using AdS/CFT [6]. The entanglement entropy in the CFT is given by the area of an extremal surface in the dual spacetime. This establishes a direct link between entanglement on one side, and geometry on the other. If an AdS geometry gives the right entanglement entropy of the CFT we can learn a lot about this geometry. Starting from this principle, it has been shown that an AdS geometry which gives the correct entropy of a perturbed CFT vacuum state must satisfy the linearized Einstein equation [7]. Research on this topic continues and hopes to discover what holographic constraints must be satisfied by the AdS metric in order for it to be a solution of the full Einstein equation. Recently, a paper was written by Oh, Park and Sin in which this problem is claimed to be solved [1]. This would a big step towards understanding of the relation between entanglement and the Einstein equation.

The main goal of this thesis is to check the validity of [1]. This requires a proper understanding of the framework introduced by [7], since the proof is a generalization of the linearized result to the full Einstein equation. This thesis will therefore also serve as a comprehensive review of the relation between gravity and entanglement entropy from a holographic viewpoint. We start by introducing the concept of entanglement entropy in Chapter 2. We define the relative entropy, which will be our quantity of interest. In Chapter 3 we consider the entanglement of a CFT. This includes setting up a scheme in which we can compute the relative entropy. We will discuss the holographic interpretation of the relative entropy in Chapter 4. This requires the identification of the CFT geometry with the boundary geometry of a dual AdS spacetime. In Chapter 5 we will focus on this AdS spacetime. Using a formalism introduced by Wald, we derive an equation which relates the gravity dual of the relative entropy to the gravitational equations of motion. We show in Chapter 6 that the AdS metric must satisfy the linearized Einstein equation. This concludes our review of [7]. In Chapter 7 we discuss recent research which goes beyond the linearized gravity result. Here we comment on [1] and review some relevant related work. Finally, in Chapter 8 we discuss the current status of the research on entanglement and gravity. This includes open problems and what we might hope to learn and achieve by solving them.

Chapter 2

Entanglement Entropy

2.1 Introduction

In a classical theory, particles can always be described by separate states in the system. If we construct a state with multiple particles, this state can be written as the product of the individual particle states. In the quantum world, this intuitive premise does not hold. Quantum theories contain the notion of quantum entanglement. Particles can be correlated, such that their quantum state can not be divided into individual states. Such a set of particles must be described using a quantum superposition. In this case, we say the particles are entangled.

A simple physical example of quantum entanglement can be constructed by two spin-$\frac{1}{2}$ particles. The Hilbert space $\mathcal{H}$ of this system is spanned by the basis $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$. According to quantum mechanics the following state is allowed

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \tag{2.1}$$

which is called a superposition. This is a pure quantum state. Since this state is a linear combination, we cannot say if the particles have spin up or down. When we measure the spin of one particle, the quantum state collapses to either $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$. Whatever spin we measure for the first particle, we immediately know that the second particle has the opposite spin. This proves the particles to be correlated. The two spins in the superposition state are entangled.

When quantum mechanics was still young, entanglement puzzled physicists. Einstein famously dubbed entanglement "spooky action at a distance", and used the seemingly paradoxical behavior as an argument against quantum mechanics. He reasoned that if we separate two entangled particles and measure one, the other particle state is determined instantaneously. This breaks causality, which dictates that signals cannot travel faster than the speed of light. In modern physics, entanglement has proven itself a useful object of study. Some examples of its applications are quantum phase transitions, black hole entropy and, by means of holography, the emergence of spacetime and gravity. The latter will be the subject of this work.

Entanglement can be quantitatively measured with entanglement entropy. In this chapter this quantity will be defined, alongside related concepts which will be important for our considerations. The discussion will be as general as possible.

2.2 Definitions

Consider a general quantum theory. The system will occupy a physical state, described by a *density matrix* ρ. A density matrix describes a quantum system as a statistical ensemble of quantum states. Each quantum state has a probability, and all probabilities add up to unity. Furthermore they are Hermitian and positive semi-definite. The von Neumann entropy of this system is defined by

$$S = -\operatorname{Tr}(\rho \log \rho), \tag{2.2}$$

which can be interpreted as the quantum generalization of the classical thermodynamic entropy.

Let us define some coordinate system in our theory, and define a spatial subregion B at a fixed time. We can describe the degrees of freedom within B with a density matrix ρ_B by integrating out all degrees of freedom within its complement $\bar{B}$. This is achieved by factoring the Hilbert space as $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_{\bar{B}}$ and tracing over $\mathcal{H}_{\bar{B}}$. We assume that this factorization exists. This procedure defines the *reduced density matrix*

$$\rho_B = \operatorname{Tr}_{\bar{B}}(\rho). \tag{2.3}$$

Reduced density matrices are both Hermitian and positive semidefinite, just like normal density matrices. As such, we can express ρ_B as

$$\rho_B = \frac{e^{-H_B}}{\operatorname{Tr}(e^{-H_B})}, \tag{2.4}$$

for some Hermitian operator H_B, which is called the *modular Hamiltonian* [8]. In general, the modular Hamiltonian is not a local operator. This means that it can not be written as a local expression in terms of the fields on B. Note that we included the denominator in the definition to ensure proper normalization. This makes sure that the trace of ρ_B equals unity, as we expect from a density matrix.

We can now define an entropy for the subregion B, analogous to the von Neumann entropy:

$$S_B = -\operatorname{Tr}(\rho_B \log \rho_B). \tag{2.5}$$

This quantity is called the *entanglement entropy*. It is a measure of the quantum entanglement between the subsystem B and its complement. In a QFT, the entanglement entropy is UV divergent due to the fact that spacetime is continuous. This is because a QFT has infinite degrees of freedom. When choosing a subregion B, there will be infinite entangled degrees of freedom between B and $\bar{B}$, causing the entropy to diverge. However, these divergences typically do not depend on the state of the system. They are well understood and can be regularized.

Another quantity of interest is the *relative entropy*. It is defined as

$$S(\rho|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma), \tag{2.6}$$

where ρ and σ denote two density matrices. The relative entropy is a measure of the distinguishability between two states, in this case that of ρ with respect to the reference state σ. One can show it to be a non-negative, monotonically increasing function, as one would expect from an entropy [9]. This paper is rather involved, so we included a proof of these properties a simple two-spin system in Appendix A. When the states are identical the relative entropy vanishes, marking its global minimum. Unlike the entanglement entropy, the relative entropy is regular in QFT. This is because the UV divergences cancel, rendering it a finite quantity. The relative entropy can also be calculated for two states described by reduced density matrices. This will be our main application.

2.3 The first law of Entanglement Entropy

The relative entropy is related to the modular Hamiltonian and the entanglement entropy.

$$\begin{aligned} S(\rho|\sigma) &= \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma) \\ &= \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma) + \mathrm{Tr}(\sigma \log \sigma) - \mathrm{Tr}(\sigma \log \sigma) \\ &= -S(\rho) + S(\sigma) + \mathrm{Tr}(\rho H) - \mathrm{Tr}(\sigma H) \\ &= \Delta \langle H \rangle - \Delta S, \end{aligned} \tag{2.7}$$

where we defined H to be the modular Hamiltonian for the reference state σ, $H = -\log \sigma$. Furthermore we defined $\Delta\langle H\rangle = \mathrm{Tr}(\rho H) - \mathrm{Tr}(\sigma H)$ and $\Delta S = S(\rho) - S(\sigma)$, the differences in modular energy and entanglement entropy between the two states. We will often write the modular energy $\langle H \rangle$ as E.

Note that the relative entropy is related to entanglement only when we are considering reduced density matrices. If we choose σ to be thermal, such that $\sigma = e^{-\beta H}$, the above expression simplifies to

$$S(\rho|\sigma) = \beta(F(\rho) - F(\sigma)) \tag{2.8}$$

where $F(\rho) = \text{Tr}(\rho H) - TS(\rho)$ is the free energy and S is the thermal entropy. This is the thermodynamical analog of equation (2.7). From here on, when we write a density matrix it will always denote a reduced density matrix implicitly.

Now consider a one parameter family of states described by $\rho(\lambda)$, such that $\rho(0) = \sigma$, the reference state. We can expand the relative entropy around the reference state σ in terms of λ as

$$S(\rho(\lambda)|\sigma) = S(\rho(0)|\sigma) + \frac{d}{d\lambda}S(\rho(\lambda)|\sigma)|_{\lambda=0}\lambda + \mathcal{O}(\lambda^2). \tag{2.9}$$

The lowest order cancels by definition of the relative entropy. The first order term also vanishes, since the relative entropy is a monotonically increasing function around σ. Therefore, the expansion of $S(\rho(\lambda)|\sigma)$ around $\lambda = 0$ starts at quadratic order. Expanding both sides of equation (2.7) up to first order in λ, we find

$$\delta S = \delta\langle H\rangle, \tag{2.10}$$

where $\delta f = \frac{d}{d\lambda}f|_{\lambda=0}$ for some function f. This is the *first law of entanglement entropy.* It relates the variation of the modular energy to the variation of the entanglement entropy. The name comes from its strong resemblance to the first law of thermodynamics.

2.4 Summary of the thesis

The rest of this thesis can be divided into two parts. Chapters 3-5 provide a derivation of the relationship between entanglement entropy and gravity. In Chapter 6 we explicitly show how the first law of entanglement entropy implies the linearized equations of motion in the dual spacetime. This concludes the first part. The second part, Chapter 7, will concern results beyond linearized gravity, which includes our comments on [1].

We will now present a brief summary of the derivations and results in Chapters 3-6. The calculations that are involved are detailed and often abstract. This summary will hopefully provide some overview when reading through the rest of the chapters. It is quite dense, so it is not necessary to understand all the details during a first read. It should however give some feeling for the structure of the arguments.

The first goal of this thesis is to relate entanglement entropy to the Einstein equations via AdS/CFT. We can achieve this by considering the gravity dual of the relative entropy. We have seen the QFT expression for the relative entropy

$$S(\rho|\sigma) = \Delta E - \Delta S. \tag{2.11}$$

We will choose the reference state σ to be the CFT vacuum state in d-dimensional Minkowski spacetime. The state ρ will denote a family of states $\rho(\lambda)$ parametrized by λ, such that $\rho(0) = \sigma$. This allows us to study perturbations around σ by expanding in λ. In order to identify the gravity dual of the relative entropy, we must determine the entanglement entropy and modular Hamiltonian and find their bulk counterparts. The

CFT state $\rho(\lambda)$ will be dual to an asymptotically AdS spacetime with metric $g(\lambda)$, such that $g(0)$ corresponds to pure AdS.

We will start with the modular Hamiltonian. It was shown in [10] that for a ball shaped region B of radius R the modular Hamiltonian H_B is given by

$$H_B = 2\pi \int_B d^{d-1}x \frac{R^2 - r^2}{2R} T_{00}. \tag{2.12}$$

where r is a radial coordinate measuring the distance to the center of the ball. This is derived by a conformal transformation relating the sphere B to the Rindler wedge, in which the expression for the modular Hamiltonian is given the the generator of Lorentz boosts. Mapping the modular Hamiltonian of the Rindler wedge back to Minkowski, we find the desired expression for H_B. To find the bulk interpretation of this quantity, we can relate the CFT stress tensor $T_{\mu\nu}$ to $T^{grav}_{\mu\nu}$, the stress tensor of the dual AdS spacetime. The modular energy E_B can then be interpreted a bulk quantity E_B^{grav}, evaluated at the asymptotic boundary of the AdS space. It is related to the energy of the AdS spacetime.

The entanglement entropy S_B can be computed using the Ryu-Takayanagi prescription. This prescription states that the entanglement entropy is given by the area of a codimension 2 extremal surface $\tilde{B}$ in the bulk spacetime, which ends on the entanglement surface. In our case, this surface is the boundary of B. By change of coordinates, we will show that the bulk region enclosed by B and $\tilde{B}$ is diffeomorphic to the exterior of a hyperbolic AdS black hole spacetime, in which $\tilde{B}$ is the horizon. The entanglement entropy S_B is then equal to the corresponding AdS black hole entropy, which we denote S_B^{grav}.

At this point we have identified E_B and S_B with bulk quantities. The next step is to relate the difference $\Delta E_B - \Delta S_B$ to the bulk equations of motion. This can be realized using a formalism developed by Wald [11] which provides a description of diffeomorphism invariant theories. We will apply this formalism to describe our bulk spacetime bounded by B and $\tilde{B}$, which we will denote Σ. Specifically, we will exploit a formula introduced in more recent work by Hollands and Wald [12],

$$\frac{d}{d\lambda}(\Delta E_B^{\text{grav}} - \Delta S_B^{\text{grav}}) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \mathcal{L}_{\xi_B} g(\lambda), \tfrac{d}{d\lambda} g(\lambda)) + \int_\Sigma \hat{E}(g(\lambda)). \tag{2.13}$$

The form $\hat{E}(g(\lambda))$ denotes an expression that vanishes when the dual AdS metric $g(\lambda)$ satisfies the equations of motion. The integral over $\boldsymbol{\omega}$ is a presymplectic form on the space of perturbations to the metric $g(\lambda)$. Using our bulk identifications of E_B and S_B, the left hand side is equal to the relative entropy in the CFT.

Equation (2.13) provides us with a holographic relation between the CFT relative entropy and the AdS equations of motion. By considering first order variations of the CFT state, we can show that the first law of entanglement entropy implies the linearized equations of motion in the dual AdS spacetime. More precisely, we show that the fluctuations around the pure AdS metric satisfy the linearized Einstein equations.

Chapter 3

Relative Entropy in a CFT

3.1 Introduction

In this chapter we focus on finding an explicit computational scheme for the relative entropy in a CFT. As we have seen, the relative entropy in terms of state ρ and reference state σ is given by

$$S(\rho|\sigma) = \Delta E - \Delta S. \tag{3.1}$$

Therefore, our goal is to find a way to compute the modular Hamiltonian and entanglement entropy of a general CFT. In both cases we will first sketch an outline of the arguments involved before proceeding with the detailed calculations.

This chapter is based on the work of Casini et al. [10]. It serves as a review of the arguments which are necessary to understand the framework in which we can relate entanglement entropy to gravity. The results should therefore not be interpreted as our own.

3.2 The Modular Hamiltonian

The modular Hamiltonian is in general a complicated object that cannot be written as a local expression of fields. One might justly wonder why it is a useful object to study at all. There are special cases in which the modular Hamiltonian simplifies and does have a local expression. These will be of our interest, as they allow us to calculate the modular energy.

We will consider a spherical entanglement surface in Minkowski space. For CFTs, we show that the domain of dependence of the sphere in the vacuum state is equivalent to Rindler space in a thermal state. In Rindler space, the modular Hamiltonian can be expressed in terms of the boost generator of the Lorentz group [13]. By inverting the conformal transformation we can map our modular Hamiltonian back to Minkowski space and obtain an explicit expression which depends on the stress tensor of the theory. The

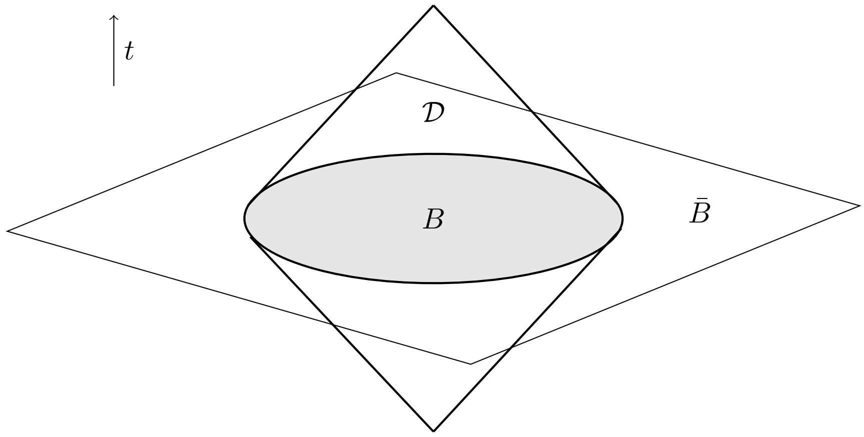

Figure 3.1 – A schematic overview of the setup. The spatial dimensions of the CFT are located in the plane. The shaded region B denotes the spatial sphere and the rest of the CFT is the complement $\bar{B}$. The entanglement surface is the boundary between B and $\bar{B}$ which we will denote ∂B. The domain of dependence $\mathcal{D}$ of B is the set of points such that every past (future) moving causal curve must intersect B.

variation of the Hamiltonian can then be computed from the variation of the stress tensor. Finally, the variation of the modular energy is given by taking the expectation value of the variation of the modular Hamiltonian.

3.2.1 Set up

Consider a CFT in d-dimensional Minkowski space $R^{1,d-1}$ in the vacuum state. We will denote the coordinates as $x^\mu = \{t, x^i\}$. We consider a spherical, spatial entanglement surface at a fixed time. Choose $t = 0$ without loss of generality. The entanglement surface divides our space into the interior of the sphere B, and its complement $\bar{B}$. The entanglement surface will be denoted as ∂B. We will denote the modular Hamiltonian of B as H_B, and the entanglement entropy as S_B. We will consider the causal development $\mathcal{D}$ of B. This is the union of the future and past domain of dependence, $\mathcal{D}^+ \cup \mathcal{D}^-$. The future (/past) domain of dependence of B is defined as the set of points for which all past (/future)-directed causal curves intersect B. A causal curve is a timelike or lightlike path, such that a signal can travel along it. The idea is that if we know the the state on B, we can evolve it to describe all points within $\mathcal{D}$. See Figure 3.1.

3.2.2 Rindler space

We will now consider Rindler space, which we will denote $\mathcal{R}$. A uniformly accelerating observer in Minkowski will undergo a hyperbolic motion. The frame in which this observer is at rest is often called the Rindler frame. We will denote the coordinates in $\mathcal{R}$ with X^μ to distinguish from the coordinates on B. Let the observer be accelerated in the X^1 direction. Rindler space, which is often called the Rindler wedge, then consists the part

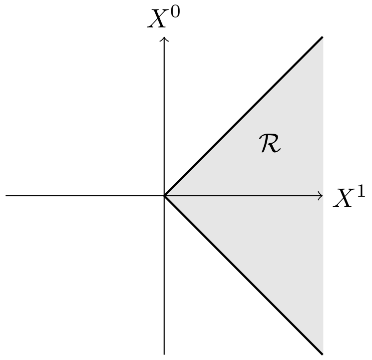

Figure 3.2 – A two dimensional sketch of the Rindler space $\mathcal{R}$. The direction of time X^0 and the direction of acceleration X^1 are shown. The wedge $\mathcal{R}$ corresponds to the region $0 < X^1 < \infty$, $-X^1 < X^0 < X^1$.

of Minkowski space given by $0 < X^1 < \infty$, $-X^1 < X^0 < X^1$. It corresponds to the domain where the coordinates of the Rindler frame are defined. See Figure 3.2.

As in any quantum theory, the physical state of the region $\mathcal{R}$ can be described by a density matrix which we will denote $\rho_{\mathcal{R}}$. Corresponding to the density matrix is a modular Hamiltonian $H_{\mathcal{R}}$. The reason that we are interested in Rindler space, is because its modular Hamiltonian takes a simple form. The Bisognano-Wichmann theorem states that $H_{\mathcal{R}}$ is simply the boost generator in the X^1 direction [13]. If the modular Hamiltonian induces evolution in a parameter s, the modular flow is given by

$$X^{\pm}(s) = X^{\pm} e^{\pm 2\pi s}, \tag{3.2}$$

where $X^{\pm} = X^1 \pm X^0$ are the null coordinates in $\mathcal{R}$. The other coordinates remain invariant under the flow, since they are not influenced by the boost. We will now consider a heuristic derivation of this statement. Start with the Lorentz boost given in terms of the hyperbolic functions

$$\begin{aligned} (X^0)' &= X^0 \cosh w + X^1 \sinh w \\ (X^1)' &= X^1 \cosh w + X^0 \sinh w \end{aligned} \tag{3.3}$$

with $w = \operatorname{arctanh}(v/c)$ the rapidity corresponding to the boost velocity v. We can rewrite this in terms of the null coordinates $X^{\pm} = X^0 \pm X^1$

$$\begin{aligned} (X^{\pm})' &= (X^0 \cosh w + X^1 \sinh w) \pm (X^1 \cosh w + X^0 \sinh w) \\ &= (X^0 \pm X^1) \cosh w + (X^1 \pm X^0) \sinh w \\ &= \tfrac{1}{2}(X^0 \pm X^1)(e^w + e^{-w}) + \tfrac{1}{2}(X^1 \pm X^0)(e^w - e^{-w}). \end{aligned} \tag{3.4}$$

From here one finds the result

$$(X^{\pm})' = X^{\pm} e^w. \tag{3.5}$$

Comparing with (3.2) we see that we find the correct answer up to a normalization factor of 2π.

The flow parametrizes orbits of constant acceleration, which coincide with the boost orbits by construction. The Hamiltonian in $\mathcal{R}$ is thermal with respect to translations along boost orbits. This phenomenon is commonly known as the Unruh effect [14]. Therefore, the modular Hamiltonian $H_\mathcal{R}$ will be thermal as well. We can show this by choosing new coordinates

$$\begin{aligned} X^0 &= z\sinh(\tau/R), \\ X^1 &= z\cosh(\tau/R), \end{aligned} \tag{3.6}$$

such that

$$X^\pm = z e^{\pm\tau/R}. \tag{3.7}$$

Rewriting the Rindler space metric in terms of these coordinates yields

$$ds^2 = dX^+dX^- + \sum_{i=2}^{d-1}(dX^i)^2 = -\frac{z^2}{R^2}d\tau^2 + dz^2 + \sum_{i=2}^{d-1}(dX^i)^2. \tag{3.8}$$

In this form, we recognize z and τ to be polar coordinates with radius z and angle $\frac{i}{R}\tau$. Through this identification, the angle is seen to be 2π-periodic, such that τ is $2\pi iR$ periodic. This is easily verified using equation (3.6). We conclude that the Rindler state is thermal with respect to H_τ, the Hamiltonian which generates translations in τ, with temperature $T = 1/2\pi R$. Hence, we can identify

$$\rho_\mathcal{R} = \frac{e^{-H_\mathcal{R}}}{\mathrm{Tr}(e^{-H_\mathcal{R}})} = \frac{e^{-\beta H_\tau}}{\mathrm{Tr}(e^{-\beta H_\tau})}. \tag{3.9}$$

We can now read off the expression for the modular Hamiltonian

$$H_\mathcal{R} = 2\pi R H_\tau. \tag{3.10}$$

We see that the modular Hamiltonian generates translations in τ. By comparing equations (3.2) and (3.7), the modular flow in terms of the new coordinates is given by $\tau \to \tau + 2\pi Rs$.

3.2.3 A Map from $\mathcal{D}$ to $\mathcal{R}$

To find an expresssion for H_B, we would like to relate H_B to $H_\mathcal{R}$. Therefore, we need a map between our causal diamond $\mathcal{D}$ and Rindler space $\mathcal{R}$. It will be useful to define null coordinates on $\mathcal{D}$ which exploit the spherical symmetry of B. Define

$$\begin{aligned} x^\pm &= r \pm t, \\ r &= \sqrt{(x^1)^2 + \ldots + (x^{(d-1)})^2}, \end{aligned} \tag{3.11}$$

where r is a radial coordinate on B. In these coordinates, the causal development of the ball $\mathcal{D}$ is given by $\{x^+ \leq R\} \cap \{x^- \leq R\}$. The map which transforms the Rindler wedge $\mathcal{R}$ to the causal diamond $\mathcal{D}$ consists of a special conformal transformation combined with a translation. It is given by [8]

$$x^\mu = \frac{X^\mu - (X\cdot X)C^\mu}{1 - 2(X\cdot C) + (X\cdot X)(C\cdot C)} + 2R^2C^\mu, \tag{3.12}$$

with $C^\mu = (0, -1/2R, 0, \dots)$[1]. It is not difficult to verify that $X^\pm \geq 0$ covers $-R \leq x^\pm \leq R$, such that we end up with $\mathcal{D}$.

Our goal is to find an expression for the modular Hamiltonian in $\mathcal{D}$. Since the modular Hamiltonian is the generator of shifts in s, we need to know what the modular flow looks like in $\mathcal{D}$. This is possible by mapping the modular flow from Rindler space back to Minkowski. The result is given by

$$x^\pm(s) = R\frac{(R + x^\pm) - e^{\mp 2\pi s}(R - x^\pm)}{(R + x^\pm) + e^{\mp 2\pi s}(R - x^\pm)}. \tag{3.13}$$

The calculation is given in Appendix B. This flow can be proven to correspond to the modular flow in $\mathcal{D}$. This was shown by [10]. It is generated by the modular Hamiltonian H_B. When we consider infinitesimal transformations, the shift δs on the surface $t = 0$ is of the form

$$\delta t = 2\pi \frac{R^2 - r^2}{2R}\delta s. \tag{3.14}$$

The corresponding operator in the CFT may now be identified as

$$H_B = 2\pi \int_B d^{d-1}x \frac{R^2 - r^2}{2R} T_{00}, \tag{3.15}$$

where $T_{\mu\nu}$ is the stress tensor of the CFT.

It is possible to rewrite this result in a more covariant form. In $\mathcal{R}$, the modular Hamiltonian $H_{\mathcal{R}}$ induces translations in the coordinate τ. Define ζ_B to be the image of the Killing vector $2\pi R\partial_\tau$ under the conformal transformation from $\mathcal{R}$ to $\mathcal{D}$. This vector will induce the flow in $\mathcal{D}$. It is given explicitly by

$$\zeta_B = \frac{\pi}{R}(R^2 - t^2 - x^i x_i)\partial_t - \frac{2\pi}{R} t x^i \partial_i. \tag{3.16}$$

This vector will set the direction of 'time' associated with the modular flow. We can now express the modular Hamiltonian as

$$H_B = \int_B d\Sigma^\mu \zeta_B^\nu T_{\mu\nu}, \tag{3.17}$$

where $d\Sigma^\mu$ is the $(d-1)$-dimensional volume form on B. The modular energy E_B is given by the expectation value of this integral. In conclusion, we have derived that the modular energy is given by

$$E_B = \int_B d\Sigma^\mu \zeta_B^\nu \langle T_{\mu\nu}\rangle. \tag{3.18}$$

3.3 The Entanglement Entropy

The entanglement entropy is generally a difficult quantity to compute. A common method is to make use of the so called replica trick. A review can be found in [15]. Computations

[1] If you are planning on working through these calculations, be careful. Reference [10] contains a typo in defining C^μ. It differs from our definition by a minus sign. The transformation as given here is the correct one.

become intractable very quickly and results are obtained in simple cases only, such as free field theories and CFTs in $(1+1)$-dimensions. We will not work out the details of the replica trick. Instead, we will rely on a different method to compute the entanglement entropy.

In 2006, Ryu and Takayanagi proposed a prescription to calculate the entanglement entropy using a holographic setup [6]. The so called Ryu-Takayanagi (RT) proposal makes use of the AdS/CFT correspondence and relates the CFT entanglement entropy to the area of an extremal codimension-2 surface in the bulk theory, which ends on the CFT entanglement surface. An extremal surface is a surface of which the area is extremal. The RT proposal considered surfaces at fixed points in time. This construction was later generalized to a covariant version, which is called the Hubeny-Rangamani-Takayanagi (HRT) proposal. Unlike the RT proposal, the HRT proposal holds in any Lorentz frame. The correctness of both proposals has been proven recently [16, 17]. The explicit expression for the entanglement entropy is

$$S_B = \frac{A(\tilde{B})}{4G_N}, \tag{3.19}$$

where A denotes the area functional, $\tilde{B}$ is the extremal surface ending on B and G_N is Newton's constant.

The HRT proposal allows us to calculate the entanglement entropy of the CFT sphere B, if we know the gravitational dual to our CFT. For an arbitrary CFT state it is not known what the gravitational dual looks like, or if it exists at all. If the CFT is in the vacuum state, the dual is pure AdS. This allows us to apply the HRT procedure to find the entanglement entropy of our spatial spherical region in the CFT vacuum. Computationally, this is a variational problem. We need to write down a bulk area functional in terms of a set of coordinates which parametrize the surface. We then vary the surface while restricting the boundary coincide with the entanglement surface, which is ∂B. For this specific case, the calculation is not too difficult. However, the relative entropy is given in terms of the variation of the entanglement entropy. We would then need to identify the gravitational dual of a perturbed CFT state, which is difficult.

As it turns out, there is another way to compute the entanglement entropy of a spherical region in the Minkowski vacuum. By a conformal transformation we can relate the vacuum state in $\mathcal{D}$ to a thermal state on a hyperbolic space $R \times H^{d-1}$. This transformation is closely related to the map from $\mathcal{D}$ to Rindler space. The entanglement entropy is then equal to the von Neumann entropy of the density matrix in $R \times H^{d-1}$, which we will show to be thermal. This is a thermodynamic entropy. In this section we show how this can be achieved.

3.3.1 A Map from $\mathcal{D}$ to $R \times H^{d-1}$

Recall the Minkowski metric in Rindler coordinates (3.8). We can divide out a scale factor and rewrite it as

$$ds^2 = \frac{z^2}{R^2}\left[-d\tau^2 + \frac{R^2}{z^2}\left(dz^2 + \sum_{i=2}^{d-1}(dX^i)^2\right)\right]. \tag{3.20}$$

Eliminate the scale factor by a Weyl transformation. Since we are considering a CFT, this is a symmetry of the theory. The result is recognized to be the metric on the $(d-1)$-dimensional hyperbolic space H^{d-1} with an additional timelike direction parametrized by τ,

$$ds^2 = -d\tau^2 + \frac{R^2}{z^2}\left(dz^2 + \sum_{i=2}^{d-1}(dX^i)^2\right). \tag{3.21}$$

We will denote this space as $R \times H^{d-1}$. The hyperbolic space H is the analog of Euclidean space with constant negative curvature. The radius of curvature in our hyperbolic space is given by R.

Since the transformation from $\mathcal{R}$ to $R \times H^{d-1}$ does not depend on τ, it does not change the modular Hamiltonian. The modular Hamiltonian is simply given by $H_{R\times H^{d-1}} = H_{\mathcal{R}} = 2\pi R H_\tau$. We conclude that our vacuum state in $\mathcal{D}$ is conformally related to a thermal state in $R \times H^{d-1}$.

The von Neumann entropy is invariant under conformal transformations. This was reasoned in [10]. They argue there exists a unitary operator which maps the CFT from $\mathcal{D}$ to $\mathcal{R}$, and similarly from $\mathcal{R}$ to $R \times H^{d-1}$. Since the von Neumann entropy is invariant under unitary transformations, it is unaffected by the conformal mapping. The entanglement entropy across the boundary of B must be equal to the von Neumann entropy of the density matrix of $R \times H^{d-1}$, which is thermal. This means our entanglement entropy is equivalent to the thermodynamic entropy in $R \times H^{d-1}$.

In [10] it is not explained why conformal transformations are unitary operations. Conformal transformations do preserve the inner product. Namely, we have $ds^2 = \eta_{\mu\nu} dX^\mu dX^\nu = \Omega^2 \eta_{\mu\nu} dx^\mu dx^\nu$ for some conformal prefactor Ω. Due to the Weyl symmetry of a CFT, the inner product is invariant. Furthermore, the mapping is invertible. We can both map $\mathcal{D} \to \mathcal{R}$ and back. These properties seem to be in favor of the claim. It would be useful to have a more mathematical rigorous proof. However, we are not able to give this in this thesis, which is why we have to assume the assertion to hold instead.

By rewriting the entanglement entropy as a thermal entropy we haven't solved anything yet, we have just rewritten our problem as a different one. This derivation is a preliminary result which we will need in Chapter 4. Via the AdS/CFT correspondence we will relate this to the entropy of an asymptotically AdS black hole spacetime. This will be the subject of Chapter 4.

Chapter 4

Holographic interpretations

In the previous chapter we discussed the modular Hamiltonian and entanglement entropy of the CFT. We should in principle be able to calculate the relative entropy in the CFT, aside from possible computational difficulties. The next step we will take makes use of the AdS/CFT correspondence. Ultimately, we are interested in determining the gravitational dual of the relative entropy $S(\rho(\lambda)|\sigma)$ in the CFT. The goal of this chapter is to find holographic interpretations of both the modular Hamiltonian and the entanglement entropy, in order to obtain a holographic interpretation of the relative entropy. We will denote these quantities as E_B^{grav} and S_B^{grav}. The superscript serves as a reminder that they are defined in the gravitational bulk theory.

First of all, we will set up some necessary framework to use the AdS/CFT correspondence. We will start by giving a quick review of AdS space and express it in different coordinate systems. We continue our discussion by identifying the dual CFT metric. This is achieved by looking at different foliations of the AdS metric and studying the metric on the AdS boundary. Assuming our theory to have a holographic dual, the metric of the CFT will correspond to the asymptotic boundary metric of the AdS space.

A thermal CFT state is dual to a black hole geometry in the AdS bulk. By considering foliations of AdS space we can identify the black hole geometry that corresponds to our CFT. Through this identification, we show that the entanglement entropy of a spherical region in the vacuum is equivalent to the entropy of the bulk black hole. We will work out this procedure in detail and show that it yields the same result as the HRT prescription.

The modular Hamiltonian also has a bulk interpretation. It can be expressed in terms of the asymptotic stress tensor of the AdS space. This will be explained in more detail below.

4.1 Geometry of AdS

In order to holographically relate our CFT to an AdS space, we need to know more about AdS. In this section we review the geometry of the AdS spacetime and look at different foliations. A foliation is a decomposition of a manifold into parallel submanifolds of smaller dimension. We will describe our AdS spacetime as a union of hypersurfaces. By taking the asymptotic boundary limit, we obtain a hypersurface which encodes the boundary geometry of the spacetime. If the AdS spacetime has a holographic CFT dual, the geometry of this theory will correspond to the geometry of this asymptotic hypersurface.

The AdS_{d+1} geometry is described by embedding the surface

$$-y_{d+1}^2 - y_0^2 + y_1^2 + \ldots + y_d^2 = -L^2 \tag{4.1}$$

within a $R^{2,d}$ space. Here, L is a constant which will turn out to be the AdS radius. There are two choices of coordinates which will be of our interest. The first choice is called the Poincaré coordinate system. It can be derived by choosing

$$\begin{aligned} y_{d+1} + y_d &= \frac{L^2}{z}, \\ y^a &= \frac{L}{z}x^a, \quad a = 0, \ldots, d-1 \end{aligned} \tag{4.2}$$

and substituting them into the hyperbolic surface,

$$\begin{aligned} -L^2 &= -y_{d+1}^2 - y_0^2 + y_1^2 + \ldots + y_d^2 \\ &= -(y_{d+1} + y_d)(y_{d+1} - y_d) - y_0^2 + \sum_{i=1}^{d-1} y_i^2 \\ &= -\frac{L^2}{z}(y_{d+1} - y_d) + \frac{L^2}{z^2}\eta_{ab}x^a x^b \end{aligned} \tag{4.3}$$

with η_{ab} the metric of $R^{1,d-1}$. This yields the following constraint,

$$(y_{d+1} - y_d) = z + \frac{1}{z}\eta_{ab}x^a x^b \tag{4.4}$$

from which we can compute the induced metric on the embedded surface.

$$\begin{aligned} ds^2 =& - dy_{d+1}^2 - dy_0^2 + \sum_{i=1}^{d} dy_i^2 \\ =& - d(y_{d+1} + y_d)d(y_{d+1} - y_d) + \eta_{ab}dy^a dy^b \\ =& - d(\frac{L^2}{z})d(z + \frac{1}{z}\eta_{ab}x^a x^b) + \eta_{ab}d(\frac{L}{z}x^a)d(\frac{L}{z}x^b) \\ =& \frac{L^2}{z^2}(dz^2 - \frac{1}{z^2}\eta_{ab}x^a x^b dz^2 + \frac{2}{z}\eta_{ab}x^a dx^b dz) + \eta_{ab}\frac{L^2}{z^2}(\frac{1}{z^2}x^a x^b dz^2 + dx^a dx^b - \frac{2}{z}x^a dx^b dz) \\ =& \frac{L^2}{z^2}(dz^2 + \eta_{ab}dx^a dx^b). \end{aligned} \tag{4.5}$$

Equation (4.5) is called the AdS_{d+1} Poincaré metric. From the metric, we see that the AdS space is foliated in sheets of Minkowski space, warped by a scale factor which depends on the z coordinate. Taking the limit $z \to 0$ allows us to study the asymptotic boundary metric which will correspond to the metric of the holographic CFT. Define a new coordinate $\tilde{z} = \log z$ and rewrite the metric in terms of this coordinate. We find

$$ds^2 = L^2(d\tilde{z}^2 + e^{-2\tilde{z}}\eta_{ab}dx^a dx^b). \tag{4.6}$$

The asymptotic limit now corresponds to $\tilde{z} \to -\infty$. In this limit the contribution of $d\tilde{z}$ becomes negligible. After removing the conformal prefactor we find that the AdS boundary has the geometry of Minkowski space

$$ds^2 = \eta_{ab}dx^a dx^b. \tag{4.7}$$

In Chapter 3 we defined the CFT setup, which was given in Minkowski coordinates. This means we can describe the holographic dual AdS space in terms of Poincaré coordinates.

The second set of coordinates we will consider are often called the global coordinates of AdS. They will turn out to induce a foliation of hyperbolic planes. They are given by

$$\begin{aligned}
y_{d+1} &= \rho\cosh(u)\\
y_0 &= \tilde{\rho}\sinh(\tau/L)\\
y_d &= \tilde{\rho}\cosh(\tau/L)\\
y_1 &= \rho\sinh(u)\cos(\phi_1)\\
y_2 &= \rho\sinh(u)\sin(\phi_1)\cos(\phi_2)\\
&\vdots\\
y_{d-1} &= \rho\sinh(u)\sin(\phi_1)\sin(\phi_2)\dots\cos(\phi_{d-2})
\end{aligned} \tag{4.8}$$

Plugging these into equation (4.1) we again derive a constraint.

$$\begin{aligned}
-L^2 &= -y_{d+1}^2 - y_0^2 + y_1^2 + \dots + y_d^2\\
&= -\rho^2\cosh^2(u) - \tilde{\rho}^2\sinh^2(\tau/L) + \tilde{\rho}^2\cosh^2(\tau/L) + \rho^2\sinh^2(u)(\dots)\\
&= \tilde{\rho}^2 - \rho^2.
\end{aligned} \tag{4.9}$$

In the second line, $(\dots)$ denotes terms from the spherical angles ϕ. These add op to unity by construction. Now compute the metric in these coordinates

$$\begin{aligned}
ds^2 =& -dy_{d+1}^2 - dy_0^2 + \sum_{i=1}^{d} dy_i^2\\
=& -d(\rho\cosh(u))^2 - d(\tilde{\rho}\sinh(\tau/L))^2 + d(\tilde{\rho}\cosh(\tau/L))^2\\
&+ d(\rho\sinh(u)\cos(\phi_1))^2 + \dots + d(\rho\sinh(u)\sin(\phi_1)\dots\sin(\phi_{d-2}))^2\\
=& \left(\frac{\rho^2}{L^2} - 1\right)^{-1} d\rho^2 - \left(\frac{\rho^2}{L^2} - 1\right) d\tau^2 + \rho^2(du^2 + \sinh^2(u)\, d\Omega_{d-2}^2).
\end{aligned} \tag{4.10}$$

In deriving the final expression, the constraint was used to get rid of $\tilde{\rho}$ in favor of ρ. The last term within brackets containing du and $d\Omega_{d-2}$ (the angular part of a $(d-2)$-sphere) is the metric of a hyperbolic plane H in $(d-1)$ dimensions. To study the behavior at the boundary, we take the limit $\rho \to \infty$. The $d\rho$ term vanishes, and after taking out a conformal factor ρ^2/L^2 we find

$$ds^2 = -d\tau^2 + L^2(du^2 + \sinh^2(u)\, d\Omega_{d-2}^2) \tag{4.11}$$

which is the metric of a $(d-1)$ dimensional hyperbolic plane with radius of curvature L, with an additional timelike direction τ. We see that in these coordinates the AdS space is foliated in terms of $R \times H^{d-1}$ surfaces.

These coordinates can be used to describe the bulk dual to the thermal state $R \times H^{d-1}$ that is conformally related to the vacuum state in $\mathcal{D}$ which we described in Chapter 3.

4.2 Identification of B with a bulk black hole

An interesting feature of the metric (4.10) is the fact that it is a AdS black hole solution [18]. More specifically, the metric we have found belongs to the class of hyperbolic black holes, or topological black holes. It is isometric to pure AdS space, since we obtained it by choosing the coordinates (4.8) on the embedding surface. Our solution is not a proper black hole in the sense that it non-singular at $\rho = 0$. The coordinate patch breaks down at $\rho = L$, which can be shown to be analogous to a Rindler horizon. Furthermore, this horizon has non-vanishing area and temperature, given by $T = 1/2\pi L$. The temperature can be calculated with the usual Euclidean prescription. Start with the metric (4.10) and go to Euclidean signature by choosing $\tau = it$. Consider the metric near the horizon, $\rho \approx L$. By expanding the prefactors in terms of $(\rho - L)$ up to linear order we have

$$ds^2 = \frac{L}{2(L-\rho)} d\rho^2 + \frac{2(L-\rho)}{L} dt^2 + L^2(du^2 + \sinh^2(u)\, d\Omega_{d-2}^2). \tag{4.12}$$

We want to rewrite the ρ and t coordinates into a form similar to polar coordinates, in order to identify the conical deficit. This is achieved by the coordinate transformation $\rho = L - \frac{r^2}{2L}$. In terms of the new coordinate r, the metric takes the form

$$ds^2 = dr^2 + \frac{r^2}{L^2} dt^2 + L^2(du^2 + \sinh^2(u)\, d\Omega_{d-2}^2). \tag{4.13}$$

From this expression we read off that t must be periodic in $t \to t + 2\pi L$. The temperature of the black hole is thus identified as $1/2\pi L$.

Assuming that the bulk theory is GR, the horizon area A is related to the black hole entropy via the Bekenstein-Hawking formula,

$$S_{\text{BH}} = \frac{A}{4G_N}. \tag{4.14}$$

The horizon area can be computed from equation (4.13) by setting $\rho = L$, or $r = 0$. We then obtain the metric on the $(d-2)$-dimensional hyperbolic space H^{d-2}. The 'black hole' therefore has a non-vanishing black hole entropy associated with it, even when it is isometric to pure AdS. This seemingly paradoxical fact is the bulk analog of the fact that the vacuum entanglement entropy is non-vanishing.

We should remark that the horizon area of the black hole is actually divergent. This is because the hyperbolic geometry extends to infinity. Mathematically, this occurs when integrating $u \to \infty$ which corresponds to the AdS boundary limit $z \to 0$. The fact that the black hole entropy diverges might seem like a problem, but it is actually consistent with our considerations in Chapter 2. In a continuous field theory, the entanglement entropy suffers UV divergences due to the entanglement of an infinite amount of degrees of freedom. We see that in the AdS language this translates to the horizon area being divergent.

We would like to identify the spherical CFT region B as the boundary of this AdS black hole space time. If this is possible, we can calculate the entanglement entropy of B as the black hole horizon entropy of the dual AdS spacetime. This follows from our derivation in Chapter 3, where we reasoned that S_B is equal to the thermal entropy in the hyperbolic space, which is in turn related to the hyperbolic black hole entropy. We will now establish the identification by working out the relation between the Poincaré and hyperbolic coordinates on the AdS boundary.

Combining the two different choices of coordinates (4.2) and (4.8) yields

$$\begin{aligned} y_{d+1} + y_d &= \frac{L^2}{z} = \rho\cosh(u) + \tilde{\rho}\cosh(\tau/L) \\ y_0 &= \frac{L}{z}t = \tilde{\rho}\sinh(\tau/L) \\ y_1^2 + \ldots + y_{d-1}^2 &= \frac{L^2}{z^2}r^2 = \rho^2\sinh^2(u) \end{aligned} \tag{4.15}$$

where we defined $x_0 = t$ and used the radial coordinate $r^2 = x_1^2 + \ldots + x_{d-1}^2$ for the CFT. We can now express the CFT coordinates (t, r) in terms of (ρ, τ, u) as

$$\begin{aligned} t &= L\frac{\tilde{\rho}\sinh(\tau/L)}{\rho\cosh(u) + \tilde{\rho}\cosh(\tau/L)} \\ r &= L\frac{\rho\sinh(u)}{\rho\cosh(u) + \tilde{\rho}\cosh(\tau/L)} \end{aligned} \tag{4.16}$$

The transformation between the two isometric boundary CFTs is obtained by taking the limit $\rho \to \infty$, which implies $\tilde{\rho} = \rho\sqrt{1 - \frac{L^2}{\rho^2}} \to \rho$,

$$\begin{aligned} t &= L\frac{\sinh(\tau/L)}{\cosh(u) + \cosh(\tau/L)} \\ r &= L\frac{\sinh(u)}{\cosh(u) + \cosh(\tau/L)}. \end{aligned} \tag{4.17}$$

One can show this transformation to be conformal. This was to be expected, since the associated AdS spaces are related by an isometry as well.

In order to make the identification complete, we need to make sure the black hole horizon asymptotes to ∂B, the boundary of our sphere. The black hole horizon is given by $\rho = L$

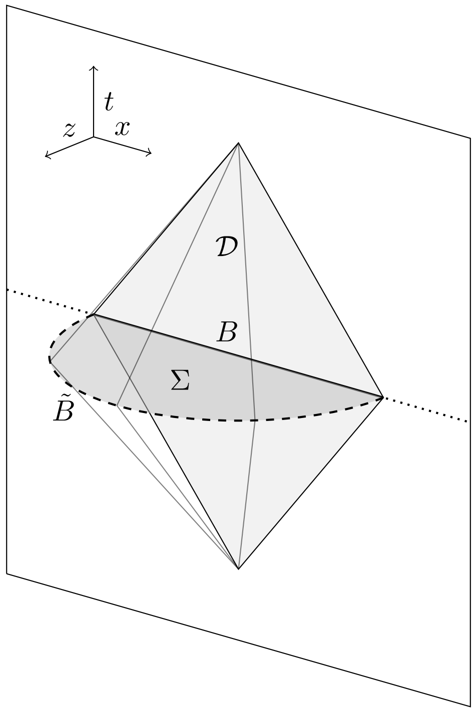

Figure 4.1 – A schematic overview of the Rindler wedge Σ associated with the sphere B. The CFT lives on the boundary and Σ extends into the AdS space. It is bounded by B and the extremal surface $\tilde{B}$, which is denoted by a dashed line. The black hole horizon coincides with $\tilde{B}$ when we consider GR.

which implies $\tilde{\rho} = 0$. From the first line of equation (4.15) we find $\frac{L}{z} = \cosh(u)$, so the limit $z \to 0$ translates to $u \to \infty$. In this limit we obtain $t \to 0$ and r asymptotes to

$$\lim_{u\to\infty} r = \lim_{u\to\infty} L \tanh(u) = L. \tag{4.18}$$

The black hole horizon intersects the AdS boundary at a circle with radius L at $t = 0$. By setting $L = R$ this circle coincides precisely with ∂B which completes our identification. It can also be checked that the transformation (4.17) with $L = R$ maps the causal diamond $\mathcal{D}$ to $R \times H^{d-1}$. This transformation was also given in [10]. This concludes our considerations. We have established that the entanglement entropy across a sphere B in flat space in the vacuum state equals the horizon entropy of a hyperbolic black hole if the AdS radius equals the radius of B. Figure 4.1 gives a schematic overview of the construction. We will denote the black hole horizon with $\tilde{B}$. The region bounded by B and $\tilde{B}$ we will call Σ, which corresponds to the black hole geometry without the black hole interior.

By comparing the HRT proposal and the black hole entropy, we can identify the extremal surface which calculates S_B to be the black hole horizon. The horizon $\tilde{B}$ has the same boundary as the entanglement surface B. If the surface is extremal, the two prescriptions to calculate the entanglement entropy agree, assuming the AdS theory is GR. This is no coincidence. The horizon $\tilde{B}$ at $t = 0$ is actually the bifurcation surface of the black hole, which is always extremal. We will explain more about this in Chapter 5. Therefore, the construction considered in this chapter can be viewed as a verification of the HRT proposal.

The HRT proposal could in principle be applied to any CFT state $\rho(\lambda)$, as long as we are able to compute the extremal surface in the bulk. However, it is only valid for GR. A generalization of the HRT formula is necessary if we are interested in more general gravity theories. For general theories, we could instead compute the entanglement entropy as the black hole entropy of some related AdS geometry. In this case we would need a more general equation for the black hole entropy, since the Bekenstein-Hawking relation is only valid in GR. This equation exists, and will be introduced below. This method only works when considering the vacuum state $\rho(0)$, since the identification of S_B with S_{BH} doesn't hold for general $\rho(\lambda)$.

4.3 The entanglement entropy

We have derived that the CFT vacuum entanglement entropy of B is given holographically by the corresponding AdS black hole entropy. In GR, the black hole entropy is given by the Bekenstein-Hawking formula S_{BH}. To keep the discussion more general, we can instead use the entropy given by Wald [11]

$$S_{\text{Wald}} = \frac{2\pi}{\kappa} \int_{\mathcal{B}} \boldsymbol{Q} \tag{4.19}$$

which calculates the entropy of any classical and covariant gravity theory. Here $\mathcal{B}$ denotes the bifurcation surface of the black hole and $\boldsymbol{Q}$ the Noether charge form. These concepts will be introduced in Chapter 5. We now have the following identification for the vacuum state for the CFT,

$$S_B = S_{\text{Wald}} \equiv S_B^{\text{grav}}. \tag{4.20}$$

The relative entropy is expressed in terms of the change in entanglement entropy. Therefore we must generalize our computational scheme of the entanglement entropy to general states which are not the CFT vacuum. When considering a small perturbation around the vacuum state, we can interpret this in the bulk as a fluctuation around the AdS black hole metric. This will induce a change in the horizon, causing the black hole entropy to change as well.

When considering finite perturbations to the CFT state, the bulk geometry is not readily found since our identification with the AdS black hole does not apply. This poses a problem when we want to write down an expression for the variation of S_B^{grav}, which is necessary to compute the dual of the relative entropy. This problem can be solved using the Hollands-Wald gauge. This gauge choice fixes the coordinate location of the black hole horizon. The extremal surface will then correspond to the black hole horizon, even for perturbed spacetimes. Using this gauge we obtain

$$\frac{d}{d\lambda} S_{\text{Wald}} = \frac{2\pi}{\kappa} \int_{\mathcal{B}} \frac{d}{d\lambda} \boldsymbol{Q}. \tag{4.21}$$

This will be explained in further detail in Chapter 5. From this expression we obtain

$$\frac{d}{d\lambda}S_B = \frac{d}{d\lambda}S_B^{\text{grav}} = \frac{2\pi}{\kappa}\int_{\mathcal{B}}\frac{d}{d\lambda}\boldsymbol{Q}. \tag{4.22}$$

4.4 The modular energy

We have derived the modular Hamiltonian of B in the CFT vacuum, which was written as a local expression. In order to proceed, we need to find its holographic interpretation E_B^{grav}. To do so, we have to define a notion of energy in the AdS space.

For a diffeomorphism invariant theory it is difficult to assign a local notion of energy-momentum to the gravitational field. The momentum is usually related to the field by its first derivative. However, by transforming to locally flat coordinates we can always make the first metric derivatives vanish. This is also the reason why GR needs the Komar integrals and ADM mass to define the energy of a spacetime. Such quantities are defined by an integral over the asymptotic boundary of a spacetime.

For an asymptotically AdS space, one similarly defines a stress tensor at the boundary of the spacetime. This was shown in [19]. The AdS stress tensor is related to the expectation value of the CFT stress tensor which lives on the boundary. This relation is simply given by

$$\langle T_{\mu\nu}^{\text{CFT}}\rangle = T_{\mu\nu}^{\text{grav}} \tag{4.23}$$

where $T_{\mu\nu}^{\text{grav}}$ is called the holographic stress tensor.

The holographic stress tensor usually diverges at the AdS boundary. The AdS/CFT correspondence solves this problem since the divergences can be related to ultraviolet divergences in the CFT. As in usual QFTs, we can take care of ultraviolet divergences by adding local counterterms to the CFT action. From the AdS perspective, these counterterms live on the boundary and will only depend on the boundary geometry. We then obtain a renormalized stress tensor for the AdS space. This procedure is called holographic renormalization.

From the holographic stress tensor, one can define a conserved charge [20]. Let S be a spacelike surface at the boundary of the spacetime and let u^μ define the local flow of time. If k^μ is a Killing vector which generates an isometry of the boundary geometry, the conserved charge is given by

$$Q_\zeta = \int_S d^{d-1}x\sqrt{\gamma}u^\mu T_{\mu\nu}^{\text{grav}}k^\nu = \int_S d\Sigma^\mu T_{\mu\nu}^{\text{grav}}k^\nu \tag{4.24}$$

where γ denotes the induced boundary metric on S. The conserved charge associated with time translation ($k^\mu \sim u^\mu$) defines the mass of the AdS space. Comparing with our result for the modular energy, we find that

$$E_B = \int_B d\Sigma^\mu\langle T_{\mu\nu}^{\text{CFT}}\rangle\zeta_B^\nu = \int_B d\Sigma^\mu T_{\mu\nu}^{\text{grav}}\zeta_B^\nu \equiv E_B^{\text{grav}} \tag{4.25}$$

where we chose the isometry to be generated by the conformal Killing vector ζ_B of the CFT and we integrate over the spatial spherical region B on the AdS boundary. Therefore, the modular energy E_B is holographically interpreted as the conserved charge associated with infinitesimal transformations generated by ζ_B at the AdS boundary.

Now we consider perturbations to our state, given by $\rho(\lambda)$. These will change the stress tensor in the CFT. The variation in the modular energy is simply given by the integral over the variation in the stress tensor. The same reasoning holds for the bulk space, where the metric is varied. Therefore, the equality $E_B = E_B^{\text{grav}}$ holds as long as the state $\rho(\lambda)$ has a dual metric. In terms of equations, we have that

$$\frac{d}{d\lambda}E_B = \int_B d\Sigma^\mu \langle \frac{d}{d\lambda}T^{\text{CFT}}_{\mu\nu}\rangle \zeta_B^\nu = \int_B d\Sigma^\mu \frac{d}{d\lambda}T^{\text{grav}}_{\mu\nu}\zeta_B^\nu = \frac{d}{d\lambda}E_B^{\text{grav}}. \tag{4.26}$$

Having identified the holographic expressions of S_B and E_B as S_B^{grav} and E_B^{grav}, we have the necessary ingredients to write down the bulk dual of the relative entropy. This gravitational quantity can be related to the gravitational equations of motion of the bulk space, which is covered in Chapter 5.

Chapter 5

Gravitation in Anti de Sitter spacetime

In this chapter we will study the gravitational bulk theory. By now we have translated the CFT modular energy and entanglement entropy into quantities that are defined in the bulk. What is left to understand is how they are related to the gravitational equations of motion. The goal of this chapter is to express the bulk dual of the relative entropy in terms of the gravitational equations of motion.

We will proceed using a formalism introduced by Robert Wald. It provides a description of general diffeomorphism invariant gravity theories. Using this formalism, Wald proved the first law of black hole mechanics for general theories and identified the black hole entropy as the Noether charge related to diffeomorphism invariance [11]. Since the entanglement entropy of our setup is given by the bulk black hole entropy, we can use these results for our present purposes.

In the first section we will introduce the formalism and use it to derive an identity which was found by Hollands and Wald in [12]. It will serve as our basis to connect the CFT relative entropy to the AdS equations of motion. To utilize this identity, we need to fix a specific gauge called the Hollands-Wald gauge. This gauge was described in [12] and will be explained in Section 5.3. Putting everything together we obtain the gravitational dual of the relative entropy. It is given by an expression involving the equations of motion and the presymplectic form, which will be defined below.

5.1 Wald's formalism

We begin by introducing Wald's formalism. Consider a general diffeomorphism-invariant gravitational Lagrangian. We will view it as a $(d+1)$-form. This will be convenient when making contact with our AdS space later on. In the following derivations all quantities are expressed as differential forms, which are written in bold-face font. Our theory is described by the action $S_{\text{bulk}} = \int \boldsymbol{L}(\phi)$, where ϕ denotes the fields in the theory, including the metric. To be precise, we define

$$\begin{aligned} \boldsymbol{L} &= \mathcal{L}\boldsymbol{\epsilon}, \\ \boldsymbol{\epsilon} &= \frac{1}{(d+1)!}\sqrt{-g}\epsilon_{a_1 a_2 \ldots a_{d+1}} dx^{a_1} \wedge \ldots \wedge dx^{a_{d+1}} \end{aligned} \tag{5.1}$$

with $\mathcal{L}$ the usual Lagrangian density and $\boldsymbol{\epsilon}$ the volume form. Consider variations of $\boldsymbol{L}$ under smooth, one-parameter variations of the fields $\phi(\lambda)$. We adapt the notation of [12],

$$\delta \boldsymbol{L} = \tfrac{d}{d\lambda}\boldsymbol{L}|_{\lambda=0}, \quad \delta\phi = \tfrac{d}{d\lambda}\phi(\lambda)|_{\lambda=0}, \quad \text{etc.} \tag{5.2}$$

to denote the first order variations of quantities. When a formula holds for general variations λ, we will keep the notation $\frac{d}{d\lambda}$.

Varying $\boldsymbol{L}$ with respect to the fields yields

$$\tfrac{d}{d\lambda}\boldsymbol{L} = \boldsymbol{E}_\phi \tfrac{d}{d\lambda}\phi + d\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi) \tag{5.3}$$

in which $\boldsymbol{E}_\phi$ denotes the equations of motion for the fields ϕ and $\boldsymbol{\Theta}$ is called the symplectic potential current density. It arises upon integrating the Lagrangian by parts to find the equations of motion, and it consists of boundary terms. Furthermore, the d in $d\boldsymbol{\Theta}$ is an exterior derivative. When minimizing the action, the boundary terms usually vanish due to asymptotic conditions of the fields. We will keep these terms as they play a role in the definition of the Noether current.

Using Noether's theorem, it is possible to write down conserved currents and charges. The symmetries we will consider are diffeomorphisms. Since our theory is diffeomorphism invariant, such a transformation will change the Lagrangian by a total derivative. Let us define the vector field X as the generator of the diffeomorphism. The variation of the fields is then given by their Lie derivative along the direction of X, which we will denote $\mathcal{L}_X\phi$. As an example, the metric will vary as $\mathcal{L}_X g_{\mu\nu} = 2\nabla_{(\mu}X_{\nu)}$. If this vanishes, it means the transformation is an isometry of the metric. In this case X is a Killing vector and the Lie derivative of the metric reproduces the Killing equation. A useful relation for the Lie derivative is Cartan's formula

$$\mathcal{L}_X = i_X d + d i_X. \tag{5.4}$$

Here i_X denotes the interior product with X. Using this formula we can express the variation of $\boldsymbol{L}$ as

$$\mathcal{L}_X \boldsymbol{L} = d(i_X \boldsymbol{L}). \tag{5.5}$$

This is because $\boldsymbol{\epsilon}$ is a top form, such that $d\boldsymbol{L} = 0$. The Noether current form related to the diffeomorphism is defined as

$$\boldsymbol{J}_X = \boldsymbol{\Theta}(\phi, \mathcal{L}_X\phi) - i_X\boldsymbol{L}. \tag{5.6}$$

The conservation of this current can be made explicit using equations (5.3) and (5.5)

$$\begin{aligned} d\boldsymbol{J}_X &= d\boldsymbol{\Theta}(\phi, \mathcal{L}_X\phi) - d(i_X\boldsymbol{L}) \\ &= \mathcal{L}_X\boldsymbol{L} - \boldsymbol{E}_\phi\mathcal{L}_X\phi - \mathcal{L}_X\boldsymbol{L} \\ &= -\boldsymbol{E}_\phi\mathcal{L}_X\phi. \end{aligned} \tag{5.7}$$

Indeed, the Noether current form is closed when on shell, for all X. This implies conservation because $d\boldsymbol{J}_X = 0$ implies $\nabla_\mu J^\mu_X = 0$ via the Hodge star, with J^μ_X the Noether current. Consequently there exists a form $\boldsymbol{Q}_X$ locally constructed from ϕ and X such that when ϕ satisfies the equations of motion, we have $\boldsymbol{J}_X = d\boldsymbol{Q}_X$ [11]. The form $\boldsymbol{Q}_X$ is referred to as the Noether charge form. It was shown by [21] that we can write an off shell relation for $\boldsymbol{J}_X$,

$$\boldsymbol{J}_X = d\boldsymbol{Q}_X + \boldsymbol{C}_X, \tag{5.8}$$

where $\boldsymbol{C}_X$ contains constraints which must vanish on shell. A proof of this relation is included in Appendix D. When the bulk theory is GR, we have that $\boldsymbol{C}_X \sim \boldsymbol{E}_\phi$.

Using equation (5.6) we can express the variation of the Noether current as

$$\begin{aligned}\tfrac{d}{d\lambda}\boldsymbol{J}_X &= \tfrac{d}{d\lambda}(\boldsymbol{\Theta}(\phi, \mathcal{L}_X\phi) - i_X\boldsymbol{L})\\ &= \tfrac{d}{d\lambda}\boldsymbol{\Theta}(\phi, \mathcal{L}_X\phi) - i_X d\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi) - i_X\boldsymbol{E}_\phi\tfrac{d}{d\lambda}\phi\\ &= \tfrac{d}{d\lambda}\boldsymbol{\Theta}(\phi, \mathcal{L}_X\phi) - \mathcal{L}_X\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi) + d(i_X\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi)) - i_X\boldsymbol{E}_\phi\tfrac{d}{d\lambda}\phi\\ &= \boldsymbol{\omega}(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_X\phi) + d(i_X\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi)) - i_X\boldsymbol{E}_\phi\tfrac{d}{d\lambda}\phi.\end{aligned} \tag{5.9}$$

In the second line, we used equation (5.3) and in the third line we used Cartan's formula. The form $\boldsymbol{\omega}$ is called the symplectic current form. It is introduced in Section 5.2 below. Combining equations (5.8) and (5.9) yields

$$d\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \boldsymbol{\omega}(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_X\phi) - (\tfrac{d}{d\lambda}\boldsymbol{C}_X + i_X\boldsymbol{E}_\phi\tfrac{d}{d\lambda}\phi), \tag{5.10}$$

where

$$\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \tfrac{d}{d\lambda}\boldsymbol{Q}_X - i_X\boldsymbol{\Theta}(\phi, \tfrac{d}{d\lambda}\phi). \tag{5.11}$$

Equation (5.10) will turn out to be very useful, since $\boldsymbol{\chi}$ is related to S_B^{grav} and E_B^{grav}. By showing this, we will have explicitly established the link between the relative entropy and the gravitational field equations.

5.2 The presymplectic form

The formalism we are introducing originates from a paper by Lee and Wald from 1990 [22]. In this work, the relation between local symmetries in the Lagrangian formalism and the corresponding constraints in the Hamiltonian formalism was studied. This allows for determination of the phase space for general physical field theories. In this section we will briefly discuss this paper, which defines the forms $\boldsymbol{\omega}$ and $\boldsymbol{\Theta}$. From $\boldsymbol{\omega}$ we will define a conserved quantity, following more recent work by Wald [23].

We consider a theory of dynamical fields ϕ, which includes a Lorentzian metric. One can define the space of field configurations $\mathcal{F}$ which contains all the "kinematically allowed" configurations. The dynamically allowed configurations are the points in $\mathcal{F}$ which are solutions to the equations of motions of the theory. We will denote this subspace as $\bar{\mathcal{F}}$.

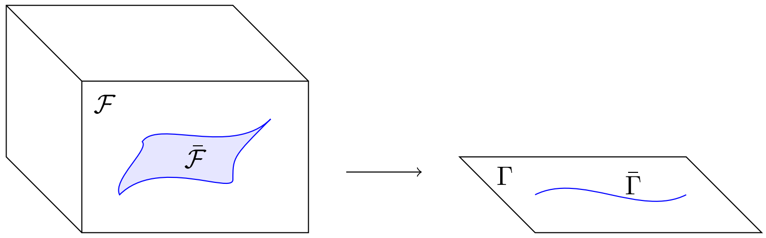

Figure 5.1 – A representation of $\mathcal{F}$ and Γ. The field configuration space $\mathcal{F}$ is larger than the phase space Γ due to the presence of local symmetries. A map from $\mathcal{F}$ to Γ must therefore reduce $\mathcal{F}$, by identifying all elements which lie within the same gauge orbit. This procedure is explained in detail by [22]. Furthermore, $\bar{\mathcal{F}}$ and $\bar{\Gamma}$ denote the solutions to the equations of motion.

When we describe our theory with the Hamiltonian formalism, we are interested in the phase space of the theory, which we denote Γ. The points in phase space which satisfy the equations of motion describe a subspace of Γ which we denote $\bar{\Gamma}$. See Figure 5.1.

Hamiltonian mechanics is naturally formulated in the language of symplectic geometry. Therefore, to obtain a relation between the Lagrangian and Hamiltonian formalisms, we would like to define a symplectic form constructed from the fields in our theory.

Let $\phi(\lambda_1, \lambda_2)$ denote a two-parameter family of field configurations. We define the presymplectic current form $\boldsymbol{\omega}$ as

$$\boldsymbol{\omega}(\phi; \tfrac{\partial}{\partial\lambda_1}\phi, \tfrac{\partial}{\partial\lambda_2}\phi) = \tfrac{\partial}{\partial\lambda_2}\boldsymbol{\Theta}(\phi, \tfrac{\partial}{\partial\lambda_1}\phi) - \tfrac{\partial}{\partial\lambda_1}\boldsymbol{\Theta}(\phi, \tfrac{\partial}{\partial\lambda_2}\phi). \tag{5.12}$$

It is a local function of a field configuration ϕ and two perturbations $\frac{\partial}{\partial\lambda_1}\phi$ and $\frac{\partial}{\partial\lambda_2}\phi$. Let Σ be a Cauchy surface (this Σ is not to be mistaken with the space bounded by B and $\tilde{B}$ from Chapter 4.). Consider now first order variations of the fields, $\delta\phi$. By integrating $\boldsymbol{\omega}$ we can define the presymplectic form

$$W_\Sigma(\phi; \delta_1\phi, \delta_2\phi) = \int_\Sigma \boldsymbol{\omega}. \tag{5.13}$$

It is a 2-form on $\mathcal{F}$ which maps two perturbations of the field configuration into the real numbers. The presymplectic form depends on the choice of Σ. However, in the case that the perturbations satisfy the linearized field equations, W_Σ does not depend on the choice of the Cauchy surface provided that Σ is compact or suitable asymptotic conditions are imposed on the dynamical fields [22, 23].

Assuming these conditions, W_Σ is also closed. This follows from the definition of $\boldsymbol{\omega}$. By taking a second variation of $\boldsymbol{L}$ we get

$$\delta_2\delta_1\boldsymbol{L} = \delta_2\boldsymbol{E}_\phi\delta_1\phi + \boldsymbol{E}_\phi\delta_2\delta_1\phi + d(\delta_2\boldsymbol{\Theta}(\phi, \delta_1\phi)). \tag{5.14}$$

A similar expression holds when reversing the order of the variations. Substracting these gives

$$0 = (\delta_2\delta_1 - \delta_1\delta_2)\boldsymbol{L} = \delta_2\boldsymbol{E}_\phi\delta_1\phi - \delta_1\boldsymbol{E}_\phi\delta_2\phi + d\boldsymbol{\omega}(\phi; \delta_1\phi, \delta_2\phi). \tag{5.15}$$

The antisymmetrized variations vanish by mixing of partial derivatives. If the variations satisfy the linearized equations, the first two terms on the right hand side vanish and we find that $\boldsymbol{\omega}$ is closed. This implies the presymplectic form is closed as well.

We have argued that W_Σ is a closed 2-form. If it is non-degenerate as well, W_Σ is a symplectic form and can be used to define a Hamiltonian. This is generally not true. If the Lagrangian possesses local symmetries, W_Σ will be degenerate. In this case different field configurations from $\mathcal{F}$ describe the same physical system, hence the same point in the phase space Γ. The space $\mathcal{F}$ is too large to be related to Γ. This is why W_Σ is called the presymplectic form instead. In order to define a proper symplectic form we must factor the space $\mathcal{F}$ by its gauge orbits. We then obtain a non-degenerate symplectic form.

For our purposes it will not be necessary to find the explicit map from $\mathcal{F}$ to Γ. It will be easier to avoid these complications and work on the field configuration space $\mathcal{F}$ with the degenerate presymplectic form W_Σ. A vector field X will induce a field variation $\mathcal{L}_X\phi$ on the dynamical fields. It is argued in [23] that if a Hamiltonian conjugate to X exists, it must satisfy the relation

$$\frac{d}{d\lambda}H_X = W_\Sigma(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_X\phi) = \int_\Sigma \boldsymbol{\omega}(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_X\phi). \tag{5.16}$$

In this paper it is also reasoned that this Hamiltonian defines a conserved quantity associated with X. This conserved quantity is related to asymptotic transformations generated by X. For instance, by choosing X to be the vector field which generates asymptotic time translation, we can define energy of the spacetime under consideration. In GR, the energy is given by the ADM mass. The agreement of these quantities was shown explicitly in [24].

The reason that the Hamiltonian is related to asymptotic conserved quantities is because it can be written as a boundary term. Using equation (5.10) we find,

$$\frac{d}{d\lambda}H_X = \int_\Sigma d\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) + \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_X + i_X\boldsymbol{E}_\phi\tfrac{d}{d\lambda}\phi). \tag{5.17}$$

If ϕ solves the equations of motions and the first order variation solves the linearized equations, the second term on the right hand side vanishes. The Hamiltonian reduces to a boundary term by virtue of Stokes theorem,

$$\frac{d}{d\lambda}H_X = \int_{\partial\Sigma} \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi). \tag{5.18}$$

For asymptotically flat spacetimes the Hamiltonian is then determined by an integral over a codimension-2 spherical surface at spatial infinity. For asymptotically AdS spaces this works differently since there is a boundary. The integration domain is then the AdS boundary.

5.3 Application to the perturbed AdS black hole

To proceed with our derivation of the gravity dual of the relative entropy, we want to make contact between $\boldsymbol{\chi}$, and S_B^{grav} and E_B^{grav}. This is achieved by applying the formalism of Sections 5.1 and 5.2 to our AdS black hole setup. Recall that the exterior of our black hole spacetime is given by Σ, which is bounded by the sphere B and the horizon $\tilde{B}$.

First, consider a diffeomorphism invariant theory described by a gravitational Lagrangian with a stationary black hole solution. Choose ξ^a to be the Killing field which generates the horizon. The surface gravity of the black hole horizon is defined by

$$\xi^a \nabla_a \xi^b = \kappa \xi^b. \tag{5.19}$$

A Killing horizon is called a bifurcate Killing horizon if it has non-vanishing surface gravity. In this case, the past and future horizons intersect at the so called bifurcation surface, at which the horizon generator $\xi = 0$ opposed to being simply null. The theories of our interest have such a surface. Physically, this restricts our discussion to non-extremal black holes.

The horizon $\tilde{B}$ is also generated by a Killing vector ξ^a. This vector is given by

$$\xi_B = -\frac{2\pi}{R}(t - t_0)[z\partial_z + (x^i - x_0^i)\partial_i] + \frac{\pi}{R}[R^2 - z^2 - (t - t_0)^2 - (x^i - x_0^i)^2]\partial_t \tag{5.20}$$

for the sphere B at coordinates (t_0, x_0^i) and radius R. One can verify that ξ_B vanishes at the horizon $\tilde{B}$ at $t = t_0$. This means our black hole horizon is a bifurcation surface. Note that the vector ξ_B is not a lucky guess, it is actually related to the generator of time translation parametrized by τ in the global AdS coordinates. It is the bulk generalization of the conformal Killing vector ζ_B we discussed in Chapter 3. Taking the asymptotic limit we indeed find that $\lim_{z\to 0} \xi_B \to \zeta_B$. Furthermore, ξ_B is normalized to have a surface gravity of $\kappa = 2\pi$ which will be useful shortly.

Recall equation (5.10). We will integrate it over our black hole exterior Σ and choose $X = \xi_B$. This yields

$$\int_\Sigma d\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \int_\Sigma \boldsymbol{\omega}(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_{\xi_B}\phi) - \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_\phi \tfrac{d}{d\lambda}\phi). \tag{5.21}$$

Focus on the left hand side of this equation. Since Σ is bounded by B and $\tilde{B}$, by Stokes theorem we can express this integral as

$$\int_\Sigma d\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \int_{\partial\Sigma} \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \int_B \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) - \int_{\tilde{B}} \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi). \tag{5.22}$$

The minus sign comes from the orientation of $\tilde{B}$ relative to B. At the bifurcation surface $\tilde{B}$, the vector ξ_B vanishes by definition. The restriction of $\boldsymbol{\chi}$ to this surface is therefore given by $\boldsymbol{\chi}(\phi, \frac{d}{d\lambda}\phi)|_{\tilde{B}} = \frac{d}{d\lambda}\boldsymbol{Q}_{\xi_B}$. The integral over $\tilde{B}$ now looks a lot like the Wald entropy defined in equation (4.19), the difference being that we are dealing with a variation of $\boldsymbol{Q}$ instead.

We would like to show that the integral over the variation of $\boldsymbol{Q}$ is equal to the variation of the entanglement entropy $\frac{d}{d\lambda}S_B^{\text{grav}}$. This is not a trivial statement since variation of the metric will cause a perturbation of the black hole geometry. This changes the horizon $\tilde{B}$

such that ξ_B will cease to generate the horizon. It was shown in [11] that the statement is valid at least for first order perturbations,

$$\delta S_B^{\text{grav}} = \int_{\tilde{B}} \delta \boldsymbol{Q}_{\xi_B}. \tag{5.23}$$

It was argued in [12] that by fixing an appropriate gauge, the equality can be generalized to arbitrary λ. This is referred to as the Hollands-Wald gauge. Without loss of generality, it imposes two conditions on our system. First of all, the coordinate location of the perturbed bifurcation surface $\tilde{B}(\lambda)$ will be fixed at the location of $\tilde{B}(0)$, rendering it independent of λ. Secondly, the Killing vector ξ_B as defined in (5.20) will still obey the Killing equation at this surface in the perturbed metric, $\mathcal{L}_{\xi_B} g(\lambda)|_{\tilde{B}} = 0$. Instead of working out the details, we will simply assume this gauge choice. More information can be found in [12] and [25]. Using the two conditions imposed by the gauge, we derive

$$\frac{d}{d\lambda} S_B^{\text{grav}} = \frac{d}{d\lambda} S_{\text{Wald}} = \int_{\tilde{B}} \frac{d}{d\lambda} \boldsymbol{Q}_{\xi_B} = \int_{\tilde{B}} \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi). \tag{5.24}$$

This completes our identification of the entanglement entropy.

The integral of $\boldsymbol{\chi}$ over B is equal variation of the conserved quantity related to asymptotic transformations induced by ξ_B. This can be seen by comparison with equation (5.18). The boundary of our Cauchy slice is the sphere B. We have

$$\frac{d}{d\lambda} H_{\xi_B} = \int_B \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi). \tag{5.25}$$

We can make contact with E_B^{grav} defined in Chapter 4 by noting that the asymptotic limit of ξ_B corresponds to the conformal Killing vector ζ_B. Recall that E_B^{grav} is the conserved quantity related to infinitesimal transformations of ζ_B. We then conclude that H_{ξ_B} and E_B^{grav} must describe the same conserved quantity. In terms of equations, we have

$$\int_B \boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \tfrac{d}{d\lambda} H_{\xi_B} = \tfrac{d}{d\lambda} E_B^{\text{grav}}. \tag{5.26}$$

Let us summarize our results. By means of the Hollands-Wald gauge we have proven that

$$\int_\Sigma d\boldsymbol{\chi}(\phi, \tfrac{d}{d\lambda}\phi) = \frac{d}{d\lambda}\left(E_B^{\text{grav}} - S_B^{\text{grav}}\right) \tag{5.27}$$

for our specific AdS black hole setup. In Chapter 4 we have shown that E_B^{grav} and S_B^{grav} are equivalent to the modular energy and entanglement entropy in the boundary CFT. The right-hand side of equation (5.27) is equal to the relative entropy of the CFT. Let us show this by writing out the λ dependencies explicitly:

$$\begin{aligned}\frac{d}{d\lambda} S(\rho(\lambda)|\rho(0)) &= \frac{d}{d\lambda}(\Delta E_B - \Delta S_B) \\ &= \frac{d}{d\lambda}(\Delta E_B^{\text{grav}} - \Delta S_B^{\text{grav}}) \\ &= \frac{d}{d\lambda}(E_B^{\text{grav}}(g(\lambda)) - S_B^{\text{grav}}(g(\lambda))).\end{aligned} \tag{5.28}$$

The last equality follows since $\Delta E_B^{\text{grav}} = E_B^{\text{grav}}(g(\lambda)) - E_B^{\text{grav}}(g(0))$ and similarly for the entropy. Plugging this into equation (5.21) we obtain the relation

$$\frac{d}{d\lambda} S(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(\phi; \tfrac{d}{d\lambda}\phi, \mathcal{L}_{\xi_B}\phi) - \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_\phi \tfrac{d}{d\lambda}\phi). \tag{5.29}$$

This finalizes our identification of the CFT relative entropy with the dual AdS equations of motion. We find that the relative entropy between two states in our sphere B is related to the symplectic current form as well as the gravitational equations. By considering first order variations in λ, we can show that the first law of entanglement entropy implies the linearized Einstein equation in the bulk. This will be discussed in Chapter 6. Attempts to derive the full non-linear Einstein equations will be discussed in Chapter 7.

Chapter 6

Linearized gravity from the first law of entanglement entropy

In this chapter we will establish the relationship between the linearized gravitational equations of motion and the relative entropy. It can be shown that spacetimes dual to small perturbations of the CFT vacuum state must obey Einstein's equations linearized around pure AdS, given that this bulk spacetime computes the correct CFT entanglement entropy using the Ryu-Takayanagi prescription. In other words, if we succeed in identifying a dual asymptotically AdS spacetime with this property, the AdS metric perturbation solves the linearized Einstein equations. In this chapter we will prove this assertion. It was first derived by [26], and worked out in more detail by [7].

The dual asymptotically AdS geometry was identified in Chapters 4 and 5. What is left to prove is that the first order perturbation δg away from pure AdS now necessarily satisfies the linearized equations of motion. Recall the final result of Chapter 5, equation (5.29). For simplicity, assume the only dynamical field is the metric.

$$\frac{d}{d\lambda}S(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \tfrac{d}{d\lambda}g(\lambda), \mathcal{L}_{\xi_B}g(\lambda)) - \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_g \tfrac{d}{d\lambda}g(\lambda)) \tag{6.1}$$

where Σ denotes the spatial surface bounded by the sphere B and the bifurcation surface $\tilde{B}$. All quantities on the right hand side now depend on the metric only and $\boldsymbol{E}_g$ denotes the equations of motion obtained by varying with respect to the metric.

By evaluating this expression for $\lambda = 0$ we obtain the first order variation. We argued in Section 2.3 that the variation of the relative entropy is given by $\delta E - \delta S = 0$, which we named the first law of entanglement entropy. To first order in λ, the symplectic current form vanishes as well. This is due to the fact that ξ_B generates an isometry of the unperturbed metric $g(0)$, so $\mathcal{L}_{\xi_B}g(0) = 0$. This implies $\boldsymbol{\omega} = 0$, which can be seen most easily from the definition of the symplectic current form in [22]. Finally, $\boldsymbol{E}_g(g(0)) = 0$ since the AdS vacuum is a solution to the Einstein equations. Equation (6.1) now reduces to the simple identity

$$\int_\Sigma \delta\boldsymbol{C}_{\xi_B} = 0. \tag{6.2}$$

The form $\boldsymbol{C}_{\xi_B}$ contains constraint equations which vanish for fields which satisfy the equations of motion. For general relativity, it is explicitly given by

$$\boldsymbol{C}_{\xi_B} = 2\xi_B^a E_{ab}^g \boldsymbol{\epsilon}^b \tag{6.3}$$

where E_{ab}^g denotes the gravitational equations of motion, in this case the Einstein tensor, and we define $\boldsymbol{\epsilon}_b = \frac{1}{d!}\sqrt{-g}\epsilon_{ba_2\ldots a_{d+1}} dx^{a_2} \wedge \ldots \wedge dx^{a_{d+1}}$. For more information about the constraints, see Appendix D. Our domain of integration Σ is defined at a fixed time $t = 0$, such that the volume element $\boldsymbol{\epsilon}_b$ on this hypersurface only has a t-component. We find

$$\int_\Sigma \xi_B^t \delta E_{tt}^g \boldsymbol{\epsilon}^t = 0 \tag{6.4}$$

where δE_{ab}^g denotes the linearized Einstein equation. To prove that these are satisfied, we need to show that $\delta E_{ab}^g = 0$. Multiply the integral by R and take a derivative with respect to R. We obtain

$$\frac{d}{dR}\int_{\Sigma(R,x_0^i)} R\xi_B^t \delta E_{tt}^g \boldsymbol{\epsilon}^t = 0 \tag{6.5}$$

where we have written the R dependence of Σ explicitly. This expression can be simplified using the Leibniz integral rule for multidimensional integrals as

$$\int_{\tilde{B}} (R\xi_B^t)\delta E_{tt}^g \hat{r}\cdot\boldsymbol{\epsilon}^t + 2\pi R\int_\Sigma \delta E_{tt}^g \boldsymbol{\epsilon}^t = 0. \tag{6.6}$$

Since ξ_B^t vanishes on the bifurcation surface $\tilde{B}$, we find the result

$$\int_\Sigma \delta E_{tt}^g \boldsymbol{\epsilon}^t = 0 \tag{6.7}$$

which holds for any surface $\Sigma(R, x_0^i)$. Thus, the integral must vanish for any entanglement sphere (R, x_0^i) we choose. This implies $\delta E_{tt}^g = 0$.

So far we have found one component of the linearized Einstein equations. We can repeat the whole argument for a different Lorentz frame since the CFT is Lorentz invariant. By demanding $\delta E_{tt}^g = 0$ to hold in any frame, we find the other components of the equation. The covariant way to do this is by defining a velocity vector u^μ which describes the frame of reference. Our equation now becomes $u^\mu u^\nu \delta E_{\mu\nu}^g = 0$ which must hold for each choice of u^μ. Note that μ and ν are indices of the CFT spacetime, which proves that the linearized equations of motion are satisfied in all boundary directions.

The remaining equations $\delta E_{\mu z}^g = 0$ and $\delta E_{zz}^g = 0$ need some more work. They can be argued to hold from the initial value formulation of gravity. This construction is rather involved, and beyond the scope of this thesis. The interested reader may refer to Wald's book on GR for an excellent review on this topic [27]. The argument why the remaining equations are satisfied is given in [7]. For the sake of completion, we present a heuristic explanation to make the statement plausible.

Any proper physical theory should have an initial value formulation. This means that by satisfying initial conditions at some point in time, we can predict what will happen at later (or earlier) times. The question whether a theory has a well posed initial value formulation is closely related to the mathematical structure of the differential equations which describe the dynamics, the equations of motion. For instance, the equations must have a unique solution, given initial data. Other physical constraints such as causality may be demanded as well.

Theories which exhibit gauge symmetry need special attention. This is because the evolution of the system doesn't have a unique mathematical solution. However, such theories can have a well posed physical initial value formulation if the mathematical solutions related by gauge transformations describe the same unique physical outcome. This is true if the initial data satisfies the initial value constraints. These are derived from the theory of interest. As a concrete example, for Maxwell theory, this constraint is given by the well known formula $\nabla \cdot \vec{E} = 0$.

In [27], the initial value constraints for GR are computed explicitly. They consist of two equations, which are both linear in the Einstein tensor. The Einstein tensor also satisfies the Bianchi identity $\nabla^a G_{ab} = 0$. It is possible to show that if the spatial components of the Einstein equation are satisfied and the constraints are satisfied initially, the constraints must be satisfied at all times. In [7] the constraints are chosen to be the z-components of the Einstein tensor. By showing they vanish at $z = 0$, they must be satisfied everywhere. It is therefore concluded that $\delta E^g_{\mu z} = \delta E^g_{zz} = 0$.

Using this result, we have argued that all components of the Einstein equation vanish. Therefore, the equations of motion are satisfied. We conclude that if we can identify an asymptotically AdS spacetime which correctly computes the entanglement entropy of a perturbed CFT vacuum state, the fluctuations around pure AdS solve the linearized Einstein equations. This is related to entanglement via the first law of entanglement entropy in the CFT.

Before we continue to the next chapter, we want to make one more remark. The relation between the relative entropy and gravity can also be reversed. Starting from equation (6.1), assume the perturbed AdS metric $g(\lambda)$ to be a solution to the Einstein equations. We then obtain a relation between the relative entropy of the dual CFT and the symplectic current form

$$\frac{d}{d\lambda} S(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \tfrac{d}{d\lambda} g(\lambda), \mathcal{L}_{\xi_B} g(\lambda)). \tag{6.8}$$

By setting $\lambda = 0$, we obtain $\delta E_B - \delta S_B = 0$. This implies that in the dual CFT the first law of entanglement must be satisfied. We conclude that if we have an asymptotically AdS space which solves the Einstein equations up to linear order in the perturbation around pure AdS, and we have identified a dual CFT, this CFT satisfies the first law of entanglement. The point we are trying to make is that gravity not just emerges from entanglement, but they are equivalent in a holographic sense. They can be regarded as two sides of the same coin.

Chapter 7

Beyond the linearized Einstein equations

Our understanding of the connection between quantum entanglement and gravity is far from complete. For first order perturbations on the CFT vacuum, we have derived the relation between entanglement entropy and the corresponding bulk equations of motion. A natural question to ask is whether this holds for all orders in the perturbation. Are finite variations related to the full Einstein equations? This question has gotten attention lately and is being investigated by different research groups in the field.

In Section 7.1 we will discuss a paper written by Eunseok Oh, I.Y. Park and Sang-Jin Sin [1]. In this work it is derived that the non-linear Einstein equation is implied by finite variations of the CFT state. The argument is based on the same formalism and setup as the present work. We will explain why their argument, in our opinion, does not prove the claimed result.

We continue in Section 7.2 to discuss other attempts to improve on the linearized result. An important result was achieved by [28], in which it is shown that the Einstein equations are satisfied up to and including second order. We also discuss a recent relevant paper from Jacobsen [29], which relates entanglement to the Einstein equations in a different setting.

7.1 Complete Einstein equation from the generalized First Law of Entanglement

In September 2017, a paper was published which claims to have derived that the full Einstein equation is equivalent to the dual of an entanglement relation [1]. This result generalizes the work of [7] to general perturbations. We first explain their arguments, after which we comment on the result, which is necessary since we are not truly convinced by their method.

The work is based on the same considerations we used to find the linearized equations. We consider a spherical region in the CFT vacuum and apply perturbations to the state. By means of holography, the relative entropy in the CFT is related to the presymplectic current form and the gravitational equations in the bulk. We start from equation (5.21), which we reproduce here for convenience

$$\int_\Sigma d\boldsymbol{\chi}(g, \tfrac{d}{d\lambda}g) = \int_\Sigma \boldsymbol{\omega}(g; \tfrac{d}{d\lambda}g, \mathcal{L}_{\xi_B}g) - \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_g \tfrac{d}{d\lambda}g). \tag{7.1}$$

This is a general identity that holds for any metric $g(\lambda)$, both on and off shell. For simplicity, we consider the metric to be the only dynamical field.

The goal is to show that the Einstein equations are equivalent to the gravity dual of the relative entropy. The argument starts by considering physical metrics which satisfy the equations of motion. In this case we have $\boldsymbol{E}_g = \boldsymbol{C}_{\xi_B} = 0$. As we have shown, the left hand side of the equation can be expressed as the difference in modular energy and entanglement entropy. We obtain

$$\frac{d}{d\lambda}(\Delta E_B^{\text{grav}} - \Delta S_B^{\text{grav}}) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \tfrac{d}{d\lambda}g(\lambda), \mathcal{L}_{\xi_B}g(\lambda)). \tag{7.2}$$

From this equation the gravity dual of the relative entropy is identified as

$$\frac{d}{d\lambda}S^{\text{grav}}(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \tfrac{d}{d\lambda}g(\lambda), \mathcal{L}_{\xi_B}g(\lambda)). \tag{7.3}$$

Using this relation, equation (7.1) reduces to

$$S^{\text{grav}}(\rho(\lambda)|\rho(0)) = \Delta E_B^{\text{grav}} - \Delta S_B^{\text{grav}}. \tag{7.4}$$

The conclusion of this reasoning is that the on-shell version of equation (7.1) gives the gravitational dual of equation (2.7), the relative entropy in terms of the modular energy and entanglement entropy.

Now consider the reversed argument. Let $g(\lambda)$ be a metric which satisfies equation (7.4). Plugging this into relation (7.1) we obtain

$$\frac{d}{d\lambda}S^{\text{grav}}(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \tfrac{d}{d\lambda}g(\lambda), \mathcal{L}_{\xi_B}g(\lambda)) - \int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_g \tfrac{d}{d\lambda}g). \tag{7.5}$$

Using the definition of the gravitational dual of the relative entropy, (7.3), we find

$$\int_\Sigma (\tfrac{d}{d\lambda}\boldsymbol{C}_{\xi_B} + i_{\xi_B}\boldsymbol{E}_g \tfrac{d}{d\lambda}g) = 0. \tag{7.6}$$

It can be shown that this implies the full equations of motion to be satisfied, in a similar fashion as the linearized case in Chapter 5. We conclude that a metric $g(\lambda)$ which satisfies the equation for the gravitational dual of the relative entropy is always on shell.

Since the argument holds in both directions, satisfying the equations of motion is equivalent to satisfying equation (7.4). The identifications in Chapter 4 then tell us this is equal to the CFT relative entropy equation (2.7). So, any metric $g(\lambda)$ which reproduces the entanglement of a perturbed CFT vacuum state must satisfy the full Einstein equation. This completes the proof.

We believe there is a flaw in this reasoning. The two proofs do not seem to be independent of each other. In equation (7.3), the gravitational dual to the relative entropy is defined for metrics on shell. This is allowed, since the proof assumes the equations of motion are solved by $g(\lambda)$. However, in the reversed argument, when considering metrics which solve the relative entropy equation, there is no reason equation (7.3) should still hold. After all, this identification was made after assuming the metric to be on shell. By using equation (7.3) as the definition of the dual relative entropy in the reversed argument, the authors implicitly force $g(\lambda)$ to be on shell. The Einstein equations are then satisfied trivially.

We stress that the question whether it is possible to relate the complete Einstein equation and entanglement entropy is still an open problem. In order to complete the proof one would need to show that for any metric which gives the right CFT entanglement entropy, the gravitational dual of the relative entropy is given by an integral over the symplectic current form. In other words, one must prove equation (7.3) to hold without constraining $g(\lambda)$ to be on shell beforehand. This has not yet been achieved.

7.2 Improving on the linearized result

We will now discuss some recent research relevant to our topic.

7.2.1 Nonlinear Gravity from Entanglement in Conformal Field Theories

This paper [28] was written in collaboration with Thomas Faulkner and Mark van Raamsdonk, who also contributed to the linear result as discussed in [7]. It demonstrates the emergence of the gravitational equations up to second order in the perturbation. We will now explain how this was achieved.

We start from the geometrical equation we derived in Chapters 2-5. Taking a derivative with respect to λ of equation (6.1) and setting $\lambda = 0$ we find

$$\delta^{(2)} S(\rho(\lambda)|\rho(0)) = \int_\Sigma \boldsymbol{\omega}(g(\lambda); \delta g(\lambda), \mathcal{L}_{\xi_B} \delta g(\lambda)) - \int_\Sigma \delta^{(2)} \boldsymbol{C}_{\xi_B} \tag{7.7}$$

where $\delta^{(2)} = \frac{d^2}{d\lambda^2}|_{\lambda=0}$. We got rid of the terms involving $\boldsymbol{E}_g$, because for $\lambda = 0$ both $\boldsymbol{E}_g = 0$ and $\delta^{(1)}\boldsymbol{E}_g = 0$. The last term on the right contains the second order Einstein tensor. The first term on the right hand side defines the canonical energy in a gravitational theory with respect to the timelike Killing vector ξ_B [12]. The main goal of [28] was to prove the equality of the first two terms, which implies the second order equations via $\delta^{(2)}\boldsymbol{C}_{\xi_B} = 0$. This is achieved by explicit calculation. Since it is very involved, we won't reproduce it here.

The conclusion is that any geometry $g(\lambda)$ which correctly gives the entanglement entropy for a CFT using the HRT proposal must solve the Einstein equations up to and including second order in the perturbation. This result is reassuring. It shows that the linearized equations were not the end of the story. This can be interpreted as a motivation researchers to find the full Einstein equation in a rigorous manner.

By considering perturbations order by order, we will not find the full nonlinear Einstein equations. However, by considering specific perturbations interesting holographic identifications can be made. The second order variation of the relative entropy must be a positive quantity, since the relative entropy it is monotonically increasing away from its reference state, in this case the vacuum $\rho(0)$. This in turn implies positivity of the canonical energy as defined by Wald [25, 11, 12]. Constraints like this one shed light on the conditions the AdS metric must satisfy to ensure the existence of a dual CFT of which the entanglement entropy is calculated using the HRT proposal.

7.2.2 Entanglement Equilibrium and the Einstein Equation

In 1995 a paper was published by Jacobson in which the Einstein equation is derived from the principles of thermodynamics [30]. It was motivated by the similarity of the laws of black hole mechanics and the laws of thermodynamics. This work can be seen as the starting point of the branch of physics which seeks to relate quantum entanglement to gravitational dynamics.

The more recent paper we will discuss here establishes a link between entanglement entropy and the full Einstein equation [29]. The main result is that the Einstein equation is implied by the so called "maximal vacuum entanglement hypothesis". To quote Jacobson:

> "*When the geometry and quantum fields are simultaneously varied from maximal symmetry, the entanglement entropy in a small geodesic ball is maximal at fixed volume.*"

This hypothesis is predicted from the assumption that the quantum vacuum has a finite and universal area density of entanglement entropy. The physical interpretation of this assumption is that the entanglement entropy of the vacuum is in an equilibrium condition similar to a thermal equilibrium. Assuming this hypothesis to hold, it can be derived that the metric must obey the Einstein equation.

The main derivation considers the entanglement entropy of regions separated by local causal horizons. It combines Jacobson's previous paper [30] and the work based on entanglement entropy and the AdS/CFT correspondence which we have reviewed in this thesis. However, Jacobson makes no use of holography whatsoever. The local causal horizons are analogous to the boundary $\partial\mathcal{D}$ of the domain of dependence $\mathcal{D}$ we defined in Chapter 3. It is not defined in a CFT, but in a more general spacetime where the metric is allowed to be dynamical. By variation of the metric away from the geometry that maximizes the entanglement entropy, the Einstein equation is derived.

The question is whether the assumption of the vacuum being in 'entanglement equilibrium' is valid. According to Jacobson, validation of this statement involves UV aspects of quantum gravity. It remains an open question.

Chapter 8

Outlook

Research on how entanglement and gravity are intertwined is still very young. Nevertheless a lot of progress has been made in a very short time. There is of course much more to be studied and there are many unexplored paths to take. As a consequence, there exist no complete comprehensive reviews on this topic yet. Our aim was to explain this topic in a pedagogical way, in as much detail as possible within the scope of the project. Needless to say, our presentation of the subject can still be improved. Still, we hope that this thesis can serve as a more easily accessible review on the relation of entanglement and gravity via holography.

There is much more left to learn about entanglement and gravity. We will now discuss some possible directions that could be studied. The first and most obvious result we would like to see is the generalization of [7] to the full Einstein equation. As we have explained in Chapter 7, the argument of [1] is in our opinion insufficient to prove this result. However we do believe it should be possible to relate the full Einstein equation to entanglement. This would tell us what set of conditions on the CFT is dual to the dynamics in an AdS spacetime governed by GR, teaching us a lot about the AdS/CFT correspondence.

Taking this reasoning a step further, one could try to apply the whole framework to general theories of gravity. Such theories are not described by the Einstein-Hilbert action, but can have any covariant Lagrangian constructed from contracted Riemann tensors. In our derivation of the Einstein equation, we made use of the HRT proposal. This formula only holds in general relativity, which is why the Einstein equation is obtained. Relating entanglement to the equations of motion of a general theory of gravity then comes down to finding the correct entanglement entropy functional. In [17], a formula is derived which applies to any gravitational theory. Aside from possible subtleties, it seems likely that by using this generalized entanglement functional, one could relate entanglement to the equations of motion of any gravitational theory.

All our considerations so far have focused on CFTs with a classical bulk description. It would be very interesting to study the relation between entanglement entropy in CFTs dual to gravity theories with quantum corrections. From the CFT perspective this is certainly possible since the first law of entanglement holds for any QFT. The first quantum corrections to the HRT proposal were derived in [31] and the bulk dual of the relative entropy in [32]. By extending the work in this direction we might learn how quantum entanglement constrains the quantum behavior of such theories.

All our considerations so far have focused on CFTs with a classical bulk description. It would be very interesting to study the relation between entanglement entropy in CFTs dual to gravity theories with quantum corrections. In the language of AdS/CFT, this amounts to relaxing the 't Hooft limit. The CFT Yang-Mills theory has a gauge group $SU(N)$ and a coupling constant g_{YM}. The coupling g_{YM} is related to the bulk string coupling g_s as $g_{YM}^2 \sim g_s$. The 't Hooft limit corresponds to $N \to \infty$ and $g_{YM} \to 0$ and selects a classical bulk dual. In order to allow for quantum corrections in the bulk, we need to consider small N and weak (non-zero) coupling. This regime of the AdS/CFT correspondence is not as well-established as the classical case, but let us assume that it is valid. Since entanglement is independent of any coupling constants, the entanglement entropy of a CFT can be computed for any limit of N and g_{YM}. On the gravity side we would need to generalize the HRT proposal. The first quantum corrections to the HRT proposal were derived in [31] and the bulk dual of the relative entropy in [32]. By extending the work in this direction we might learn how quantum entanglement constrains the quantum behavior of such theories.

In the long run, the hope is that by exploiting the relation between entanglement and gravity we can deduce information about the nature of quantum gravity. It could be that entanglement is the best way to learn about this theory. The question whether entanglement will contribute to finding a theory of quantum gravity is one that cannot be answered with certainty right now. Personally, I am very curious to see where the developments in this field of research will lead us in the future.

Appendix A

Relative entropy properties for a two-spin system

In this Appendix we will work out the properties of the relative entropy of a simple spin system. The system consists of two spin-$\frac{1}{2}$ particles which both assume an orientation. For simplicity, we assume both spins will point in the same direction. The reference state will be described by the density matrix

$$\sigma = a\left|\uparrow\right\rangle\left\langle\uparrow\right| + (1-a)\left|\downarrow\right\rangle\left\langle\downarrow\right| = \begin{pmatrix} a & 0 \\ 0 & 1-a \end{pmatrix} \tag{A.1}$$

where $0 < a < 1$. It is interpreted as a statistical ensemble of the up and down quantum states. We are interested in the relative entropy between this state and a state with a different distribution of up and down spins. We define

$$\rho = b\left|\uparrow\right\rangle\left\langle\uparrow\right| + (1-b)\left|\downarrow\right\rangle\left\langle\downarrow\right| = \begin{pmatrix} b & 0 \\ 0 & 1-b \end{pmatrix} \tag{A.2}$$

where $0 < b < 1$. Now we compute the relative entropy.

$$\begin{aligned} S(\rho|\sigma) &= \mathrm{Tr}(\rho\log\rho) - \mathrm{Tr}(\rho\log\sigma) \\ &= \mathrm{Tr}\begin{pmatrix} b\log b & 0 \\ 0 & (1-b)\log(1-b) \end{pmatrix} - \mathrm{Tr}\begin{pmatrix} b\log a & 0 \\ 0 & (1-b)\log(1-a) \end{pmatrix} \\ &= b\log\left(\frac{b}{a}\right) + (1-b)\log\left(\frac{1-b}{1-a}\right). \end{aligned} \tag{A.3}$$

The first derivatives are easily calculated to be

$$\begin{aligned} \frac{\partial S}{\partial a} &= \frac{a-b}{a(1-a)} \\ \frac{\partial S}{\partial b} &= \log\left(\frac{b}{1-b}\frac{1-a}{a}\right), \end{aligned} \tag{A.4}$$

which both vanish if and only if $a = b$. At $a = b$ the relative entropy $S(\rho|\sigma) = 0$ for all values of a. If we can show that these extremal points are minima, they must be global minima which proves the relative entropy to be a non-negative and monotonically increasing function of the state ρ. We will consider the second derivatives of the relative entropy. They are given by

$$\begin{aligned} \frac{\partial^2 S}{\partial a^2} &= \frac{b(1-a)^2 + (1-b)a^2}{a^2(1-a)^2} \\ \frac{\partial^2 S}{\partial b^2} &= \frac{1}{b(1-b)}. \end{aligned} \tag{A.5}$$

The second derivatives are positive for all $0 < a, b < 1$. This is enough to prove the points $a = b$ to be global minima, which proves the relative entropy is non-negative and monotonically increasing.

In this derivation, we have excluded the possibility that either σ or ρ describes a pure state, $a, b = 0, 1$. For these states the relative entropy will be ill-defined because we cannot take the logarithm of a non-invertible matrix. When a and b are both 0 or both 1, the relative entropy will vanish since σ and ρ will be equal. When $a = 0$ and $b = 1$ or vice versa, the relative entropy diverges. This is due to the fact that $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ are independent. Therefore, when comparing these states, they are "infinitely different". Note that this does not contradict our claim in Chapter 2 that the relative entropy is finite. The divergences we discussed there are UV divergences related to the continuous nature of a field theory, which will indeed cancel.

Appendix B

Derivation of modular flow in $\mathcal{D}$

We want to translate the flow

$$X^{\pm}(s) = X^{\pm} e^{\pm 2\pi s}, \tag{B.1}$$

using the transformation

$$x^{\mu} = \frac{X^{\mu} - (X \cdot X) C^{\mu}}{1 - 2(X \cdot C) + (X \cdot X)(C \cdot C)} + 2R^2 C^{\mu}. \tag{B.2}$$

with $C^{\mu} = (0, -1/2R, 0, \ldots)$. We will start by rewriting the transformation. First note that the modular flow in $\mathcal{R}$ only affects the $X^{\pm}$ coordinates. Therefore, we will just consider these directions. One easily calculates $C^{\pm} = C^1 \pm C^0 = -1/2R$, $(C \cdot C) = 1/4R^2$. Assume for now that there are two dimensions. This yields

$$\begin{aligned} x^{\pm} &= \frac{X^{\pm} + \frac{X^+ X^-}{2R}}{1 + \frac{(X^+ + X^-)}{2R} + \frac{X^+ X^-}{4R^2}} - R \\ &= \frac{X^{\pm}(1 + \frac{X^{\mp}}{2R})}{(1 + \frac{X^+}{2R})(1 + \frac{X^-}{2R})} - R \\ &= R\frac{2X^{\pm}}{(2R + X^{\pm})} - R \\ &= R\frac{(X^{\pm} - 2R)}{(X^{\pm} + 2R)}. \end{aligned} \tag{B.3}$$

We will also need the inverse of this relation, which is given by

$$X^{\pm} = -2R\frac{(x^{\pm} + R)}{(x^{\pm} - R)}. \tag{B.4}$$

The flow in $\mathcal{D}$ is given by

$$
\begin{aligned}
x^{\pm}(s) &= R\frac{(X^{\pm}(s) - 2R)}{(X^{\pm}(s) + 2R)} \\
&= R\frac{(X^{\pm}e^{\pm 2\pi s} - 2R)}{(X^{\pm}e^{\pm 2\pi s} + 2R)} \\
&= R\frac{(-2R\frac{(x^{\pm}+R)}{(x^{\pm}-R)}e^{\pm 2\pi s} - 2R)}{(-2R\frac{(x^{\pm}+R)}{(x^{\pm}-R)}e^{\pm 2\pi s} + 2R)} \\
&= R\frac{(x^{\pm} + R)e^{\pm 2\pi s} + (x^{\pm} - R))}{((x^{\pm} + R)e^{\pm 2\pi s} - (x^{\pm} - R))} \\
&= R\frac{(R + x^{\pm}) - e^{\mp 2\pi s}(R - x^{\pm})}{(R + x^{\pm}) + e^{\mp 2\pi s}(R - x^{\pm})} .
\end{aligned}
\tag{B.5}
$$

Appendix C

Derivation of H_B

Our goal is to construct an explicit expression for the modular Hamiltonian. Consider the surface $x^0 = 0$. H_B will induce some infinitesimal shift δs away from this surface. Recall the expression for the flow in $\mathcal{D}$

$$x^{\pm}(s) = R\frac{(R + x^{\pm}) - e^{\mp 2\pi s}(R - x^{\pm})}{(R + x^{\pm}) + e^{\mp 2\pi s}(R - x^{\pm})}. \tag{C.1}$$

We want see how the shift δs influences our coordinates x^{μ}. Express $x^0(s)$ at $x^0 = 0$ in terms of r and s as

$$\begin{aligned} x^0(s) &= \frac{1}{2}(x^+(s) - x^-(s)) \\ &= \frac{1}{2}R\left(\frac{(R+r) - e^{-2\pi s}(R-r)}{(R+r) + e^{-2\pi s}(R-r)} - \frac{(R+r) - e^{2\pi s}(R-r)}{(R+r) + e^{2\pi s}(R-r)}\right). \end{aligned} \tag{C.2}$$

Consider an infinitesimal shift δs away from our surface. This yields

$$\begin{aligned} \delta x^0(s) &= \frac{d}{ds}x^0(s)|_{s=0}\ \delta s \\ &= \frac{1}{2}R \cdot 2\pi(R-r)\left(\frac{1}{(R+r)+(R-r)} + \frac{(R+r)-(R-r)}{((R+r)+(R-r))^2}\right. \\ &\qquad\qquad \left. + \frac{1}{(R+r)+(R-r)} + \frac{(R+r)-(R-r)}{((R+r)+(R-r))^2}\right)\delta s \\ &= R \cdot 2\pi(R-r)\ \left(\frac{1}{(R+r)+(R-r)} + \frac{(R+r)-(R-r)}{((R+r)+(R-r))^2}\right)\delta s \\ &= R \cdot 2\pi(R-r)\ \left(\frac{2R}{4R^2} + \frac{2r}{4R^2}\right)\delta s \\ &= 2\pi\frac{R^2 - r^2}{2R}\delta s \end{aligned}$$

Now repeat the same derivation for $x^1(s) = \frac{1}{2}(x^+(s) + x^-(s))$. Due to the plus sign, the terms cancel after taking the derivative and setting $s = 0$. Therefore, we find that $\delta r = 0$.

We conclude that the modular Hamilion induces an infinitesimal flow away from $x^0 = 0$. We can identify the corresponding operator in the CFT as

$$H_B = 2\pi\int_B d^{d-1}x\frac{R^2 - r^2}{2R}T^{00}(x). \tag{C.3}$$

Appendix D

The off shell Noether current form

In this appendix we will show that $\boldsymbol{J}_X = d\boldsymbol{Q}_X + \boldsymbol{C}_X$ as claimed in Chapter 5 and derive the explicit form of $\boldsymbol{C}_X$. We will follow the argument presented in Appendix B of [7].

We start with an arbitrary gravitational Lagrangian in $(d+1)$-dimensions, which is a function of a collection of fields ϕ. Any proper gravitational theory is diffeomorphism invariant. Under a diffeomorphism generated by vector X, the variation of the corresponding action is given by

$$\delta_X S = \int \boldsymbol{\epsilon} E_\phi \delta_X \phi = \int \boldsymbol{\epsilon} E_\phi \mathcal{L}_X \phi. \tag{D.1}$$

where ϵ denotes the volume form and E_ϕ the equations of motion. The boundary terms do not contribute since ϕ vanishes at infinity. In order for the invariance to hold, the integrand must vanish. We will now work out how this constrains the fields. The Lie derivative of a tensor T of rank (i,j) is given by

$$\begin{aligned}(\mathcal{L}_X T)^{a_1\ldots a_i}{}_{b_1\ldots b_j} =& X^c(\nabla_c T^{a_1\ldots a_i}{}_{b_1\ldots b_j}) \\ &- \sum_{k=1}^{i} (\nabla_c X^{a_k}) T^{a_1\ldots c\ldots a_i}{}_{b_1\ldots b_j} \\ &+ \sum_{l=1}^{j} (\nabla_{b_l} X^c) T^{a_1\ldots a_i}{}_{b_1\ldots c\ldots b_j}\end{aligned} \tag{D.2}$$

To prevent us from complicating the derivation with lots of indices, we consider one rank-$(1,0)$ field such that $\phi = \phi^a$. In this case, the Lie derivative of our field just has two terms. The integrand is given by

$$\begin{aligned}\boldsymbol{\epsilon}(E_\phi)_a \mathcal{L}_X \phi^a &= \boldsymbol{\epsilon}(E_\phi)_a [X^b(\nabla_b \phi^a) - (\nabla_b X^a)\phi^b] \\ &= \boldsymbol{\epsilon} X^b (E_\phi)_a (\nabla_b \phi^a) + \boldsymbol{\epsilon} X^b \nabla_a((E_\phi)_b \phi^a) - \boldsymbol{\epsilon} \nabla_a (X^b (E_\phi)_b \phi^a)\end{aligned} \tag{D.3}$$

where we used Leibniz rule in the second line. The third term is a total derivative so it vanishes when we integrate. In order for the Lagrangian to be diffeomorphism invariant we must have

$$\boldsymbol{\epsilon} X^b (E_\phi)_a (\nabla_b \phi^a) + \boldsymbol{\epsilon} X^b \nabla_a((E_\phi)_b \phi^a) = 0. \tag{D.4}$$

Plugging this result back into the previous equation, we derive

$$\boldsymbol{\epsilon}(E_\phi)_a \mathcal{L}_X \phi^a = -\boldsymbol{\epsilon}\nabla_a (X^b (E_\phi)_b \phi^a). \tag{D.5}$$

We will now use this result to rewrite the Noether current. Recall that it is was given off shell by equation (5.7). Combining with the results in this appendix we find

$$\begin{aligned} d\boldsymbol{J}_X &= -\boldsymbol{E}_\phi \mathcal{L}_X \phi \\ &= -\boldsymbol{\epsilon}(E_\phi)_a \mathcal{L}_X \phi^a \\ &= \boldsymbol{\epsilon}\nabla_a (X^b (E_\phi)_b \phi^a) \\ &= d(X^b (E_\phi)_b \phi^a \boldsymbol{\epsilon}_a) \\ &\equiv d\boldsymbol{C}_X \end{aligned} \tag{D.6}$$

where $\boldsymbol{\epsilon}_a = \frac{1}{d!}\epsilon_{ab_2\ldots b_{d+1}} dx^{b_2} \wedge \ldots \wedge dx^{b_{d+1}}$. From this expression it follows that

$$\boldsymbol{J}_X = d\boldsymbol{Q}_X + \boldsymbol{C}_X. \tag{D.7}$$

This expression now defines the Noether current off shell.

The derivation of $\boldsymbol{C}_X$ can easily be extended to tensors of arbitrary rank. For our purposes, it is useful to derive what $\boldsymbol{C}_X$ looks like in GR. In this case we would have one rank-$(0,2)$ field, namely the metric. By comparing with our result for ϕ^a we can write it down immediately

$$\boldsymbol{C}_X = 2X^a (E_g)_{ab} g^{bc} \boldsymbol{\epsilon}_c = 2X^a (E_g)_{ab} \boldsymbol{\epsilon}^b. \tag{D.8}$$

The factor two comes from the fact that the metric has two indices. Therefore there are two terms from the Lie derivative, which add up since the metric is symmetric. There is no extra minus sign as the convention is to vary with respect to the inverse metric, which has raised indices.

References

[1] Eunseok Oh, IY Park, and Sang-Jin Sin. Full einstein from entanglement first law and sewing the space with entanglement flux line. *arXiv preprint arXiv:1709.05752*, 2017.

[2] Jacob D Bekenstein. Black holes and entropy. *Physical Review D*, 7(8):2333, 1973.

[3] Jacob D Bekenstein. Generalized second law of thermodynamics in black-hole physics. *Physical Review D*, 9(12):3292, 1974.

[4] Stephen W Hawking. Particle creation by black holes. *Communications in mathematical physics*, 43(3):199–220, 1975.

[5] Juan Maldacena. The large-n limit of superconformal field theories and supergravity. *International journal of theoretical physics*, 38(4):1113–1133, 1999.

[6] Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from the anti–de sitter space/conformal field theory correspondence. *Physical review letters*, 96(18):181602, 2006.

[7] Thomas Faulkner, Monica Guica, Thomas Hartman, Robert C Myers, and Mark Van Raamsdonk. Gravitation from entanglement in holographic cfts. *Journal of High Energy Physics*, 2014(3):1–41, 2014.

[8] Rudolf Haag. *Local quantum physics: Fields, particles, algebras*. Springer Science & Business Media, 2012.

[9] Huzihiro Araki. Relative entropy of states of von neumann algebras. *Publications of the Research Institute for Mathematical Sciences*, 11(3):809–833, 1976.

[10] Horacio Casini, Marina Huerta, and Robert C Myers. Towards a derivation of holographic entanglement entropy. *Journal of High Energy Physics*, 2011(5):36, 2011.

[11] Robert M Wald. Black hole entropy is the noether charge. *Physical Review D*, 48(8):R3427, 1993.

[12] Stefan Hollands and Robert M Wald. Stability of black holes and black branes. *Communications in Mathematical Physics*, 321(3):629–680, 2013.

[13] Joseph J Bisognano and Eyvind H Wichmann. On the duality condition for quantum fields. *Journal of mathematical physics*, 17(3):303–321, 1976.

[14] William G Unruh. Notes on black-hole evaporation. *Physical Review D*, 14(4):870, 1976.

[15] Mukund Rangamani and Tadashi Takayanagi. Holographic entanglement entropy. In *Holographic Entanglement Entropy*, pages 35–47. Springer, 2017.

[16] Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 2013(8):90, 2013.

[17] Xi Dong, Aitor Lewkowycz, and Mukund Rangamani. Deriving covariant holographic entanglement. *Journal of High Energy Physics*, 2016(11):28, 2016.

[18] Roberto Emparan. Ads/cft duals of topological black holes and the entropy of zero-energy states. *Journal of High Energy Physics*, 1999(06):036, 1999.

[19] Vijay Balasubramanian and Per Kraus. A stress tensor for anti-de sitter gravity. *Communications in Mathematical Physics*, 208(2):413–428, 1999.

[20] J David Brown and James W York Jr. Quasilocal energy and conserved charges derived from the gravitational action. *Physical Review D*, 47(4):1407, 1993.

[21] Vivek Iyer and Robert M Wald. Comparison of the noether charge and euclidean methods for computing the entropy of stationary black holes. *Physical Review D*, 52(8):4430, 1995.

[22] Joohan Lee and Robert M Wald. Local symmetries and constraints. *Journal of Mathematical Physics*, 31(3):725–743, 1990.

[23] Robert M Wald and Andreas Zoupas. General definition of "conserved quantities" in general relativity and other theories of gravity. *Physical Review D*, 61(8):084027, 2000.

[24] Vivek Iyer and Robert M Wald. Some properties of the noether charge and a proposal for dynamical black hole entropy. *Physical review D*, 50(2):846, 1994.

[25] Nima Lashkari and Mark Van Raamsdonk. Canonical energy is quantum fisher information. *Journal of High Energy Physics*, 2016(4):153, 2016.

[26] Nima Lashkari, Michael B McDermott, and Mark Van Raamsdonk. Gravitational dynamics from entanglement "thermodynamics". *Journal of High Energy Physics*, 2014(4):195, 2014.

[27] Robert M Wald. *General relativity.* University of Chicago press, 2010.

[28] Thomas Faulkner, Felix M Haehl, Eliot Hijano, Onkar Parrikar, Charles Rabideau, and Mark Van Raamsdonk. Nonlinear gravity from entanglement in conformal field theories. *Journal of High Energy Physics*, 2017(8):57, 2017.

[29] Ted Jacobson. Entanglement equilibrium and the einstein equation. *Physical review letters*, 116(20):201101, 2016.

[30] Ted Jacobson. Thermodynamics of spacetime: the einstein equation of state. *Physical Review Letters*, 75(7):1260, 1995.

[31] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *Journal of High Energy Physics*, 2013(11):74, 2013.

[32] Daniel L Jafferis, Aitor Lewkowycz, Juan Maldacena, and S Josephine Suh. Relative entropy equals bulk relative entropy. *Journal of High Energy Physics*, 2016(6):4, 2016.